



19th
Amsterdam Colloquium
2013

Proceedings of the 19th Amsterdam Colloquium

Edited by Maria Aloni, Michael Franke and Floris Roelofsen

Foreword

This is a collection of papers presented at the 19th Amsterdam Colloquium, organized by the Institute for Logic, Language, and Computation (ILLC) at the University of Amsterdam, December 18–20, 2013. The bi-annual Amsterdam Colloquia aim at bringing together linguists, philosophers, logicians, cognitive scientists and computer scientists who share an interest in the formal study of the semantics and pragmatics of natural and formal languages.

Besides the regular programme, the 2013 edition featured two workshops on Quantitative Data and Modals, respectively, a special session on the Semantics and Pragmatics of Dialogue, jointly organized with SemDial, one evening lecture, jointly organized with the E.W. Beth Foundation, a special event in honor of Jeroen Groenendijk, Martin Stokhof, and Frank Veltman, with plenary addresses by Johan van Benthem, Barbara Partee and Hans Kamp, and a round-table debate on the Future of Semantics, jointly organized with the SMART Cognitive Science lecture series. The programme included 10 invited talks and 32 contributed talks.

We would like to thank the members of the programme committee and all the reviewers, listed below, for their efforts in selecting the contributed talks. We would also like to thank Peter van Ormondt, Ben Rodenhäuser, Ivano Ciardelli, Michele Herbstritt, Masa Mocnik and Ben Sparkes for their help in organizing the colloquium.

Finally, we would like to thank the ILLC, the E.W. Beth Foundation, and the City Council of Amsterdam for providing financial support.

Maria Aloni
Michael Franke
Floris Roelofsen

Programme Committee

General programme

Robert van Rooij (chair)	ILLC, University of Amsterdam
Paul Dekker	ILLC, University of Amsterdam
Rick Nouwen	Utrecht University

Workshop on Quantitative Data

Michael Franke	ILLC, University of Amsterdam
----------------	-------------------------------

Workshop on Modals

Ben Rodenhäuser	ILLC, University of Amsterdam
Frank Veltman	ILLC, University of Amsterdam
Hedde Zeijlstra	Universität Göttingen

Reviewers

Dorit Abusch	Cornell University
Luis Alonso-Ovalle	McGill University
Scott Anderbois	Brown University
Anton Benz	ZAS Berlin
Adrian Brasoveanu	University of California Santa Cruz
Richard Breheny	University College London
Seth Cable	University of Massachusetts
Ivano Capomigro	University of California, San Diego
Lucas Champollion	New York University
Emmanuel Chemla	LSCP Paris
Gennaro Chierchia	Harvard University
Charles Clifton	University of Massachusetts Amherst
Ariel Cohen	Ben-Gurion University
Chris Cummins	Universität Bielefeld
Henriëtte De Swart	Utrecht University
Judith Degen	University of Rochester
Paul Dekker	ILLC, University of Amsterdam
Donka Farkas	University of California Santa Cruz
Kit Fine	New York University
Lyn Frazier	University of Massachusetts Amherst
Jon Gajewski	University of Connecticut
Bart Geurts	Radboud University Nijmegen
Jeroen Groenendijk	ILLC, University of Amsterdam
Atle Grønn	University of Oslo
Valentine Haccuard	University of Maryland
Jeff Horty	University of Maryland
Julie Hunter	Institut Jean-Nicod & EHESS, Paris

Gerhard Jaeger	University of Tuebingen
Napoleon Katsos	University of Cambridge, RCEAL
Magdalena Kaufmann	University of Connecticut
Stefan Kaufmann	University of Connecticut
Nathan Klinedinst	University College London
Angelika Kratzer	University of Massachusetts Amherst
Fred Landman	Tel Aviv University
Daniel Lassiter	Stanford University
Roger Levy	University of California San Diego
Louise McNally	Universitat Pompeu Fabra
Paul Portner	Georgetown University
Chris Potts	Stanford University
Jessica Rett	UCLA
Susan Rothstein	Bar-Ilan University
Uli Sauerland	ZAS Berlin
Hans-Christian Schmitz	IDS Mannheim
Roger Schwarzschild	Rutgers University
Chungchieh Shan	Indiana University
Yael Sharvit	UCLA
Stephanie Solt	ZAS Berlin
Benjamin Spector	Institut Jean-Nicod, CNRS, ENS, EHESS Paris
Martin Stokhof	ILLC, University of Amsterdam
Matthew Stone	Rutgers University
Kristen Syrett	Rutgers University
Anna Szabolcsi	New York University
Jakub Szymanik	ILLC, University of Amsterdam
Kjell Sæbø	University of Oslo
Frank Veltman	ILLC, University of Amsterdam
Kai Von Fintel	MIT
Klaus Von Heusinger	Universität zu Köln
Galit W. Sassoon	Bar Ilan University
Yoad Winter	Utrecht University
Henk Zeevat	ILLC, University of Amsterdam
Hedde Zeijlstra	Georg August University Goettingen
Ede Zimmermann	Goethe Universität Frankfurt

Contents

Invited

What a Rational Interpreter Would Do: Building, Ranking, and Updating Quantifier Scope Representations in Discourse.....	1
<i>Adrian Brasoveanu and Jakub Dotlačil</i>	
On the negation of indicative conditionals.....	10
<i>Paul Egré and Guy Politzer</i>	
The grammatical category of modality	19
<i>Valentine Hacquard</i>	
Quantifier Particles and Compositionality	27
<i>Anna Szabolcsi</i>	

Contributed

Epistemic Indefinites: Are We Ignorant About Ignorance?	35
<i>Luis Alonso-Ovalle and Paula Menendez-Benito</i>	
Correlating cessation with double access.....	43
<i>Daniel Altshuler and Roger Schwarzschild</i>	
Modification of DPs by epistemic modal adverbs	51
<i>Elizabeth Bogal-Allbritten</i>	
Witness Sets, Polarity Reversal and the Processing of Quantified Sentences.....	59
<i>Oliver Bott, Udo Klein and Fabian Schlotterbeck</i>	
Universal quantification as iterated conjunction	67
<i>Dylan Bumford</i>	
'Than' = 'More' + Exhaustivity: Evidence from Circassian.....	75
<i>Lisa Bylinina and Yuri Lander</i>	
Man and Woman: the Last Obstacle for Boolean Coordination	83
<i>Lucas Champollion</i>	
On the meaning of Intensifiers.....	91
<i>Harris Constantinou</i>	
Most: the View from Mass Quantification	99
<i>Carmen Dobrovie-Sorin</i>	
Dowty's aspect hypothesis segmented	107
<i>Tim Fernando</i>	
A two dimensional analysis of the future: modal adverbs and speakers bias	115
<i>Anastasia Giannakidou and Alda Mari</i>	
No ordered arguments needed for nouns	123
<i>Scott Grimm and Louise McNally</i>	

The Intervention effect: Focus alternatives or indefinite alternatives? Experimental evidence	131
<i>Andreas Haida and Sophie Repp</i>	
Monadic Quantifiers Recognized by Deterministic Pushdown Automata.....	139
<i>Makoto Kanazawa</i>	
An experimental investigation of interrogative syntax/semantics	147
<i>Hadas Kotek and Martin Hackl</i>	
Tense and Aspect in Swing Conditionals	155
<i>Fabienne Martin</i>	
An Analysis of Quantifier Scope Restrictions in Dependence Logic.....	163
<i>Ralf Naumann and Wiebke Petersen</i>	
Attaching NRCs to Plural Quantificational Heads	171
<i>Claudia Poschmann</i>	
A Fregean Semantics for Number Words.....	179
<i>Susan Rothstein</i>	
<i>At least</i> and Quantity Implicature: Choices and Consequences.....	187
<i>Bernhard Schwarz</i>	
Semantics of DP islands.....	195
<i>Alexandra Simonenko</i>	
Towards a Formal Theory of Explanatory Biases in Discourse	203
<i>Torgrim Solstad and Oliver Bott</i>	
Epistemic Modals, Qualitative Probability, and Nonstandard Probability	211
<i>Satoru Suzuki</i>	
Monotonicity has only a relative effect on the complexity of quantifier verification	219
<i>Jakub Szymanik and Marcin Zajenkowski</i>	
Focus association in superlatives and the semantics of <i>-est</i>	226
<i>Barbara Tomaszewicz</i>	
The type of adjectives.....	234
<i>Galit Weidman Sassoon</i>	
Distributivity and agreement: new evidence for groups as sets	241
<i>Hanna De Vries</i>	
Indicative Scorekeeping	249
<i>Malte Willer</i>	
Deontic scope restrictions beyond polarity.....	257
<i>Igor Yanovich</i>	
Questioning and Asserting at the same time: the L% tone in A-not-A questions	265
<i>Mengxi Yuan and Yurie Hara</i>	
Universal Quantifier PPIs	273
<i>Hedde Zeijlstra</i>	

What a Rational Interpreter Would Do: Building, Ranking, and Updating Quantifier Scope Representations in Discourse

Adrian Brasoveanu and Jakub Dotlačil*

¹ UC Santa Cruz, abrsvn@gmail.com

² Utrecht University/University of Groningen, j.dotlacil@gmail.com

Abstract

We frame the general problem of ‘rationally’ (in the sense of Anderson et al’s ACT-R framework) integrating semantic theories and processing, and indicate how this integrated theory could be explicitly formalized; an explicit formalization enables us to empirically evaluate semantic and processing theories both qualitatively and quantitatively. We then introduce the problem of quantifier scope, the processing difficulty of inverse scope, and two types of theories of scope, and discuss the results of a self-paced reading experiment and its consequences for these two types of theories. Finally, we outline how probabilities for LF construction rules could be computed based on the experimental results.

1 Introduction: ‘Rational’ theories of cognition

Anderson (1990) and much subsequent work argues for the following ‘rational cognition’ hypothesis (a.k.a. general principle of rationality): the cognitive system operates at all times to optimize the adaptation of the behavior of the organism. ‘Rationality’ is not used here in the sense of engaging in logically correct reasoning when deciding what to do. It is used in the sense of ‘adaptation’: human behavior is optimal in terms of achieving human goals. A ‘rational’, as opposed to ‘mechanistic’, approach to cognition is closely related to aiming for explanatory adequacy in addition to descriptive adequacy.

Developing a theory along the lines of the rational cognition hypothesis requires one to follow the six steps discussed in Anderson (1990: 29-30): **(1)** begin by precisely specify the goals of the cognitive system; **(2)** develop a formal model of the environment to which the system is adapted; **(3)** make minimal assumptions about computational limitations; **(4)** derive the optimal behavioral function given steps **1-3**; **(5)** examine the empirical literature to see if the predictions of the behavioral function are confirmed (if available; else do the empirical investigation); **(6)** finally, if the predictions are off, iterate. The theoretical commitments are made in steps **1-3**. They provide the “framing of the information-processing problem”. Steps **4-5** are about deriving and dis/confirming predictions. Finally, theory building is iterative: if one framing does not work, we try another.

Our goal in this paper is to get started with the first iteration of our rational analysis for a classical problem in formal semantics: *quantifier scope ambiguities*. In particular, we will study how interpreters deal with scope ambiguities during actual comprehension. The specific questions we are interested in are as follows. **(Q1)** How are quantifier scope ambiguities *represented*

*We want to thank Pranav Anand, Nate Arnett, Amy Rose Deal, Donka Farkas, John Hale, Roger Levy, Anna Szabolcsi, Matt Wagers and the UCSC S-Circle audience (Nov. 15, 2013). Adrian Brasoveanu was supported by a UCSC CoR SRG grant for part of this research. Jakub Dotlačil was supported by a Rubicon grant from the Netherlands Organization for Scientific Research for part of this research. The usual disclaimers apply.

by the interpreter? **(Q2)** How are these representations *built and maintained/updated* as the discourse is incrementally processed/interpreted? **(Q3)** Finally, how are these representations *ranked* so that the ambiguities are *resolved*? But what would it mean to provide a rational analysis for the problem of processing quantifier scope ambiguities? Paraphrasing the title of Hale (2011): what would a rational interpreter do?

In §2, we introduce the problem of quantifier scope and the difficulty of inverse scope, and we describe the results of a self-paced reading experiment targeting questions **Q1-Q3** above. In §3, we pick up the ‘rational’ analysis thread again and frame our information-processing problem, i.e., the parsing/interpretation problem, in detail. The main payoff of the detailed ‘framing’ is a much clearer understanding of the relation between semantic theories and the processor, so clear that explicit formalization of the connection between semantic theory and processing, as well as ways to do quantitative empirical evaluation, will be within reach. Finally, we will briefly outline how probabilities for LF construction rules could be computed.

2 Experimental investigations of quantifier scope

Consider the sentence in (1) below. The surface scope interpretation of this sentence is that there is a boy that lifted every box (the same boy lifted all of them); the inverse scope interpretation is that every box is such that a boy lifted it (a possibly different boy for each box).

- (1) A boy lifted every box.

A working definition of inverse scope that will suffice for this paper is that the interpretation of a quantifier is dependent on another quantifier that was introduced later (see Szabolcsi 1997 a. o. for a more precise definition). Importantly for us, inverse scope is costly relative to surface scope: it is harder to process (Pylkkänen and McElree 2006 and references therein). This is shown, for example, by the fact that a plural follow-up to (1) above, e.g., *The boys were looking for a marble* – which forces the inverse-scope reading – leads to increased reading times (RTs; Tunstall 1998 a.o.). The inverse scope interpretation is costly, hence the increase in RTs.

The previous literature leaves several issues open. Crucially, it focuses on sentences with only 2 quantifiers, as in (1) above. This might suffice to establish the cost of inverse scope readings but it doesn’t substantially help us understand how quantifier scope ambiguities are represented and maintained/updated by the interpreter. One could imagine at least two possibilities, which are often assumed in the literature: (a) the interpreter builds an LF representation that disambiguates scope readings; if the continuation is incompatible with it, the LF representation is revised accordingly (Pylkkänen and McElree 2006 and references therein); or (b) the interpreter builds a (mental/discourse) model structure, which is revised if the continuation is incompatible with it (Fodor 1982). One way to specify the model-based approach is to take indefinites to denote Skolem functions (or Skolemized choice functions) of variable arity (Steedman 2012): what gets revised then is the arity – and consequently the function.¹

We conducted two new experimental studies (eye-tracking and self-paced reading) to decide between these two possibilities. Here, we report only the self-paced reading experiment (see Dotlačil and Brasoveanu 2013 for the other experiment and details about the experimental designs of both experiments). The main novelty of the tasks: we examined the interaction of 3 quantifiers, 2 singular indefinites and 1 universal, in two-sentence discourses like (2) below:

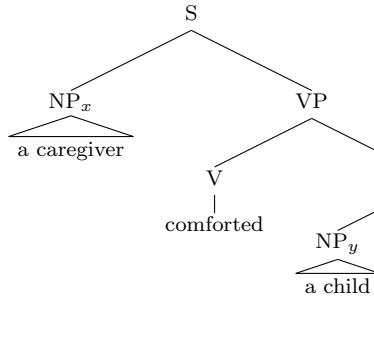
¹The interpreter could also operate with underspecified structures/models (Ebert 2005 and references therein), but these theories have no clear way to explain inverse scope difficulty unless something else is added, e.g., that specifying scope relations is sometimes forced (mid-sentence) and is at least sometimes costly, so we’ll set them aside. See Radó and Bott (2012) for an experimental investigation of underspecification theories.

- (2) a. A caregiver comforted a child every night.
 b. The $\{\begin{matrix} \text{caregiver} \\ \text{caregivers} \end{matrix}\}$ wanted the $\{\begin{matrix} \text{child} \\ \text{children} \end{matrix}\}$ to get some rest.

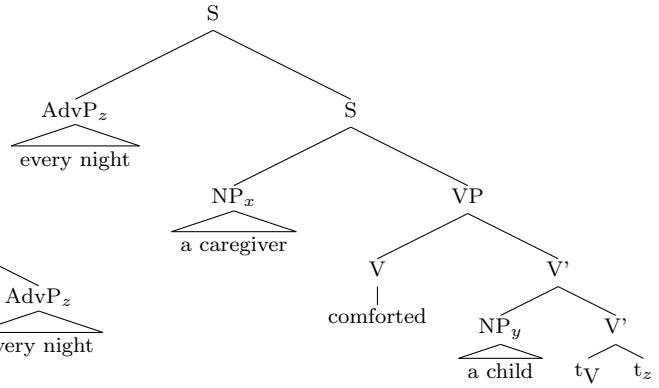
The first sentence has 2 indefinites in SU and DO position and a universal quantifier as a sentence-final adverb. The second sentence elaborates on the entities brought to salience by the 2 indefinites. The only manipulation (that is relevant for our purposes; see Dotlačil and Brasoveanu 2013 for a much more detailed discussion) is morphological number on the SU and DO definites in the second sentence (2×2 design): the idea is that *singular* definite \Rightarrow *wide-scope* indefinite,² while *plural* definite \Rightarrow *narrow-scope* indefinite. The two theories of (inverse) scope we outlined above make the following predictions for this type of examples.

- (3) Predictions of the covert LF operations theory: (a) assume a base-generated structure with the universal adverb in the lowest position (Frazier and Fodor 1978, Larson 1988); (b) assume that the more operations we need to apply to obtain an LF, the less plausible/salient it is (Frazier 1978); (c) then: narrow scope SU \Rightarrow narrow scope DO.

Wide scope SU, wide scope DO:

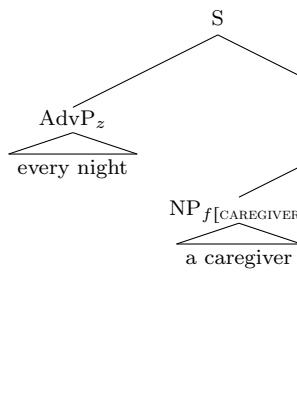


Narrow scope SU \Rightarrow narrow scope DO:

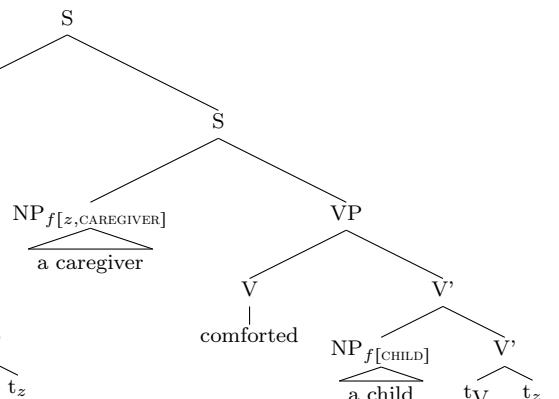


- (4) Predictions of the model revision theory: (a) assume that giving widest scope to the universal is costless, but setting the arities of the two Skolem functions is costly; (b) assume that the arities of the two Skolem functions are independently specified; (c) then: narrow scope SU \Rightarrow wide scope DO.

Wide scope SU, wide scope DO:



Narrow scope SU \Rightarrow wide scope DO:



²Not necessarily wide scope: maybe narrow with ‘accidental’ coreference; we ignore this complication here.

We note that for presentational clarity, we postulated a very specific, LF-based theory but any theory that assumes a *scope hierarchy* (a strict total order: asymmetric, total and transitive) that has the DO by default in the scope of the SU will predict that narrow scope SU \Rightarrow narrow scope DO. This prediction is not made by theories that directly operate on models since DO scope is then independent of SU scope (as exemplified above using Skolem functions).

The experiment examined two-sentence discourses like (2a-2b) (henceforth CONTEXT: YES), but also their one-sentence counterparts consisting only of the second sentence (2b) (henceforth CONTEXT: NO). The main finding relevant for us is that in the CONTEXT: YES condition, the narrow-scope reading of the SU or the DO led to increased RTs, but there was a clear facilitation (observable in decreased RTs) when a narrow-scope DO followed a narrow-scope SU. That is, the inverse scope of the universal over the SU makes it easier to also interpret the DO as taking narrow scope. This facilitation cannot be due to the repetition of two plural forms because there is no facilitation in the CONTEXT: NO condition (in fact, this condition showed a borderline-significant *slowdown* in the region following the object for SU:PL & DO:PL condition). Thus, PL on the SU facilitates PL on the DO *but only* when the PL disambiguates scope. So the facilitation is (likely) due to the disambiguation role played by PL morphology.

These results are *incompatible*: (a) with the assumption that readers do not use disambiguating information quickly to reanalyze scope, (b) with (discourse/mental) model based theories of inverse scope – to the extent these theories do not keep track of some basic remnant of a grammatical/thematic scope hierarchy, (c) with related theories of scope, e.g., theories that take indefinites to denote Skolem functions/Skolemized choice functions of variable arity, or underspecification theories of scope – again, to the extent that specifying the scope of the DO is independent of specifying the scope of the SU in these theories.

The results are *compatible*: (a) with the assumption that the reanalysis is done on scope representations that can be specified in terms of LF/grammatical/thematic/linear order hierarchies, and (b) more generally, with the assumption that the processor builds hierarchical scope representations and updates/maintains them across sentential boundaries. Because of this, the results favor dynamic systems that have rich interpretation contexts like DRT (Kamp and Reyle 1993) rather than ‘less representational’ systems like DPL (Groenendijk and Stokhof 1991).

3 Framing the parsing/interpretation problem

These experimental results and their consequences help us understand how the interpreter builds and maintains scope representations, but we might want to do better. Theoretically, we left the connection between semantic theories and processing implicit but our conclusions/generalizations relied on a fairly tight connection between semantic theory and processing – how else could we link *behavioral measurements* in the experimental task and the *mental representations* postulated by our semantic theories? We don’t need to make this connection formally explicit for the conclusions to be acceptable, but it would be good to do it for all the usual reasons. Empirically, we only focused on whether the RTs for the different conditions are different or not (while taking into account sampling error etc.), but *the relative magnitudes of the RTs contain additional information that we largely ignored*. They might tell us something about the relative likelihood of the different scope representations investigated in the experiment.

So let’s ‘frame’ our information-processing problem, i.e., the parsing/interpretation problem, in more detail. A rational analysis of this problem is a *minimal* formally explicit theory of parsing/interpretation: it explicitly tries to make minimal assumptions about processing mechanisms and syntactic/semantic theories. We start with some basic, and largely uncontroversial, assumptions about the human processor (Marslen-Wilson 1973, Frazier and Fodor

1978, Tanenhaus et al. 1995, Steedman 2001, Hale 2011 a.o.): the human processor (*a*) is incremental – syntactic parsing and semantic interpretation do not lag significantly behind the perception of individual words, (*b*) is predictive – the processor forms explicit representations of words and phrases that have not yet been heard, and (*c*) satisfies the competence hypothesis – understanding a sentence/discourse involves the recovery of the structural description of that sentence/discourse on the syntax side, and of the meaning representation on the semantic side.

We will now go through the first 3 steps of rational theory construction for parsing/interpretation (see Hale 2011: 403 et seqq). First, we take the goal of the parser/interpreter to be that it should rapidly arrive at the syntactic and meaning representation intended by the speaker. This goal weaves together two competing demands: *be quick* and *be accurate*. Given the competence hypothesis, we can formulate this as follows:

- (5) The goal of the parser/interpreter (step 1): search through the space of syntactic structures and meaning representations quickly (the end state is reached fast) and accurately (the end state is the interpretation intended by the speaker).

This formulation is an instance of a general approach (Newell and Simon 1972): cognition as problem solving, and problem solving as search through a state space.

We turn now to step 2: identifying a formal model of the environment to which the parser/interpreter is adapted. Since (*i*) sentence/discourse comprehension occurs in a speech community, and (*ii*) grammars describe the knowledge shared by the community, we take grammars to be models of the environment to which comprehenders are adapted (Hale 2011):

- (6) The formal model of the environment (step 2): the parser/interpreter is adapted to categorical and gradient information specified in *the grammars (syntax and semantics)* of particular languages.

This step says nothing about what counts as a grammar (a syntactic or a semantic theory), which theory is best, etc. But it provides a clear link between processing and grammar: this step and the competence hypothesis provide the two central assumptions we relied on when interpreting our experimental results.

We finally turn to step 3: computational limitations/specifications. Given a grammar (let's focus on syntax and semantics only), the parser/interpreter has to:

- (7) Computational limitations/specifications (step 3): the parser/interpreter has to
 - a. define a way of *applying* the syntax and semantics *rules*;
 - b. define a way of *resolving conflict* when more than one rule is applicable.

Conflicts should be resolved in such a way that the estimated distance to completion is minimized (be quick) and the estimated correctness of the analysis is maximized (be accurate).

What does it mean to apply a rule? As already indicated, we take parsing/interpretation to be search through a state space. For syntax, a state is a partially completed syntactic structure (Hale 2011 and references therein). For semantics, a state is a partially constructed DRS (more broadly, LF) and/or a partially evaluated DRS/LF. Applying a grammar rule takes us from one state to another (strong competence hypothesis): *rule applications are transition/accessibility relations between states*. For syntax, we apply phrase structure rules. For semantics, we can take transitions to consist of (*i*) applying a DRS/LF construction rule and/or (*ii*) evaluating a sub-DRS/sub-LF and updating the current interpretation context in the process.

How do we resolve conflict to minimize distance to completion and maximize accuracy? Trying to maximize accuracy is very hard because we can't guess what the speaker intends to

say in the future. That is, it is hard to define heuristic values to maximize accuracy (Hale 2011): an analysis for the first few words may be very good if they're followed by one continuation, very bad if followed by another. So let's focus conflict resolution on minimizing distance to completion: assume that the current partial analysis is right; now let's choose between two paths of analysis. We can estimate how far we are from completion based on previous experience, i.e., based on analyses that we completed before and that have the same initial subpart.

But how do we estimate distance to completion? For syntax, we can do it empirically: we can use a treebank, simulate the actions of a given parser (e.g., left-corner) for the sentences in the treebank and record how far particular intermediate states are from the correct end state. We can use those average distances to resolve conflict: select the analysis path with the smallest expected distance to completion. Hale (2011) uses *A* search*, which is *best-first* – try the best path first, keep a priority queue of alternates, and *informed* – it uses problem-specific knowledge (heuristic values) rather than a fixed policy (e.g., breadth first, depth first). The heuristic value at a state has 2 components: how far we traveled from the initial state + estimated distance to the goal; using both components minimizes *overall* path length.

But the empirical way is not really possible for semantics (yet). So let's look at alternatives. An alternative way to compute heuristic values for the syntactic processor is to assume that our phrase structure rules are weighted (probabilistic grammars) and to derive expected distances to the end state based on those weights. The basic idea: the more uncertain an analysis path is, the more likely that path is to be far from the end state. Uncertainty is based on the weights themselves, but also on how many choices we have at a particular point. For example, big/complex phrases are more ‘uncertain’, hence avoided, because they can be expanded in many ways – and the more alternatives, the longer it takes to disconfirm the incorrect ones.

The exact procedure is less important, see Hale (2011: 430-432) for more details and a syntactic application. But the main moral for semantics is that estimating probabilities for DRS/LF construction and/or evaluation rules enables our semantic theories to make (more) precise predictions about processing. Thus, our proposal is as follows: (i) we can estimate probabilities experimentally based on RTs; (ii) once we estimate probabilities from one experiment, we can derive predictions for another;³ (iii) we can then evaluate these predictions with respect to their overall qualitative pattern, but we can also quantitatively evaluate them; (iv) things will probably not work out the first time around, so we go to step 6: iterate.

Here's how estimating probabilities could go. Take a simple two-sentence discourse with 2 quantifiers in the first sentence: *A boy climbed every tree. The boy/boys wanted to catch blue jays.* And suppose we measure the RTs on the word *wanted*. Assume (following Hale 2001 and Levy 2008) that the RTs vary according to how unexpected/surprising the SG *boy* is relative to the PL *boys*. In particular, assume that the difference in difficulty between SG, i.e., *SS* (surface scope), and PL, i.e., *IS* (inverse scope) is proportional to the difference between the surprisal of *SS*, i.e., $-\log(\Pr(SS))$, and the surprisal of *IS*, i.e., $-\log(\Pr(IS))$. Let's make this precise by taking the difficulty of SG/PL to be measured in log(RTs), with RTs measured in ms.:

$$(8) \quad \begin{aligned} \log(\mathbf{RT}(SS)) - \log(\mathbf{RT}(IS)) &\propto (-\log(\Pr(SS))) - (-\log(\Pr(IS))), \text{ i.e.,} \\ \log(\mathbf{RT}(SS)) - \log(\mathbf{RT}(IS)) &= c \cdot [\log(\Pr(IS)) - \log(\Pr(SS))], \text{ hence:} \\ \log\left(\frac{\mathbf{RT}(SS)}{\mathbf{RT}(IS)}\right) &= \log\left(\left(\frac{\Pr(IS)}{\Pr(SS)}\right)^c\right), \text{ i.e., } \frac{\mathbf{RT}(SS)}{\mathbf{RT}(IS)} = \left(\frac{\Pr(SS)}{\Pr(IS)}\right)^{-c} \text{ (where } c > 0\text{)} \end{aligned}$$

That is, RTs and probabilities are inversely related: the higher the probability of *SS* is relative to *IS*, the shorter the RTs for *SS* relative to *IS* because *SS* is less surprising/more

³Using probabilities does not necessarily mean that we commit to the fact that they are part of mental representations. They are useful theoretical constructs, just like possible worlds are in formal semantics.

predictable. The free parameter c allows for a flexible relation between ratios of RTs and odds (ratios of probabilities) and should be estimated from the data. Now let's take the RTs from the CONTEXT:YES condition of the self-paced reading experiment and estimate probabilities. We estimate 6 probabilities, 2 for the SU and 4 for the DO: $\Pr(SU = SS)$ (*caregiver*) – the probability that the SU takes wide scope (call it *SS* for uniformity) relative to the universal; $\Pr(SU = IS)$ (*caregivers*) – the probability that the SU takes narrow scope (call it *IS* for uniformity) relative to the universal; $\Pr(DO = SS|SU = SS)$ (*child|caregiver*) – the probability that the DO takes wide scope given that the SU takes wide scope; $\Pr(DO = IS|SU = SS)$ (*children|caregiver*) – the probability that the DO takes narrow scope given that the SU takes wide scope; $\Pr(DO = SS|SU = IS)$ (*child|caregivers*) – the probability that the DO takes wide scope given that the SU takes narrow scope; $\Pr(DO = IS|SU = IS)$ (*children|caregivers*) – the probability that the DO takes narrow scope given that the SU takes narrow scope.

To keep things simple, we will sum the RTs for the relevant regions of interest and obtain one measurement for each of the 42 participants by averaging over items. A serviceable basic Bayesian model with low information priors to estimate these probabilities can be constructed as follows. The data y consists of 42 $\frac{\text{RT}(SS)}{\text{RT}(IS)}$ ratios (one per participant), and $y_i \sim \text{Gamma}(\alpha, \beta)$. *Gamma* is a convenient distribution to use because the RT ratios are always positive. We reparametrize it in terms of its mean μ and standard deviation σ so that we can link it to probability ratios: α (shape) = $\frac{\mu^2}{\sigma^2}$ and β (rate) = $\frac{\mu}{\sigma^2}$. The mean of the *Gamma* distribution is then specified as $\mu = \left(\frac{\Pr(SS)}{\Pr(IS)} \right)^{-c}$, and we assume a *Unif*(0.01, 10) prior for c . Furthermore, we assume a uniform *Beta*(1, 1) prior for $\Pr(SS)$ and take $\Pr(IS) = 1 - \Pr(SS)$. Finally, we assume an *IGamma*($10^{-3}, 10^{-3}$) prior for the variance σ^2 . We also add random effects for participants, not listed in the model above for simplicity.

These are the means of the posterior distributions estimated using this model:⁴

$$(9) \quad \begin{aligned} \Pr(SU = SS) &= 0.59 & \Pr(SU = IS) &= 0.41 \\ \Pr(DO = SS|SU = SS) &= 0.55 & \Pr(DO = SS|SU = IS) &= 0.51 \\ \Pr(DO = IS|SU = SS) &= 0.45 & \Pr(DO = IS|SU = IS) &= 0.49 \end{aligned}$$

We can now calculate joint probabilities, i.e., the probabilities of the 4 scope configurations for the initial sentence. In general, $\Pr(X, Y) = \Pr(X|Y) \cdot \Pr(Y)$, hence:

$$(10) \quad \begin{aligned} \Pr(SU = SS, DO = SS) &= 0.33 & \Pr(SU = IS, DO = SS) &= 0.21 \\ \Pr(SU = SS, DO = IS) &= 0.26 & \Pr(SU = IS, DO = IS) &= 0.20 \end{aligned}$$

We see that $SU = SS, DO = SS$ is about 6% more likely than $SU = SS, DO = IS$, which in turn is about 6% more likely than the two configurations in which $SU = IS$. It seems that every quantifier movement up the tree makes the resulting configuration 6% less likely.

There is basically no difference between the last two configurations $SU = IS, DO = SS$ and $SU = IS, DO = IS$. This unexpected result is due to the fact that we did not take into account the ‘baseline’ RTs provided by the CONTEXT:No condition. But note that taking CONTEXT:No into account would only make the probability of $SU = IS, DO = SS$ lower, *definitely not null*. In fact, our model assumed that $SU = IS, DO = SS$ is *a priori* possible: we did not build a probability of 0 for this configuration into the prior. This is right for the LF theory since we can imagine $SU = IS, DO = SS$ being derived from $SU = IS, DO = IS$ via an additional movement of the DO indefinite.

However, once we assume we have weights for LF rules that are reflected in RTs (because the heuristic values for the processor are derived from those weights), Skolem-function approaches

⁴We used R (R Core Team 2013) and JAGS (Plummer 2013) to estimate the posterior distributions.

and related approaches, e.g., Dependence Logic, become a viable option again. If covert LF operations are weighted, why not add weights to the arity specification procedure? We can specify the weights so that if a Skolem function is relativized to a variable x , Skolem functions lower in tree are by default also relativized to x . But note that the Skolem approach really needs the processor to enforce an ordering over scope configurations. In contrast, the LF approach provides the ordering on its own, and the processor only specifies particular weights.

The last observation shows that the theoretical relevance of experimental data is hard to assess without being even minimally explicit about processing, i.e., the structure of the parser/interpreter. Our minimal rational analysis indicated that we need some heuristic values/weights for the processor. But once we have those, we do not need semantic theories to induce orderings over scope representations, since the weights themselves can induce such orderings.

4 Conclusion

In sum, we outlined a rational (in the sense of ACT-R) analysis of the interpretation problem: we indicated how the relation between semantic and processing theories could be explicitly formalized. We introduced the specific problem of quantifier scope and the processing difficulty of inverse scope, and discussed two types of theories of scope. We presented the results of one experiment and its consequences for these two types of theories. We outlined how probabilities for scope representations, and for the LF rules used to build them, could be computed based on the experimental results. Associating weights/probabilities with our semantic representations enables our theories to make quantitative, not only qualitative, predictions. In addition, being formally explicit about processing can have a substantial impact on the interpretation of experimental results, and their (presumed) consequences for semantic theories.

References

- Anderson, John R.: 1990, *The adaptive character of thought*. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Dotlačil, Jakub and Adrian Brasoveanu: 2013, ‘The manner and time course of updating quantifier scope representations in discourse’, *Language and Cognitive Processes*. accepted with revisions.
- Ebert, Christian: 2005, *Formal investigations of underspecified representations*, Doctoral Dissertation, King’s College, London.
- Fodor, Janet Dean: 1982, ‘The mental representation of quantifiers’, in S. Peters and E. Saarinen (eds.), *Processes, Beliefs and Questions*, 129–164. Reidel, Dordrecht.
- Frazier, Lyn: 1978, *On comprehending sentences: Syntactic parsing strategies*, Doctoral Dissertation, University of Connecticut.
- Frazier, Lyn and Janet Dean Fodor: 1978, ‘The sausage machine: A new two-stage parsing model’, *Cognition* **6**, 291–325.
- Groenendijk, Jeroen and Martin Stokhof: 1991, ‘Dynamic Predicate Logic’, *Linguistics and Philosophy* **14**, 39–100.
- Hale, John: 2001, ‘A Probabilistic Earley Parser as a Psycholinguistic Model’, in *Proceedings of the 2nd Meeting of the North American Association for Computational Linguistics*, 159–166.
- Hale, John: 2011, ‘What a rational parser would do’, *Cognitive Science* **35**, 399–443.
- Kamp, Hans and Uwe Reyle: 1993, *From Discourse to Logic. Introduction to Model theoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Kluwer, Dordrecht.
- Larson, Richard K.: 1988, ‘On the Double Object Construction’, *Linguistic Inquiry* **19**, 335–391.
- Levy, Roger: 2008, ‘Expectation-based syntactic comprehension’, *Cognition* **106**, 1126–1177.
- Marslen-Wilson, William: 1973, ‘Linguistic Structure and Speech Shadowing at Very Short Latencies’, *Nature* **244**, 522–523.
- Newell, Alan and Herbert A. Simon: 1972, *Human Problem Solving*. Prentice-Hall, Englewood Cliffs, NJ.
- Plummer, Martyn: 2013, ‘rjags: Bayesian graphical models using MCMC’. R package version 3-10.
- Pylkkänen, Liina and Brian McElree: 2006, ‘The syntax-semantic interface: On-line composition of sentence meaning’, in *Handbook of Psycholinguistics*, 537–577. Elsevier, New York.

- R Core Team: 2013, ‘R: A Language and Environment for Statistical Computing’. R Foundation for Statistical Computing, Vienna, Austria.
- Radó, Janina and Oliver Bott: 2012, ‘Underspecified representations of quantifier scope?’, in M. Aloni, V. Kimmelman, F. Roelofsen, G. W. Sassoon, K. Schulz, and M. Westera (eds.), *Logic, Language and Meaning: 18th Amsterdam Colloquium*. Springer, The Netherlands.
- Steedman, Mark: 2001, *The Syntactic Process*. MIT Press, Cambridge, MA.
- Steedman, Mark: 2012, *Taking Scope*. MIT Press, Cambridge, MA.
- Szabolcsi, Anna: 1997, *Ways of scope taking*. Kluwer, Dordrecht.
- Tanenhaus, M. K., M. J. Spivey-Knowlton, K. M. Eberhard, and J. C. Sedivy: 1995, ‘Integration of visual and linguistic information in spoken language comprehension’, *Science* **268**, 1632–1634.
- Tunstall, Susanne: 1998, *The interpretation of quantifiers: Semantics and processing*, Doctoral Dissertation, University of Massachusetts, Amherst.

On the negation of indicative conditionals*

Paul Egré^{1,2} and Guy Politzer¹

¹ Institut Jean-Nicod, CNRS

² NYU, Philosophy Department

paulegre@gmail.com

politzer@univ-paris8.fr

Abstract

A debated aspect of the analysis of indicative conditionals of the form “if A then C” concerns whether they have as their negation the conjunction “A and not C” or the conditional negation “if A then not C”. We argue that neither theory is adequate, but that both forms of negation can be pragmatically retrieved from a Kratzer-style analysis of conditionals in which the negation of “if A then C” is equivalent to the weak negation “if A possibly not C”. This paper lays out the relevant pragmatic hypotheses and presents the results of one experimental study intended to test those predictions.

1 Negating conditionals

By an indicative conditional sentence, we mean an *if-then* sentence in which both the antecedent and the consequent are in the indicative mood, as in the following examples:

- (1) If John was in Paris, then Mary was in New York.
- (2) If John visits Paris tomorrow, then Mary will be pleased.
- (3) If this figure is a rectangle, then it is a square.

The question we are investigating is how the insertion of a sentential negation operator (such as “it is not the case that”) is understood, and more specifically how the denial of such sentences is expressed (as in a response starting by “No,...”), depending on the sentence and on the context.

A large part of the recent literature on the negation of conditionals has been focused on the opposition between two families of theories: on the one hand accounts based on the material conditional analysis, predicting the negation of “if A then C” to be the conjunction “A and not C” (see [6], [8], [10], [9]), and on the other suppositional theories predicting the negation to be the conditional negation “if A then not C” (including possible-world theories [20], trivalent theories [15], [1], and probabilistic theories [2], [4], [16], [5]).¹ Several experiments have been conducted in recent years by psychologists of reasoning to advance this debate (viz. [7], [3], [11]), indicating a preference for conditional negation, but with systematic exceptions (for example, [7] report differences depending on whether the conditional sentence is in the past tense, or the future tense).

*This is an abridged version of a more extended paper in preparation. The full version, in particular, also includes the results of another experimental study. We are indebted to Jean Baratgin, David Over, Philippe Schlenker, and Benjamin Spector for valuable discussions, and to Robert Stalnaker for helpful comments and criticisms. We also thank Jean-Louis Stilgenbauer for his technical assistance in the experiment, and the ANR project BETAFFDOC (“Beyond Degrees of Truth, Degrees of Confidence”) for support.

¹Suppositional theories are all inspired from Ramsey’s remarks in[18], who already concluded, from what has since been called the Ramsey Test: “in a sense If p, q and If $p, \neg q$ are contradictories”.

Our claim is that the debate has been unduly restricted to that opposition. In particular, most accounts of the psychology of conditionals have ignored a third family of accounts, predicting the negation of a conditional sentence to be equivalent to “if A , possibly not C ” (viz. [13], [12], [14] and other analyses in the family of strict conditionals, without the uniqueness assumption). In our view, this modal negation is the baseline negation for all indicative conditionals, but the other two kinds of negations are pragmatically recoverable from it based on additional features of the context and on constraints regarding the information shared between the participants in a conversation. In what follows we first give a more precise operationalized version of those hypotheses, and then present the result of one experimental study in which they were tested.

2 Basic framework

Our approach assumes a version of Krater’s modal account of conditionals, which provides a compositional version of Lewis’s theory. Specifically, Lewis predicts the following relations between the negation of a conditional, the conditional negation of the consequent, and the conjunction of the antecedent with the negation of the consequent (“ $>$ ” stands for Lewis’s conditional connective):

$$(4) \quad A \wedge \neg C \models (A > \neg C) \models \neg(A > C)$$

The first sentence is true provided A is true and C false, the second provided some $A \neg C$ -world is closer than any AC -world (or if there are no possible A -world), and the third if for every AC -world there is an $A \neg C$ -world at least as close. Moreover, $\neg(A > C)$ is in Lewis’s theory exactly a way of expressing “if A then it might be that not C ” ([13]:21), namely a conditional negation in which only the possibility of the consequent is asserted conditional on the antecedent.²

Assuming this pattern of asymmetric entailments, standard pragmatic considerations suggest that the semantic negation of a conditional would by default be the weakest negation, but that in some contexts, it may be more informative to pick a stronger negation provided some additional information is at hand. Specifically, we are making the following two hypotheses:

- (1) the conjunctive negation “ A and not C ” of a conditional of the form “if A then C ” should be favored over a conditional negation (“if A then not C ”, or “if A , it might be that not C ”) only if the speaker is sufficiently informed or sufficiently confident about both the truth status of the antecedent and the consequent. In particular, in case the speaker is uncertain about the antecedent, she should preferably use a conditional negation.
- (2) among conditional negations, the strong negation “if A then not C ” should be favored over the weak negation “if A then it might be that not C ” only if the speaker has sufficient information or sufficient reason to believe that the possibilities described by the antecedent exclude the possibilities described by the consequent.

For example, assuming geometric competence, and for a context in which one is reasoning about a figure about whose shape the only information one has is that it is a polygon, one can predict from (1) the fact that it would be too strong to deny “if this figure is a rectangle, it is a triangle” by “this figure is a rectangle and it is not a triangle”. From (2) one can predict that it would be too weak to deny the same conditional sentence by “if this figure is a rectangle,

²Within Kratzer’s framework, which we detail in the more extended version of this paper, the occurrence of “might” in negated conditionals is directly predicted, instead of stipulated.

it might not be a triangle”, but also that it would be too strong to deny “if this figure is a rectangle, then it is a square” by “if this figure is a rectangle, then it is not a square”.

3 Experimental hypotheses

Our proposal relies on a more specific typology of the denials of conditional sentences of the form “if A then C ”. First, we distinguish between *conjunctive* and *conditional* denials, depending on whether the antecedent is unconditionally asserted or not; and secondly, we distinguish between *weak* and *strong* denials, depending on whether the negation of the consequent is under the scope of a possibility operator (*possibly/need/may not*), or equivalently, scoping above a necessity operator (*not necessarily*). One substantial assumption we are making is that, when no overt modality is used to mark a weak negation, we are dealing with a strong denial. The typology is illustrated in Table 1.

	Weak	Strong
Conjunctive	‘ A and possibly not C ’	‘ A and not C ’
Conditional	‘if A then possibly not C ’	‘if A then not C ’

Table 1: Four types of denials

In the experiment we present in the next section, we each time presented our participants with a dialogue between two interlocutors, one of whom is asserting a conditional, and the other denying it, starting a sentence with ‘No...’ (following a methodology first used by [17] from an unpublished pilot study reported therein). The purpose of the experiment was to elicit negations of conditional sentences in order to get a comparative measure of the occurrence of each negation depending on the environment.

The main relational predictions we tested for were the following:

- (1) The higher the unconditional probability of the antecedent, the more we expect to see conjunctive negations. This corresponds to the idea that a conjunctive negation will be favored if the contradictor is in a position to know the truth of the antecedent.
- (2) The lower the conditional probability of the consequent given the antecedent, the more we expect to see strong negations. This corresponds to the idea that a strong negation will be favored if the contradictor is in a position to assume that the antecedent excludes the consequent.

In the experiment, we manipulated two main variables. One variable concerns the probability of occurrence of the antecedent; the other concerns the conditional probability of the consequent given the antecedent. Both probabilities are presented as transparent to the contradictor of the conditional, but not necessarily to the assertor of the conditional, who is always presumed uncertain about the realization of the antecedent.

An important caveat is that we never presented the antecedent as having probability zero (it is never an impossible event). Because of that, we did not expect to see a fifth form of denial of the conditional in terms of the plain negation of the antecedent. Suppose for example that I am talking about a game that happened yesterday and say: “if the Reds won, then it must have been a terrific party”. If the hearer knows that the Reds did not win, then the most natural way of rejecting the sentence is to update my information by saying: “actually, the Reds did not win”. Such cases would be easy to incorporate as controls, but were deliberately set aside.

P(square)	P(if square then black)			
	0	1/6	3/6	5/6
1/4	□□□□□□ ●●●○○○ ●●●○○○ ●●●○○○	■□□□□□ ●●●○○○ ●●●○○○ ●●●○○○	■■■□□□ ●●●○○○ ●●●○○○ ●●●○○○	■■■■■□ ●●●○○○ ●●●○○○
	□□□□□□ ●○	■□□□□□ ●○	■■■□□□ ●○	■■■■■□ ●○
1	□□□□□□	■□□□□□	■■■□□□	■■■■■□

Table 2: The 12 conditions defined by the combination of the probability of the conditional sentence and the probability of its antecedent.

4 Experimental Study

4.1 Method

4.1.1 Material and design

We compared denials for two kinds of sentences, a conditional sentence of the form “if A then C ”, and a control consisting of the categorical assertion of the consequent “ C ”. Each sentence was administered to a different group of subjects.

The conditional sentence referred to a chip drawn at random from a display of chips of different colors and shapes. The target sentence was “if the chip is square, it will be black” (here abbreviated as: ‘if A, C ’). The probability of the conditional $P(C|A)$ and the probability of its antecedent $P(A)$ were determined by choosing specific numbers of chips in each color and shape. There was always a first row of six square chips with 0, 1, 3, or 5 black chips allowing the values of 0, 1/6, 3/6, and 5/6 for $P(C|A)$. To determine the values of $P(A)$ an appropriate number of round chips (half of them black in all the cases) was added to the first row. $P(A)$ was set to 1/4, 3/4, or 1 by adding 18 chips, two chips, or none, respectively. In brief, there were four values for $P(C|A)$ crossed with three values for $P(A)$, hence twelve conditions (see Table 2).

The categorical sentence was “the chip will be black”. Only one row of six chips was presented, identical with the $P(A) = 1$ conditions of the conditional group, that is, zero, one, three, or five black chips out of 6, defining four levels for the probability of the target sentence.

In summary, the design was mixed. In the conditional group there were two factors: the probability for the sentence to be true (four levels) crossed with the probability of the antecedent (three levels). In the categorical group only the first factor with the same four levels was considered.

To get the expression of the denial we used the following scenario reproduced here in full (translated from French):

Pierre and Marie own a set of chips that consists of black chips and white chips that can be square or round. They have displayed the following chips :

[there follows one specific display for each condition]

Someone is going to draw a chip at random.

Pierre says : [conditional group] “if it is a square chip, it will be black”.
[categorical group] “the chip will be black”.

Marie disagrees. She wants to contradict Pierre.

But the place is noisy. Here are the only words that can be heard, following the order in which Marie has uttered them: [conditional group] *No; square; chip; black.*
[categorical group] *No; chip; black.*

Your task : Observe the chips carefully, then restore the missing words to make a full sentence. It is likely that Marie said: “No !.....” Please fill in the dots.

4.1.2 Participants

Participants were recruited from the same pool of students as in the first experiment (in which none of them had participated). They were randomly allocated to one of the two groups, and within the conditional group, to one condition. They received the scenario on a sheet of paper. Nonnative speakers of French were discarded, leaving 264 answers (conditional group N = 198; categorical group N = 66).

4.2 Results

4.2.1 Classification of denials

A substantial proportion of participants took an argumentative point of view (29%), giving reasons to indirectly predict the outcome instead of giving direct predictions. The great majority of these justifications were quantified sentences such as *all/some/ square chips are not black; not all square chips are black; no square chip is black* that expressed generalizations; a minority referred to some specific evidence (e. g. *there is one square chip that is not black*). These answers have in common that they did not refer to the singular chip to be drawn but provided reasons to make a prediction about it. We did not include such answers in the analysis, in order to focus on denials making direct reference to “the chip”, or “it”. After they were discarded, there were 187 answers to classify, out of which nine were either incomprehensible or showed incomprehension of the task (e.g., *this is not a chip*), hence a total of 178 answers (conditional group N = 123 ; categorical group N = 55).

Conditional group. Altogether, 93% of all the denials belonged to one of the four types predicted. The following percentages are relative to these predicted responses (N = 115), leaving aside the 7% that were unpredicted. There were more conditional (61%) than conjunctive responses (39%); and more weak (63%) than strong responses (37%). The corresponding cross-classification was: weak conditional 40%; strong conditional 21%; weak conjunctive 23%; strong conjunctive 16%.

- *Conditional responses.* The most frequent response was the weak form < if A, neg NEC C > = *if it is a square chip, it will not necessarily be black* (24% of the predicted responses). The remaining weak forms (16%) were various expressions of probability for C to occur, of possibility for not-C, or for C or not-C to occur. The strong conditional, < if A, neg C > = *if it is a square chip, it will not be black* was the second most frequent response (21%).

- *Conjunctive responses.* The majority of the weak forms was < A and neg NEC C > = *the square chip will not necessarily be black* (13%). The remaining weak forms (10%) were various

expressions of possibility for not- C , or for C or not- C to occur, or of likelihood for C to occur, e. g. *there is /little/more/ chance for the square chip to be black*. The majority of the strong forms was $\langle A \text{ and not-}C \rangle = \text{the square chip will not be black}$ (13%) and there were a few $\langle A \text{ and neg POSS } C \rangle = \text{the square chip cannot be black}$ (3%).

Figure 2 shows the distribution of the answers based on all the observations ($N = 123$). The unpredicted responses constituted 7% of the responses and were of the type “not- A and C ”, or “if not- A then not- C ”, or “if C then A ”, with roughly equal frequency.

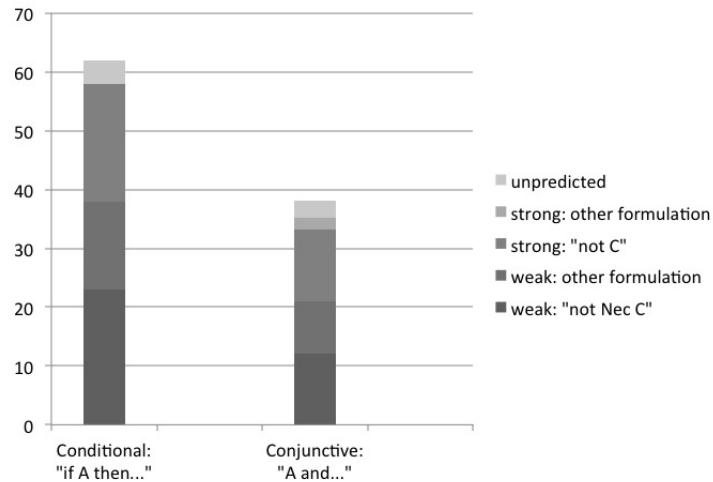


Figure 1: Percentage of conditional and conjunctive responses showing the weak and strong formulations.

Categorical group. The strong responses $\langle \neg C \rangle = \text{the chip will not be black}$ constituted the majority (64% of the observations). The majority of the weak responses (22%) were various expressions of probability or possibility for not- C , or for C to occur similar to the categorical responses above, and the remaining (14%) were $\langle \neg \text{NEC } C \rangle = \text{the chip will not necessarily be black}$.

4.2.2 Test of the predictions

We first compare the frequency of weak responses in the two groups, which differ by the sentence to be denied (conditional versus categorical). The data are: conditional sentence = $73/123=59\%$, categorical sentence = $20/55=36\%$. The difference is significant (chi-square = 9.46, $p < .01$). In the context that we have chosen, the denial of a conditional sentence is more frequently expressed by a weak negation, as opposed to a strong negation in the case in which the consequent is categorically asserted.

We now turn to the test of the relational predictions in the case of the conditional group. The rate of responses in which A is categorically affirmed, and conjunctions uttered, is predicted

$P(A)$	conditional	conjunctive	Total
1/4	27	12	39
3/4	26	10	36
1	19	23	42
Total	72	45	117

Table 3: Frequency distribution of conditional and conjunctive denials as a function of the probability of the antecedent of the conditional sentence to be denied.

$P(C A)$	weak	strong	Total
5/6	21	9	30
3/6	27	8	35
1/6	18	12	30
0	7	15	22
Total	73	44	117

Table 4: Frequency distribution of weak and strong denials as a function of the probability of the conditional sentence to be denied.

to be an increasing function of the probability of the antecedent of the conditional sentence. Kendall’s S test (with Jonckheere’s correction for continuity) applied to the data in Table 3 confirmed the general trend ($z = 2.14, p < .02$). However, the figures in Table 3 suggest that the increase is not uniform. A decomposition of the total chi-square (= 7.43, $df = 2$) after partitioning the table between the first two rows and the bottom row yielded a chi-square of 7.36 ($df = 1$), while the top two rows contributed only for .07 ($df = 1$), showing that the rate for the two lower values of $P(A)$ do not differ and that the level $P(A) = 1$ is responsible for the general trend.

The second relational prediction is that the rate of strong negations should be a decreasing function of the probability of the conditional sentence to be denied. Table 4 displays the relevant data.

There was indeed a significant trend for strong denials to be more frequent as $P(C|A)$ decreases (Kendall’s S test with Jonckheere’s correction, $z = 2.77, p < .005$). The trend was marked essentially for the null value, as shown by the partition of the total chi-square (12.82, $df = 3, p < .005$) in which the opposition of the first three rows against the last one ($df = 1$) yielded a value (10.79) close to the total chi-square.

4.3 Discussion

The results of this experiment support both the prediction that the four types of denials would occur, and the relational predictions we made. In particular, we see that conjunctive negations are favored as the antecedent of the conditional becomes more certain, and secondly, that strong negations are favored as the consequent becomes less probable relative to the antecedent. Furthermore, the comparison between denials of conditional sentences and denials of categorical sentences, which shows a significant preference for weak over strong negations in the case of conditional sentences, supports the main hypothesis whereby weak negations are semantically fundamental for indicative conditionals (assuming the semantic negation of the future tense sentence “the chip will be black” to be “the chip will not be black”).

5 Further work

One aspect we have not investigated concerns the interaction of negation with tense in conditionals. Handley et al. in [7] point out in their first study that “participants were more likely to indicate that p and not- q followed in the past compared to the future tense condition”. We surmise that the contrast, assuming it is robust, can be pragmatically explained along the lines we have outlined here (specifically, for past tense clauses favoring the implicature that the event has happened). Secondly, in this paper we have investigated how the negation of conditional sentences is expressed, but a different task would be to investigate experimentally how negated conditionals of the form “if A then not C ” and “it’s not the case that if A then C ” are interpreted. In other words, the task would be to directly compare the interpretations of the outer and inner negation of a conditional. In the framework we have adopted, “if A then not C ” and “not (if A then C)” have different truth-conditions, which is what justified our distinction between strong and weak conditional negation. Other analyses, notably Schlenker’s analysis in terms of plural definite descriptions [19], also make the distinction at the semantic level, but for the latter it has been suggested that a presupposition of homogeneity could pragmatically make “not (if A then C)” equivalent to “if A then not C ” (see [21]). We conjecture that “not (if A then C)” will generally be interpreted as weaker than “if A then not C ”, but we leave an empirical investigation of this question for further work.

References

- [1] J. Cantwell. Indicative conditionals: Factual or epistemic? *Studia Logica*, 88(1):157–194, 2008.
- [2] D. Edgington. On conditionals. *Mind*, 104(414):235, 1995.
- [3] O. Espino and R. MJ Byrne. It is not the case that if you understand a conditional you know how to negate it. *Journal of Cognitive Psychology*, 24(3):329–334, 2012.
- [4] J. Evans, S.J Handley, and D. Over. Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(2):321, 2003.
- [5] J. Evans and D. Over. *If*. Oxford: Oxford University Press, 2004.
- [6] P. Grice. Indicative conditionals. In *Studies in the Way of Words*, pages 58–85. Harvard University Press, 1989. Chapter 4.
- [7] S.J. Handley, J. Evans, and V. Thomson. The negated conditional: a litmus test for the suppositional conditional. *Journal of Experimental Psychology: Learning Memory and Cognition*, 13:378–95, 2006.
- [8] F. Jackson. *Conditionals*. Oxford University Press, 1987.
- [9] P. N. Johnson-Laird, R. MJ Byrne, and V. Girotto. The mental model theory of conditionals: A reply to Guy Politzer. *Topoi*, 28(1):75–80, 2009.
- [10] P.N. Johnson-Laird and R.M.J. Byrne. Conditionals: A theory of meaning, pragmatics, and inference. *Psychological review*, 109(4):646, 2002.
- [11] S. Khemlani, I. Orenes, and P. N Johnson-Laird. Negation: A theory of its meaning, representation, and use. *Journal of Cognitive Psychology*, 24(5):541–559, 2012.
- [12] A. Kratzer. Conditionals. in von Stechow, A. and Wunderlich, D., editors, *Semantik/Semantics: An International Handbook of Contemporary Research*, 1991.
- [13] D.K. Lewis. *Counterfactuals*. Wiley-Blackwell, 1973.
- [14] W.G. Lycan. *Real conditionals*. Oxford University Press, USA, 2001.
- [15] M. McDermott. On the truth conditions of certain ‘if’-sentences. *The Philosophical Review*, 105(1):1–37, 1996.

- [16] D.E. Over and J. Evans. The probability of conditionals: The psychological evidence. *Mind and Language*, 18(4):340–358, 2003.
- [17] G. Politzer. Could it be the case that if I am right my opponents will be pleased? a rejoinder to Johnson-Laird, Byrne and Girotto. *Topoi*, 28(1):81–85, 2009.
- [18] F. P. Ramsey. General propositions and causality. In D. H. Mellor, editor, *Foundations of Mathematics and other Logical Essays*, pages 237–257. Routledge & Kegan Paul: London, 1929.
- [19] P. Schlenker. Conditionals as definite descriptions. *Research on Language & Computation*, 2(3):417–462, 2004.
- [20] R. Stalnaker. A theory of conditionals. *Studies in logical theory*, 2:98–112, 1968.
- [21] K. von Fintel. Bare plurals, bare conditionals, and only. *Journal of Semantics*, 14(1):1–56, 1997.

The grammatical category of modality*

Valentine Hacquard

University of Maryland, College Park, Maryland, United States
hacquard@umd.edu

Abstract

In many languages, the same words are used to express epistemic and root modality. These modals further tend to interact with tense and aspect in systematic ways, based on their interpretation. Is this pattern accidental, or a consequence of grammar or meaning? I address this question by: (i) comparing ‘grammatical’ modals to verbs/adjectives that share meanings with modals, but not the same scope constraints; (ii) examining patterns of grammaticalization from ‘lexical’ to ‘grammatical’ modality; (iii) comparing scope interactions in languages where modals are ‘polysemous’ and in those where they are not.

1 Introduction

In many languages, the same modal words are used to express a variety of ‘root’ and ‘epistemic’ meanings. English *may*, for instance, can express deontic or epistemic possibility. About half of the 200+ languages in [vdAA05] have a single form that is used to express both kinds of modality. Yet, in many other languages, modal markers are unambiguously determined for meaning. In the Kratzerian tradition ([Kra81, Kra91, Kra12]), modals are lexically specified only for force (as existential or universal quantifiers over worlds), and the various meanings a ‘polysemous’¹ modal expresses arise from the modal combining with various modal bases and ordering sources. This account, based on the case of polysemous languages, easily extends to non-polysemous ones: a modal can further lexically specify the kind of modal base and ordering source it allows, restricting its meaning to a single epistemic or root meaning. An alternative account based on *non*-polysemous cases would provide separate lexical entries for the various modals, and extend to polysemous cases by postulating ambiguity. All else equal, such an account seems lacking in treating as synchronically accidental the fact that the same ambiguity is found in unrelated languages such as French, Arabic or Malay.

Interestingly, in languages that show modal polysemy, the different modal meanings a modal expresses interact differently with elements of the sentence it appears in, notably tense and aspect. A modal with an *epistemic* meaning scopes above tense and aspect, but the same modal with a root (*deontic, bouletic...*) meaning scopes below tense and aspect ([Pic90], [Sto04], [Hac06], [BC07], [Lac08], a.o.). In (1), French *pouvoir* can either express an epistemic or a root possibility. When epistemic, the modal is interpreted *above* tense and aspect: it describes a *current* possibility for a past event. When root, it is interpreted *below* tense and aspect, and yields an ‘actuality entailment’ with perfective ([Bha99]): it describes an (actualized) *past* possibility.

- (1) Jean a pu prendre le train.
Jean has could take the train

Epis: It is possible given what we *now* know that Jean took the train at some past time.

Root: At some past time, given Jean’s circumstances *then*, it was possible for him to take the train [and he did].

*Many thanks to Pranav Anand, Thomas Grano, Aynat Rubinstein, Alexander Williams, and the audience at the *TRAIT* workshop for discussion and suggestions, and to Sayaka Funakoshi for Japanese judgments.

¹I use the term polysemous as a shorthand: under a Kratzerian account, each modal is strictly monosemous.

Assuming that the differences in interaction with tense and aspect between flavors of modality can be reduced to scope (cf. [Hac06, Hac10] for arguments)², one can ask why modal flavor should correlate with modal scope: why should a modal like *pouvoir* scope ‘high’ (above tense and aspect) when it receives an epistemic interpretation, but ‘low’ (below tense and aspect) with a root interpretation? In my attempts at answering this question, I have largely ignored the many non-polysemous languages. However, the fact that languages differ in whether their modal markers are polysemous or not matters to this question. If all modals, of both the polysemous and the non-polysemous kind, show the same constraints, then we might be able to explain the interactions with tense and aspect purely in terms of meaning.³ If, however, polysemous languages do not show the same constraints, then we need to explain why polysemy should correlate with flavor/scope interactions.

[Cin99] proposes that the ordering of functional heads is fixed universally (as shown in (2)). If that is true, *all* modals, polysemous or not, should scope below or above tense and aspect as a function of their meaning. Does this ordering truly hold universally, and if so, does it follow from meaning considerations only?

- (2) **Cinque’s Hierarchy** (simplified): $\text{Mod}_{\text{epis}} > \text{Tense} > \text{Aspect} > \text{Mod}_{\text{root}} > \dots$

It seems clear that the flavor/scope correlation cannot be based on meaning alone, since *lexical* verbs and adjectives that express the same flavors of modality do not show the same constraints. Adjectives (*possible*, *likely*) and verbs (*seem*, *know*, *think*) that express *epistemic* meanings can be interpreted under a past tense (3). Nouns like *permission* and verbs like *want* that express root meanings do not yield actuality entailments with perfective (4):

- (3) a. It **seemed** to John/It was **possible/likely** that Mary was home.
- b. John **thought/knew** that Mary was home.
- (4) a. Jean **a voulu** s’enfuir, mais il n’est pas parti.
Jean **has wanted** escape, but he didn’t leave.
- b. Jean **a eu la permission** de partir, mais il n’est pas parti.
Jean **has had** the permission to leave, but he didn’t leave.

We thus need to explain why modal auxiliaries show aspectual and temporal restrictions that are unexpected from their meaning alone, and what sets “grammatical” modality (expressed by modal auxiliaries) apart from “lexical” modality (expressed by verbs and adjectives). We will then be in a position to address the question of why languages differ in the polysemy status of their modal markers, and the implications of such a dichotomy for acquisition and diachrony. In section 2, I review the account I proposed for ‘polysemous’ modals. Section 3 contrasts lexical and grammatical modality. Section 4 discusses grammaticalization, section 5 non-polysemous languages, and section 6 implications for acquisition and diachrony.

²In a nutshell, having epistemics scope above tense, and roots below, explains why the modals’ time of evaluation is the speech time (in matrix context) for the former, and the time provided by tense for the latter ([Pic90], [Sto04]). The presence/absence of actuality entailments can also be explained in terms of scope: An actuality entailment arises when the aspect that quantifies over the VP event scopes above a modal, as happens with root modals. In such a configuration, given that aspect scopes *above* the modal, its world of evaluation has to be the matrix world of evaluation, forcing the VP event to occur in the actual world. Since an epistemic modal scopes *above* aspect, it determines the world of evaluation of aspect, thereby forcing the VP event to occur in the modal worlds, but not necessarily in the actual world, as sketched below.

(1) **Asp** > **Mod_{root}**: There is an event in **w*** which in some w is an event of John taking the train.
Mod_{epis} > **Asp**: In some world **w**, there is an event in **w** of John taking the train.

³For reasons of space, I will simply assume that the generalizations with tense, aspect and modal flavors hold. For further data and discussion of counterexamples, cf. [Hac13] and references therein.

2 Scope/flavor correlations

In [Hac06, Hac10], I proposed an account in which modals are lexically unspecified for flavor, in Kratzeanian fashion, but where they can appear in two positions (TP and VP). Modals need to combine with a proposition, which both TP and VP can provide (under certain assumptions about tense and aspect). To derive the scope/flavor constraints, I proposed that modals are event-relative: their modal bases and ordering sources are determined relative to an event, rather than a world of evaluation. I argued that this event-relativity is responsible for the association between TP-level modality and epistemic flavors, and VP-level modality and root ones. Different syntactic positions make available different events that anchor modals, and these different events make available different kinds of modal bases.

Specifically, modals are quantifiers over possible worlds, determined by a modal base f , itself relative to an event e . e is an event variable (a dependent variable in the sense of [Gia98]), which needs to be bound locally. The type of modal base the modal receives will depend on the type of event that it is anchored to. In the following lexical entries, I assume, following [Kra13], that the modal base projects from the event the modal is anchored to.⁴

- (5) a. $[[\text{must}]] = \lambda q \forall w (w \in f(e) : q(w))$
- b. $[[\text{can}]] = \lambda q \exists w (w \in f(e) : q(w))$

The position in which a modal appears makes it relative to different events. At the VP-level (i.e., right below aspect), the modal's event variable gets bound by the aspect that quantifies over the VP event. This anchors the modal to the VP event. At the TP-level, the modal's event variable cannot be bound by the aspect that scopes below the modal: in matrix contexts, it is bound by the speech event e_0 ; in embedded contexts, the event variable is bound by the matrix attitude event (more specifically, by the aspect that quantifies over that attitude event):

- (6) Low (VP-level) modal: anchored to VP event
 $[CP e_0 \lambda e_0 [TP T [Asp_P Asp_1 \lambda e_1 \textbf{Mod-}e_1 [VP VP(e_1)]]]]$
- (7) High (TP-level) modal: anchored to speech event (matrix), or attitude event (embedded)
 $[CP e_0 \lambda e_0 \textbf{Mod-}e_0 [TP T [Asp_P Asp_1 \lambda e_1 [VP VP(e_1)]]]]$
 $[CP e_0 \lambda e_0 [TP T \textbf{Asp}_2 \lambda e_2 \textbf{Att}(e_2) [CP \textbf{Mod-}e_2 [TP T Asp_1 \lambda e_1 [VP VP(e_1)]]]]]]$

There are two types of modal bases: factual vs. content-based [Kra12, Kra13], responsible for root and epistemic meanings, respectively:

- (8) a. $f_{content}(e) = \{w : w \text{ is compatible with the content of } e\}$
- b. $f_{fact}(e) = \{w : w \text{ is compatible with the circumstances of } e\}$

A content-based modal base will be projected from an event that has ‘content’ (i.e., describes an information state), which speech and attitude (but not VP) events do. We obtain that only modals relative to speech and attitude events (‘high modals’, i.e., merged above tense and aspect) receive epistemic interpretations, while modals relative to VP events (‘low modals’, i.e., merged below tense and aspect) do not. When *pouvoir* scopes above tense, it gets anchored to the speech event, a contentful event: the modal base is *content-based*, and the modal quantifies over the information state associated with that event (the speaker’s commitments). When it scopes below aspect, it is anchored to the VP event. The modal base that projects from that event is *factual*: hence, the modal receives a root interpretation. Actuality entailments arise by having the aspect that quantifies over the VP event scope above the modal, forcing the event to occur in the actual world, and not merely in the worlds provided by the modal.

- (9) Jean a pu prendre le train

⁴In [Kra13], a modal takes an argument, e.g., an event or situation, which uniquely determines a *domain fixing function*, which provides the set of worlds quantified over by the modal.

- a. $[CP e_0 \lambda e_0 \text{ can-}e_0 [TP \text{ Past } [AspP \text{ Perf}_1 \lambda e_1 [VP \text{ J. take-train}(e_1)]]]]$ ‘high’ = epistemic
In some world compatible with the content of the speech event [i.e., speaker’s commitments], Jean took the train
- b. $[CP e_0 \lambda e_0 [TP \text{ Past } [AspP \text{ Perf}_1 \lambda e_1 \text{ can-}e_1 [VP \text{ J. take-train}(e_1)]]]]$ ‘low’ = root
There is an actual past event e , which in some world compatible with the circumstances of e is an event of John taking the train.

To sum up, ‘polysemous’ modals come in single lexical entries. The interpretation they can receive is however constrained by the position in which they appear. This is because modals are relative to an event. Different events project different modal bases: only speech and attitude events are contentful, hence only modals relative to speech and attitude events (i.e., modals that scope above tense and aspect) are epistemic. In this framework, then, the modal meaning/modal scope correlations are due to modals’ anaphoric nature.

3 Lexical vs. grammatical modality

As the examples in (3) and (4) show, the aspectual/temporal constraints that modal flavors exhibit cannot be blamed on meaning alone. Indeed, lexical items like *seem* or *probable* scope below past tense, despite expressing epistemic meanings. Similarly, lexical items like *want* or *permission* do not yield actuality entailments, despite expressing root meanings. Why should ‘lexical’ modals differ from ‘grammatical’ modals in this way?

As argued in section 2, grammatical modals interact the way they do with tense and aspect because of their anaphoric nature. Epistemic modals scope above tense not because they express epistemic meanings, but the other way around: modals that scope above tense and aspect happen to have epistemic meanings because they are relative to events that happen to be contentful. Similarly, actuality entailments arise when aspect scopes over a modal within its clause. This happens when modals are anchored to VP events, which happen to yield root meanings.

Why shouldn’t a lexical verb like *seem* with an epistemic meaning scope above tense? Because such a lexical item is a predicate of events, and as such, it has to appear below tense and aspect. Why shouldn’t a lexical verb with a root (bouleptic) meaning like *want* yield actuality entailments with perfective in a language like French? Again, because it is a predicate of events: as such, it requires its own aspect and tense projections, separate from those of its complement (the embedded event is quantified over by an aspect embedded under the attitude verb, anchoring it in the desire worlds, and not necessarily the actual world).⁵

- (10) $[CP e_0 [TP T [AspP Asp_1 \lambda e_1 [VP \text{ seem}(e_1) [CP \dots]]]]]$
- (11) $[CP e_0 [TP T [AspP Asp_1 \lambda e_1 \text{ want}(e_1) [CP [TP T Asp_2 \lambda e_2 [VP VP(e_2)]]]]]]$

To sum up, the main difference between lexical and grammatical modality is that lexical modals are fully specified predicates of events; grammatical modals are not, and thus they can appear in different positions. The aspectual/temporal constraints that grammatical modals are subject to are due to their anaphoric nature, rather than to the meanings they express.

Verbs and modal auxiliaries thus differ in two respects: (i) whether they are fully specified for meaning, and (ii) how they interact with tense and aspect. In the proposed account, these two differences are correlated: modal auxiliaries interact with tense and aspect the way they do due to the fact that they are not fully specified for meaning: their domain is determined by the event that they are anchored to. But what underlies the differences between ‘lexical’ and ‘functional’ modality? And could we have *grammatical* modals that are fully specified for meaning, but still obey the modal flavor/scope constraints? We turn to the first question in section 4, and the second in section 5.

⁵See [Hac06, Hac08] for actuality entailments with the Italian *restructuring* predicate *volere* (*want*), which differ from its French non-restructuring counterpart in involving a monoclausal structure (i.e., single tense and aspect projection).

4 Grammaticalization

While polysemous modal auxiliaries are relatively common, verbs⁶, nouns, and adjectives⁷, are overwhelmingly fully specified for meaning. Why should full meaning specification correlate with the functional/lexical divide, and what is this functional/lexical divide in the first place? Looking at predicates that straddle the functional/lexical line, such as restructuring verbs, and processes of grammaticalization where lexical morphemes turn into functional ones may provide useful clues.

Restructuring verbs behave like functional heads in forming a single clause with their complements, with a single tense and aspect projection ([Wur01, Cin04]). They however differ from modals in having a fixed position and meaning. [Gra12] proposes an interesting restructuring analysis of verbs like *try* and *manage*, whereby when these predicates restructure, their individual argument gets converted into a dependent variable. They are thus not predicates of individuals, but they contain a free variable over individuals, and are thus assignment-dependent. This variable gets bound when the subject raises from the VP below it, creating the illusion of control. I would like to propose that patterns of grammaticalization in which verbs turn into modal auxiliaries take this sort of functionalization one step further.

Modal auxiliaries commonly develop from lexical predicates. Processes of grammaticalization and semantic change follow ‘regular’ patterns cross-linguistically (cf. [Tra11, Nar12], and references therein), and tend to go “upwards” in the structural hierarchy of functional categories [RR03], with lexical words turning functional, and with root meanings often extending to epistemic meanings (e.g., [BPP94]). More generally, modals with an “event orientation” (i.e., having to do with the VP event and its participants) extend to modal meanings with “speech act” orientation (i.e., having to do with the speaker/hearer; [Nar12]). For instance, English *can* developed from Old English *cunnan* (‘know (how)'). Its meaning then extended to ability, to circumstantial, to epistemic possibility (for some speakers). Such a grammaticalization process might work as follows in our current framework. The verb's event argument would first be converted into a dependent variable, similarly to what Grano proposes for the individual argument of restructuring verbs. This would preserve the verb's meaning, but the verb would lose its status as predicate of events. This morpheme would still appear below aspect, and receive a root interpretation. At this point, it might allow various ordering sources, yielding different root meanings. Over time, however, being free to move, it might venture up the tree. When its event variable would get bound by the speech event, and thus be anchored to a contentful event, the modal would start yielding epistemic meanings, as sketched below.⁸

(12)	<i>know (how):</i>	$\lambda p \lambda e. \text{know-how}'(e) \ \& \ \exists w(w \text{ comp. } w / \text{know-how}(\text{Exp}(e)) : q(w))$
	<i>can-ability:</i>	$\lambda p \exists w(w \text{ comp. } w / \text{know-how}(\text{Exp}(e)) : q(w))$
	<i>can-root:</i>	$\lambda p \exists w(w \in f_{fact}(e) : q(w))$
	<i>can:</i>	$\lambda p \exists w(w \in f(e) : q(w))$

⁶One interesting polysemous verb is Mandarin *xiang* [TW00, Ngu13].

⁷The adjective *possible* looks like a ‘polysemous’ *lexical* modal: like a grammatical modal, it seems to express both epistemic and circumstantial modality, but like a lexical modal, it is able to scope under a past tense with an epistemic interpretation. Interestingly, there seems to be some correlation between meaning and syntax: when *possible* takes a nonfinite complement, it favors a circumstantial interpretation; with a finite complement, it seems more epistemic. It is not entirely clear, however, that the ‘polysemy’ of *possible* is really of the same nature than that of a modal auxiliary. For one, it is particularly difficult to distinguish separate kinds of root possibilities (bouletic or permission). One possibility is that the ‘polysemy’ of *possible* is really a matter of generality, rather than anaphoricity. Whereas for modal auxiliaries, the domain of quantification is determined anaphorically, it would be lexically determined for *possible*, and describe a more general possibility. (Thanks to Alexander Williams for this suggestion.) I leave a fuller investigation of this interesting case for future research.

⁸According to the proposal just sketched, why shouldn't an epistemic *verb* turn into an epistemic, and then circumstantial/polysemous, *modal* (a “downward” process)? Tom Grano (p.c.) suggests that a verb with a root meaning under aspect can easily be reanalyzed as a root modal in virtue of its position. However, when an epistemic verb appears below aspect, it cannot be so reanalyzed, as this would require a simultaneous shift in flavor, which would be too big of a step for reanalysis.

5 Polysemous vs. non polysemous languages

We see that what differentiates ‘grammatical’ modals from ‘lexical’ modals is both their polysemy and their interactions with tense and aspect. What about modal markers that are not polysemous? There are two logical possibilities. They could either lack anaphoricity altogether: their lexical entry could then hard-wire both meaning and scope. Alternatively, they could be just like their polysemous counterparts: event-relative, but with lexical restrictions on the kinds of modal bases they allow. For instance, an event-relative modal could be specified to only allow content-based modal bases. Such modals would have to appear in a position where they would be anchored to a contentful event (e.g., speech event). We would expect modals’ scope to correlate with meaning with the second option, but not necessarily with the first. How do polysemous languages actually behave?

In a language like Japanese, modals are fully specified for meaning, and they do not seem to show scope/meaning correlations. *kamosirenai* and *nitigainai*, for instance, express epistemic possibility and necessity, respectively, but unlike their French and English counterparts, they easily scope under tense:

- (13) John-ga hannin-de aru {**kamosirenakat-ta/nitigainakat-ta**}
John-Nom culprit-Pred be **might-past/must-past**
(At some point) It was possible/necessary that John was the murderer.

However, in many other languages, the scope of modals seems to be rather strict. [Nau08] argues that modals from diverse language families show the same Cinquean ordering of epistemic over root modals, regardless of whether they are fully-specified for meaning, at least with respect to one other. The epistemic enclitic *-kəl*, and the circumstantial circumfix *kaa* in Lilloet, for instance, are fully-specified for flavor, and only allow the epistemic to scope over the circumstantial. Whether this ordering of modals can be explained solely on conceptual grounds (and whether it really holds cross-linguistically) is not entirely settled (cf. [vFI04, Kra76, Hac11, Hac13], a.o.). But to the extent that it holds, it supports the anaphoric view. Grammatical modal markers, whether polysemous or not, would all be event-relative, and project modal bases from the event they are relative to, explaining the correlation between scope and meaning. Non-polysemous languages would differ from polysemous ones in lexically imposing restrictions on the kinds of modal bases their modals can combine with.

Clearly, much more empirical work is needed before a definite answer can be provided, as much more is known about the interaction between modals, tense and aspect in polysemous languages than in non-polysemous ones. What gets classified as modal markers in grammars do not necessarily form a uniform grammatical class. Thus, the answer may vary based on the functional vs. lexical status of such markers. As for the Japanese modals in (13), one might argue that they are in fact lexical, based on their adjectival morphology, the relative temporal independence of their complement, and the fact that epistemics can scope below deontics, as shown in (14) (translated from [vFI04] English example by S. Funakoshi, p.c.):

- (14) Kensa hiyoo-no harai-modosi-no tameni-wa, kanja-wa arutuhaimaa(-ni
Test cost-Gen reimbursement-Gen for-Top, patient-Top Alzheimer's(-Dat
kakattei-ru) kamosirenake rebanaranai.
be.taken.with-present) possible must
'For reimbursement of the test costs, it must be possible that the patient has Alzheimer's.'

6 Acquisition and diachrony

Let’s take stock by taking the point of view of a learner, trying to figure out how modality works in her language. The child has to figure out whether modals in her language are polysemous or not, and has to associate certain modal meanings with certain scope positions: modals with epistemic meanings scope high, modals with root meanings scope low. Such a mapping should be particularly difficult when the *same* words are used to express both epistemic and root modality. Hence, the hope is that

the meaning/scope association is principled—if indirect—as I have argued here. If the association were arbitrary, we would expect incorrect meaning/scope mappings, and the ordering of modals and tense and aspect to flip over time. However, diachronic patterns show that this is not the case: modals develop in systematic ways.

Languages may differ in whether modality is expressed by functional morphemes or solely by lexical ones. They further differ in whether their modals show polysemy or not. The first choice point for the child then will be whether modality in her language is expressed by functional or only lexical vocabulary. Hopefully the clues for such a distinction (morphological and other) will be clear enough. If modals are functional, the second choice point will be to figure out whether they are polysemous or not. If the child figures out that the same modal can scope in two positions (which should be easy at least in languages where syntactic scope transparently reflects semantic scope), knowing that this modal is anaphoric (because functional) should help her figure out the correct flavor/scope mappings.

As for why root modals tend to extend to epistemic meanings in some languages (Romance and Germanic), but not others (Japanese), it may largely be due to historical idiosyncracies. [Nar12] proposes that while both types of languages faced at some point a need for new morphemes to express epistemic modality, the items available for meaning extensions were not the same. English and German used verbs that already had a deontic meaning, but such verbs did not (and still do not) exist in Japanese. Japanese's epistemetics thus evolved from various other lexical items (e.g., *nitigaina* derives from a construction meaning 'there is no mistake that').

Thus, whether modals in a given language show polysemy or not may be largely idiosyncratic. However, the way polysemy develops and how the resulting modal meanings scopally interact is principled. Languages may not all necessarily have grammatical modals, let alone polysemous ones. Yet, given the way polysemy arises diachronically, and works synchronically (if this account is on the right track), the correct scope/flavor correlations should be recoverable by the learner.⁹

7 Conclusion

Modal markers are unique in language in exhibiting systematic polysemy in many languages, and having their meanings further correlate with scope. I have argued that modals' 'polysemy' and their associated temporal and aspectual restrictions follow from their anaphoric dependency. That lexical verbs/adjectives do not show the same temporal and aspectual constraints support this fact: the correlations are not purely meaning-based. I have argued that the main difference between 'lexical' and 'grammatical' modals is whether they are predicates of events or merely event-dependent, and suggested that patterns of grammaticalization support this view: lexical modals turn functional when they lose their event argument. This loss, or rather conversion, leads to anaphoricity, which leads to scope/flavor constraints. Finally, while modals in some non-polysemous languages show similar scope/meaning constraints as in polysemous languages, they do not in others. Ultimately, the answer to the universality of scope and meaning of modal markers will depend on the grammatical category and status of such markers in their respective language.

References

- [BC07] C. Borgonovo and S. Cummins. Tensed modals. In *Coreference, modality, and focus: Studies on the syntax-semantics interface*, Linguistik Aktuell, pages 1–18. John Benjamins, 2007.
- [Bha99] R. Bhatt. *Covert modality in non-finite contexts*. PhD thesis, UPenn, 1999.
- [BPP94] J. Bybee, R. Perkins, and W. Pagliuca. *The evolution of grammar: Tense, aspect and modality in the languages of the world*. University of Chicago Press, Chicago, 1994.

⁹Cross-linguistically, modals also differ in whether they encode *force* of quantification (cf. e.g., [RM08]). This variability raises important further questions about acquisition, diachrony ([Yan13]), and universal properties of modality ([Nau08] proposes the tentative universal that modals can either express variability in flavor or force, but not both). I leave the integration of these questions for future research.

- [Cin99] G. Cinque. *Adverbs and functional heads: A cross-linguistic perspective*. Oxford University Press, Oxford, 1999.
- [Cin04] G. Cinque. Restructuring and functional structure. In A. Belletti, editor, *Structures and beyond: The cartography of syntactic structures*, volume 3, pages 132–191. 2004.
- [Gia98] A. Giannakidou. *Polarity sensitivity as (non)veridical dependency*. John Benjamins, Amsterdam, 1998.
- [Gra12] T. Grano. *Control and restructuring at the syntax-semantics interface*. PhD thesis, University of Chicago, 2012.
- [Hac06] V. Hacquard. *Aspects of modality*. PhD thesis, MIT, 2006.
- [Hac08] V. Hacquard. Restructuring and implicative properties of *volere*. In A. Gronn, editor, *Proceedings of Sinn und Bedeutung 12*, 2008.
- [Hac10] V. Hacquard. On the event relativity of modal auxiliaries. *Natural Language Semantics*, 18/1, 2010.
- [Hac11] V. Hacquard. Modality. In C. Maienborn et al, editor, *Semantics: An international handbook of natural language meaning*, volume HSK 33.2, pages 1484–1515. de Gruyter, Berlin, 2011.
- [Hac13] V. Hacquard. Modals: Meaning categories? University of Maryland, m.s., 2013.
- [Kra76] A. Kratzer. Was “können” und “müssen” bedeuten können müssen. *Linguistische Berichten*, 42:128–160, 1976.
- [Kra81] A. Kratzer. The notional category of modality. In *Words, worlds, and contexts*. de Gruyter, Berlin, 1981.
- [Kra91] A. Kratzer. Modality. In A. von Stechow and D. Wunderlich, editors, *Semantik: Ein internationales Handbuch zeitgenössischer Forschung*, pages 639–650. de Gruyter, Berlin, 1991.
- [Kra12] A. Kratzer. *Modals and Conditionals*. Oxford University Press, Oxford, 2012.
- [Kra13] A. Kratzer. Modality for the 21st century. In *Keynote address for ICL 2013 Geneva*, 2013.
- [Lac08] B. Laca. On modal tenses and tensed modals. Université Paris 8/CNRS, m.s., 2008.
- [Nar12] H. Narrog. *Modality, Subjectivity, and semantic change*. Oxford University Press, Oxford, 2012.
- [Nau08] F. Nauze. *Modality in typological perspective*. PhD thesis, Amsterdam University, 2008.
- [Ngu13] E. Nguyen. Observations on mandarin *xiang*. University of Connecticut, m.s., 2013.
- [Pic90] C. Piccalo. Modal verbs in Catalan. *Natural Language Linguistic Theory*, 8(2):285–312, 1990.
- [RM08] H. Rullmann and L. Matthewson. Modals as distributive indefinites. *Natural Language Semantics*, 16:317–357, 2008.
- [RR03] I. Roberts and A. Roussou. *Syntactic change: a minimalist approach to grammaticalization*. Cambridge University Press, Cambridge, 2003.
- [Sto04] T. Stowell. Tense and modals. In J. Guéron and J. Lecarme, editors, *The syntax of time*, pages 495–537. MIT Press, Cambridge, MA, 2004.
- [Tra11] E. Traugott. Modality from a historical perspective. *Language and Linguistics Compass*, 5/6:381–396, 2011.
- [TW00] T. Tardif and J. Wellman. Acquisition of mental state language in Mandarin- and Cantonese-speaking children. *Developmental Psychology*, (36):25–43, 2000.
- [vdAA05] J. van der Auwera and A. Ammann. Overlap between epistemic and situational possibility. In M. Haspelmath et al, editor, *World atlas of language structures*. Oxford University Press, Oxford, 2005.
- [vFI04] K. von Fintel and S. Iatridou. Seminar notes. MIT, ms., 2004.
- [Wur01] S. Wurmbrand. *Infinitives*. Mouton de Gruyter, New York, 2001.
- [Yan13] I. Yanovich. *Four pieces for modality, context and usage*. PhD thesis, MIT, 2013.

Quantifier Particles and Compositionality

Anna Szabolcsi*

New York University, NY
as109@nyu.edu

Abstract

In many languages, the same particles build quantifier words and serve as connectives, additive and scalar particles, question markers, existential verbs, and so on. Do the roles of each particle form a natural class with a stable semantics? Are the particles aided by additional elements, overt or covert, in fulfilling their varied roles? I propose a unified analysis, according to which the particles impose partial ordering requirements (glb and lub) on the interpretations of their hosts and the immediate larger contexts, but do not embody algebraic operations themselves.

1 The compositionality question

Formal semanticists often treat even multi-morphemic words as compositional primitives. This paper examines a domain of data in which extending compositionality below the word level seems especially rewarding.

English *some*, *or*, *whether*, *every*, *both*, *and*, *even*, *too*, and *also* look like a motley crew. But in many other languages, the same particles build quantifier words and serve as connectives, additive and scalar particles, question markers, existential verbs, and so on. It is natural to ask if they are really “the same” across their varied environments.

Consider the following samples. Hungarian *ki* and Japanese *dare*, usually translated as ‘who’, are indeterminate pronouns in the terminology of Kuroda 1965. *Ki* and *dare* form ‘someone’ and ‘everyone’ with the aid of morphemes whose more general distribution is exemplified below. The joint distribution of Hungarian *vala/vagy* and etymologically unrelated *-e* corresponds, roughly, to that of Japanese *-ka*. The joint distribution of *mind* and *is* corresponds to that of *-mo* (see further in Szabolcsi, Whang & Zu 2013, Szabolcsi 2013).

Slavic languages, Malayalam, Sinhala, and many others exhibit similar patterns. I will use upper-case KA and MO as generic cross-linguistic representatives.

(1)	<i>Hungarian</i>	<i>Japanese</i>	<i>Gloss</i>
a.	vala-ki	dare-ka	‘someone’
b.	A vagy B	A- ka B(- ka)	‘A or B’
c.	vagy száz	hyaku-nin-to- ka	‘some 100 = approx. 100’
d.	val- , vagy-	—	‘be’ participial & finite stems
e.	—	dare-ga VP- ka	‘Who is VP-ing?’
f.	[S-e]	S- ka	‘whether S’

(2)	<i>Hungarian</i>	<i>Japanese</i>	<i>Gloss</i>
a.	mind-en-ki	dare-mo	‘everyone/anyone’
b.	mind A mind B [A is (és) B is]	A- mo B- mo	‘A as well as B, both A and B’
c.	[A is]	A- mo	‘also/even A’

*I thank Ivano Ciardelli, Marcel den Dikken, Salvador Mascarenhas, and Benjamin Slade for discussions.

2 A promising perspective: join and meet

There is a beautiful generalization that caught the eyes of many linguists working with these data (Gil 2008, Haspelmath 1997, Jayaseelan 2001, 2011, among others; see Szabolcsi 2010: Ch 12). In one way or another, the roles of KA involve existential quantification or disjunction, and the roles of MO involve universal quantification or conjunction. Generalizing,

- (3) KA is lattice-theoretic join (\cup), MO is a lattice-theoretic meet (\cap).

Alternative Semantics has thrown a new light on the signature environments of KA. Hamblin 1973, Kratzer & Shimoyama 2002, Alonso-Ovalle 2006, Aloni 2007, AnderBois 2012, and others proposed that not only polar and wh-questions but also declaratives with indefinite pronouns or disjunctions contribute sets of multiple classical propositions to interpretation. They contrast with declaratives that are atomic or whose main operations are negation, conjunction, or universal quantification, and contribute singleton sets of classical propositions. If the universe consists of Kate, Mary, and Joe, we have,

- (4) a. Who dances?, Someone dances, Kate or Mary or Joe dances
 $\{\{w : \text{dance}_w(k)\}, \{w : \text{dance}_w(m)\}, \{w : \text{dance}_w(j)\}\}$
- b. whether Joe dances
 $\{\{w : \text{dance}_w(j)\}, \{w : \neg\text{dance}_w(j)\}\}$
- (5) a. Joe dances
 $\{\{w : \text{dance}_w(j)\}\}$
- b. Everyone dances
 $\{\{w : \text{dance}_w(k) \& \text{dance}_w(m) \& \text{dance}_w(j)\}\}$

Inquisitive Semantics (say, Ciardelli et al. 2012) develops a notion of propositions as non-empty, downward closed sets of information states. The sentences in (4) and (5) are recognized as inquisitive and non-inquisitive propositions, respectively; disjunction and conjunction re-emerge as Heyting-algebraic join and meet.

The upshot is that the Alternative/Inquisitive Semantic perspective offers an even more interesting way to unify KA's environments, and maintains the possibility to treat KA as a join and MO as a meet operator, although in a slightly modified algebraic setting. In other words, it looks like the core roles of KA and MO can be assigned a stable semantics, and a simple one at that.

3 Mismatches: Too few arguments, too many operators

There are two general problems with this beautiful approach. The first problem is that both KA and MO may have just one argument. Schematically,

- (6) Hungarian (KA = *vagy*, MO = *is*), Russian (MO = *i*), Japanese (MO = *mo*):
- | | |
|-----------------|---------------------------------------|
| 10-KA boys ran. | 'Approximately/at least ten boys ran' |
| John-MO ran. | 'Also/even John ran' |

The flip-side problem is that in some cases KA and MO occur on all their alleged arguments. In Sinhala, both inclusive disjunction *hari* and alternative question forming disjunction *da*

attach to each disjunct, as illustrated in (7). Japanese *mo*, Russian *i*, and Hungarian *mind* as well as *is* all attach to each conjunct in the distributive construction illustrated in (8).

- (7) Sinhala (KA = *hari* / *də*):

John-KA Mary-KA ran.	'John or Mary ran'
John-KA Mary-KA ran?	'Did John run, or did Mary?'

- (8) Japanese (MO = *mo*), Russian (MO = *i*), Hungarian (MO = *mind* / *is*)

John-MO Mary-MO ran	'John as well as Mary ran'
---------------------	----------------------------

Russian *li* and Hungarian *-e*, the morphemes that mark alternative questions alternate with ‘or not’ in glorious justification of the Hamblin/Karttunen analysis of *whether*. But, embarrassingly, they also co-occur with ‘or not’ – which is in fact equally possible in the case of *whether*.

- (9) Russian (KA = *li*), Hungarian (KA = *-e*)

... John ran-KA	
... John ran or not	'whether (or not) John ran'
... John ran-KA or not	

The first problem might be explained away by assuming that the single argument represents or evokes a set of alternatives, to which join and meet can sensibly apply. But it is not clear how that assumption would explain the cases where KA and MO attach to each of the dis/conjuncts, i.e. where we have too many actors for one role.

I conclude that KA is not join, and MO is not meet. But, in solving the problems I would like to preserve the insight that KA and MO occur precisely in contexts that are the least upper bound / greatest lower bound of the contribution of the host of KA/MO and something else.

4 The gist of the solution: KA/MO impose semantic requirements on the context

There are three basic strategies for solving the mismatch problems:

- (10) a. KA and MO are meaningful, but their mission in the compositional process is not directly related to \cup and \cap .
- b. KA and MO are meaningless syntactic elements that point to (possibly silent) meaningful \cup and \cap operators. Compare \pm interpretable features.
- c. KA and MO are meaningful elements that point to least upper bounds (join) and greatest lower bounds (meet) in a semantic way. Compare presuppositions.

The analysis of KA in Hagstrom 1998, Yatsushiro 2009, Cable 2010, and Slade 2011 can be seen to represent option (a). On this view, KA is a choice-function variable that eliminates alternatives. I will not pursue this analysis here, because it inherits the problems of choice-functional analyses of indefinites, it offers no parallel insight for MO’s role, it assumes that alternatives (in general, sets as opposed to individuals) are bad for you, and it does not help with the “too many actors” problem.

Variants of option (b) have been proposed in Carlson 1983, 2006 for all functional categories, in Ladusaw 1992 for negative concord, in Beghelli & Stowell 1997 for *every/each*, and in Kratzer 2005 for *ka*, *mo*, and more concord phenomena. Taking KA and MO to be meaningless syntactic pointers could be viable. But I'm going to argue that the semantic route is also viable and interesting.

Option (c) says that KA and MO are meaningful elements that point to joins and meets in a semantic way. This is what I am going to pursue. My approach draws from Kobuchi-Philip 2009 and Slade 2011, works that took seriously some problems that other literature glossed over, and provided important elements of the solutions.¹

MO is a good starting point, because the standard analysis of *too* easily extends to MO in *John-MO ran* ‘John, too, ran’ (I put scalar ‘even’ aside). *John-MO ran* is thought to assert that John ran and to presuppose that a salient individual distinct from John ran. So MO can be seen as a “semantic pointer” — it points to a fact not mentioned in the sentence, and ensures that the context is such that both John and another individual ran. Kobuchi-Philip’s insight is that in *John-MO Mary-MO ran* ‘John as well as Mary ran’, both MO’s can be seen as doing the same thing. John’s running and Mary’s running mutually satisfy the requirements of the two MOs. Similarly for *Person-MO ran* ‘Everyone ran’, with generalized conjunction.

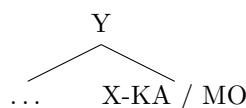
The mutual satisfaction of requirements is reminiscent of presupposition projection, and so a small amendment is called for. Presupposition projection works left-to-right, at least when it is effortless (Chemla & Schlenker 2012). If so, the symmetrical character of these constructions is a problem. I reclassify these definedness conditions as presuppositions in the sense of Brasoveanu 2013: tests that are delayed and checked simultaneously after the at-issue content is established. This is utilized in *John-MO Mary-MO ran*. In contrast, if nothing in the at-issue content satisfies the test, it is imposed on the input context and emerges as a presupposition. The traditional analysis of *John-MO ran* is reproduced. For details see Brasoveanu & Szabolcsi 2013.

Below, I will use the neutral term “requirement” instead of pre- or postsupposition. I assume that the same reasoning carries over to KA, whose semantics will be detailed in Section 6.

5 Ingredients of the analysis: Greatest lower bound / least upper bound requirements, pair formation, silent meet and join, defaults

To summarize, the “mismatch cases” offer the best insight into the working of the particles. The particles do not embody algebraic operations, as examples of the form *A Particle B* would lead us to believe. Instead, I suggest, the particles require that interpretations of their hosts and of the immediately larger contexts stand in particular partial ordering relations. The core of the proposal is this. For simplicity, I pretend that the hosts of KA and MO are propositions.

(11)



¹In particular, Kobuchi-Philip 2009 proposed a unified analysis of Japanese *mo* that I follow with minor modifications. Slade 2011 was the first to take up the challenge posed by the multiple occurrences of KA-style particles and to invoke den Dikken’s 2006 Junction; see Section 5. Also, Slade 2011 is my source for Sinhala.

- (12) a. MO requires that another proposition parallel to $\llbracket X \rrbracket$ hold in $\llbracket Y \rrbracket$. MO's requirement is trivially satisfied if $\llbracket Y \rrbracket$ is the meet (glb) of $\llbracket X \rrbracket$ and something else.
 b. KA requires that the alternatives in $\llbracket X \rrbracket$ be preserved and boosted in $\llbracket Y \rrbracket$. KA's requirement is trivially satisfied if $\llbracket Y \rrbracket$ is the join (lub) of $\llbracket X \rrbracket$ and something else.

But MO and KA do not perform meet and join operations. Who does, then?

Winter (1995, 1998: Ch 8) argued that the word *and* in *A and B* is not conjunction; it merely forms pairs:

- (13) $A \text{ and } B = A \bullet B = \langle A, B \rangle$

Pairs grow pointwise, much like Hamblinian alternatives. At some point silent meet (\cap) applies, creating the illusion that *and* scopes there. The pair-former can be silent (asyndetic conjunction); in many languages it is always silent. In contrast, *or* is cross-linguistically almost never silent (no asyndetic disjunction). I adopt the decomposition of conjunction into pair-formation and silent meet, with the following modifications:

- (14) a. Identify the pair-former, whether overt or silent, with the syntactic Junction (J) head that den Dikken 2006 detected in *either ... or ...* disjunctions.
 b. Delimit the pointwise growth of pairs to avoid scopal overgeneration.
 c. Replace Winter's plain Boolean \cap with Dekker's 2012 order-sensitive \cap , which interprets the 2nd conjunct strictly in the context of the 1st (cf. anaphora).
 d. Recognize silent order-sensitive \cap as the default propositional operation.

Canonizing order-sensitive \cap as the default will be important in regulating the use of KA, but it has independent motivation. All languages use silent order-sensitive \cap for text-level sequencing. Bumford 2013 even shows that the definition of universal quantifiers involves the iteration of order-sensitive \cap .

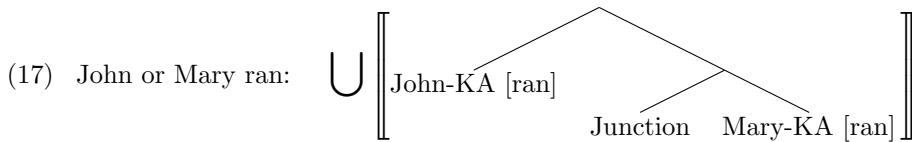
To wrap up the emergent analysis of *A-MO B-MO* 'A as well as B,' it contains a pair-forming \bullet , overt or silent, silent \cap , and the MOs, which impose postsuppositional requirements on the context and make the construction irrevocably distributive in the process.

Compare the following Hungarian examples. In both, *és* 'and' is the pair-former Junction, and silent \cap applies. The collective reading of (15) is due to a shift, as in Winter 1998, 2002. The collective shift is bled by the MO-particles, as in (16).

- (15) *Kati és Mari felemezte az asztalt.*
 Kate and Mary up-lifted the table-ACC
 'Kate and Mary lifted up the table, individually or together'
 (16) *Kati is (és) Mari is felemezte az asztalt.*
 Kate MO (and) Mary MO up-lifted the table-acc
 'Kate as well as Mary lifted up the table, individually'

See Mitrović & Sauerland 2013 for connectives that incorporate both Junction and MO or KA, among them Latin *at-que* 'and, nevertheless' and Russian *i-li* 'or'.

We now turn to KA. I take the iterated KA of Sinhala to be paradigmatic, cf. (7), and add silent join (\cup) to the inventory. Winter assumed that conjunctions and disjunctions have completely different structures. I assume that they have the same structure, with silent \cap versus silent \cup .



The significant fact, I suggest, is that cross-linguistically, KA is only mandated in disjunctions ($A \vee B$ never means ‘ A or B ’), but not in wh-questions or in indefinites. German offers good examples of the absence of any KA-style particle from the latter contexts (Haida 2007):

- (18) *Wer MAG was?*
 who likes what
 ‘Who likes something?’

KA must be present in disjunctions, because the default operation on pairs is meet, \cap . KA’s semantic requirement forces the pair to be fed to join, \cup .

In contrast, when join is the default, KA’s presence is subject to variation. One reason for non-disjunctive KA to be present can be to signal that join composes with another operator, as in the various flavors of epistemic indefinites.

Why KA/MO is sometimes present on each “junct” and sometimes on just one requires further morpho-syntactic research. But, my assumption is that there is no “completely silent” KA/MO. If there is not a single overt copy in the construction, KA/MO is not there, and the construction is interpreted via semantic defaults.

6 Alternatives in KA’s host must be preserved and boosted

In line with (12b), KA imposes the following requirement:

- (19) KA requires that $\llbracket X \rrbracket$, the alternatives introduced by its host X , be preserved and boosted in $\llbracket Y \rrbracket$, the interpretation of the immediately larger context: $\llbracket X \rrbracket < \llbracket Y \rrbracket$. This ensures that $\llbracket Y \rrbracket$ is the least upper bound of $\llbracket X \rrbracket$ and something else.

Let us first see how things pan out with linguistic data beyond plain disjunctions. The core indefinite and wh-question examples reproduce (4a):

	Hungarian	Japanese	Gloss	Meaning
(20)	-ki táncol	dare- odorimasu		$\{\{w : \text{dance}_w(x)\} : \text{person}(x)\}$
	Vala-ki táncol	dare-ka odorimasu	‘Someone dances’	
	Ki táncol?	dare-ga odorimasu-ka	‘Who dances?’	$\bigcup\{\{w : \text{dance}_w(x)\} : \text{person}(x)\}$

Krifka 2001, arguing for structured meanings for questions, distinguishes polarity questions, which can be answered by plain *Yes* or *No*, from alternative questions, which require repeating one of the alternatives, possibly accompanied by *Yes* or *No*. (He notes that Karttunen 1977 considers polarity questions a subclass of alternative questions.) Let us look at Hungarian data. ↑ indicates final rising intonation, and ↓ falling, declarative intonation. (No such distinction exists in complements.) I consider the *-e* suffix on the finite verb a KA-particle, although it is etymologically unrelated to *vala/vagy*:

- (21) a. Alszik? ↑ ‘Is he asleep?’ **compl:** *... alszik.
 b. Alszik vagy nem? ‘Is he asleep or not?’ ... alszik vagy nem.
 c. Alszik-e? ↓ ‘Is he asleep-KA?’ ... alszik-e.
 d. Alszik-e vagy nem? ‘Is he asleep-KA or not?’ ... alszik-e vagy nem.

I propose that only *Alszik?* ↑ is a Krifkean polarity question and that polarity questions are a main-clause phenomenon, interpreted via the Inquisitive Semantic ? operator (non-informative closure). Alternative questions, being disjunctions, always contain either one KA (-e or *vagy*) or two (-e and *vagy*). The KA-particle -e requires $\llbracket X \rrbracket < \llbracket Y \rrbracket$, as in (19). In (21c) the only possible exclusive alternative is “accommodated” — in (21d) it is spelled out.

The approximate-number constructions in (1c) corroborate the existence of “unary KA” alongside “unary MO”. *Hyaku-nin-to-ka-, vagy száz,* and *some one hundred* mean, ‘100 or another number in the vicinity of 100.’ Dutch *een* (M. den Dikken, p.c.) and Hebrew *eyze* (D. Farkas, p.c., with reference to Kagan & Spector 2008) also attach to full arguments:

- (22) a. Neem een Chomsky.
 ‘Take for example Chomky’ = ‘Take Chomsky or someone like him’
 b. Kxi eyze tapuax!
 ‘Take an apple or something’ = ‘Take an apple or some other fruit’

Let us now make the interpretation more precise. Switching from Alternative Semantics to Inquisitive Semantics does not have a major effect on the linguistic analysis of KA, but it gives access to well-defined algebraic tools, and to the ? and ! operators (non-informative and non-inquisitive closure, resp.). I assume some basic definitions in Ciardelli et al. 2012.²

An Inquisitive Semantic definition of $\llbracket X \rrbracket < \llbracket Y \rrbracket$ must ensure that the alternatives of KA’s host X are “preserved” and “boosted”. It must approve $\llbracket Y \rrbracket = \bigcup(\llbracket X \rrbracket, \llbracket Z \rrbracket)$ and possibly also $\llbracket Y \rrbracket = ((\bigcup(\llbracket X \rrbracket, \llbracket Z \rrbracket))^*)^*$ i.e. one-fell-swoop non-inquisitive join, as satisfiers of KA. It must exclude shrinking, $\llbracket Y \rrbracket = \bigcap(\llbracket X \rrbracket, \llbracket Z \rrbracket)$ and what I dub endogamy, $\llbracket Y \rrbracket = ((\llbracket X \rrbracket)^*)^*$, as satisfiers of KA. The following definition will do; other definitions are imaginable if we abandon downward closure.

- (23) $\llbracket X \rrbracket < \llbracket Y \rrbracket$ iff every possibility in $\llbracket X \rrbracket$ is a possibility in $\llbracket Y \rrbracket$, and $\llbracket Y \rrbracket$ contains a possibility that is excluded in $\llbracket X \rrbracket$.

(24)-(25) examine the effect from the perspective of just one KA, which suffices for the formal demonstration. Both (24a) and (24b) preserve all possibilities in $\llbracket \text{Mary runs} \rrbracket$, and add a possibility excluded in $\llbracket \text{Mary runs} \rrbracket$, e.g. $\{k \neg m\}$ = only Kate runs. KA is happy.

- (24) a. $\llbracket Y \rrbracket = \bigcup(\llbracket \text{KA}(\text{Mary runs}) \rrbracket, \llbracket \text{Kate runs} \rrbracket)$
 = $\wp\{w : \text{run}_w(m)\} \cup \wp\{w : \text{run}_w(k)\}$
 = $\{\emptyset, \{m \neg k\}, \{mk\}, \{m \neg k, mk\}, \{k \neg m\}, \{k \neg m, mk\}\}$

²A **proposition** is a non-empty, downward closed set of possibilities. A **possibility** is a set of worlds. $\llbracket \phi \rrbracket = \llbracket \text{John runs} \rrbracket = \wp\{w : \text{run}_w(j)\}$. The informative content of ϕ , $\text{info}(\phi) = \bigcup[\phi]$. **Meet:** $A \cap B$. **Join** = $A \cup B$. **Pseudo-comp:** $A^* = \{\beta : \text{disjoint}(\beta, \cup A)\}$. $A \cap A^* = \perp$, but $A \cup A^*$ may or may not be \top . ϕ is **informative** iff $\text{info}(\phi) \neq W$; ϕ excludes something in W . ϕ is **inquisitive** iff $\text{info}(\phi) \notin [\phi]$; ϕ has > 1 maximal possibility. An **alternative** is a maximal possibility. Non-inquisitive closure $[\! [\phi] \!] = ([\phi]^*)^* = \wp(\text{info}(\phi))$. Non-informative closure $[? \phi] = [\phi] \cup [\phi]^*$.

- b. $\llbracket Y \rrbracket = ((\bigcup(\llbracket KA(Mary \text{ runs}) \rrbracket, \llbracket Kate \text{ runs} \rrbracket))^*)^*$
 $= \wp\{w : \text{run}_w(m) \vee \text{run}_w(k)\}$
 $= \left\{ \begin{array}{l} \emptyset, \{m \neg k\}, \{mk\}, \{m \neg k, mk\}, \\ \{k \neg m\}, \{k \neg m, mk\}, \{m \neg k, k \neg m\}, \{m \neg k, k \neg m, mk\} \end{array} \right\}$
- (25) a. $\llbracket Y \rrbracket = \bigcap(\llbracket KA(Mary \text{ runs}) \rrbracket, \llbracket Kate \text{ runs} \rrbracket)$
 $= \wp\{w : \text{run}_w(m) \wedge \text{run}_w(k)\} = \{\emptyset, \{mk\}\}$
KA: \cap eliminates $\{m \neg k\}$ from $\llbracket Mary \text{ runs} \rrbracket$; possibilities are shrinking.
- b. $\llbracket Y \rrbracket = ((\llbracket KA(Mary \text{ runs or Kate runs}) \rrbracket)^*)^*$
 $= ((\wp\{w : \text{run}_w(m)\} \cup \wp\{w : \text{run}_w(k)\})^*)^*$
 $= \left\{ \begin{array}{l} \emptyset, \{m \neg k\}, \{mk\}, \{m \neg k, mk\}, \{k \neg m\}, \\ \{k \neg m, mk\}, \{m \neg k, k \neg m\}, \{m \neg k, k \neg m, mk\} \end{array} \right\}$
KA in this position: ! preserves the possibilities in inquisitive $\llbracket Mary \text{ or Kate runs} \rrbracket$, but the new possibilities are all joins of old possibilities; endogamy.
- It is important to have ! in interpreting indefinites and disjunctions. Generally, ! attaches to the scopes of externally static operators (27a), although not to all clause-boundaries (27b):
- (26) Bill saw Joe or some girl, but I forget who/which
‘whether Bill saw Joe or Bill saw some girl’
won’t suffice: $[\phi_{joe}] \cup [\phi_{kate}] \cup [\phi_{mary}]$ — needed: $[\phi_{joe}] \cup (([\phi_{kate}] \cup [\phi_{mary}])^*)^*$
- (27) a. Bill didn’t invite Kate or Mary, but I forget which
‘didn’t invite !(Kate or Mary)’
‘didn’t invite Kate or didn’t invite Mary’
b. Bill thinks that they hired Kate or Mary, but I forget which
‘whether Bill thinks that they hired Kate or he thinks that they hired Mary’

Selected References

- [1] Dylan Bumford. Universal quantification as iterated dynamic conjunction. In *Proceedings of the 19th Amsterdam Colloquium*, Lecture Notes in Computer Science, 2013.
- [2] Marcel den Dikken. Either-float and the syntax of co-or-dination. *Natural Language & Linguistic Theory*, 24(3):689–749, 2006.
- [3] Mana Kobuchi-Philip. Japanese MO: Universal, additive and NPI. *Journal of Cognitive Science*, 10(2):172–194, 2009. <http://semanticsarchive.net/Archive/mIxZDk0N/>.
- [4] Benjamin Slade. *Formal and Philological Inquiries Into the Nature of Interrogatives, Indefinites, Disjunction, and Focus in Sinhala and Other Languages*. Ph.D. dissertation, University of Illinois, 2011. <http://semanticsarchive.net/Archive/TdhYTM3Y/>.
- [5] Anna Szabolcsi. What do quantifier particles do? <http://ling.auf.net/lingbuzz/001857>, 2013.
- [6] Anna Szabolcsi, James Whang, and Vera Zu. Quantifier words and their multi-functional(?) parts. *Language and Linguistics*, 15(1), 2013. <http://semanticsarchive.net/Archive/GIxYWYyY/>.
- [7] Yoad Winter. Syncategorematic conjunction and structured meanings. In *Proceedings of Semantics and Linguistic Theory (SALT) 5*, pages 387–404, 1995. <http://elanguage.net/journals/salt/article/view/5.387>.

Epistemic Indefinites: Are We Ignorant About Ignorance?

Luis Alonso-Ovalle¹ and Paula Menéndez-Benito^{2*}

¹ McGill University, Montréal, Québec, Canada

luis.alonso-ovalle@mcgill.ca

² University of Göttingen, Göttingen, Germany

Paula.Menendez-Benito@phil.uni-goettingen.de

Abstract

Epistemic indefinites make an existential claim and convey that the speaker does not know which individual makes this claim true. The account put forward by Aloni and Port [3, 2] —which we will dub ‘the Lack of Relevant Identification Approach’— defends that ‘not knowing who’ means that the speaker cannot identify the individual that satisfies the existential claim in a contextually relevant way. The Variation Approach (see, e.g., Alonso-Ovalle and Menéndez Benito [5], Chierchia [7], Fäläuš [9], , Giannakidou and Quer [16]) assumes that ‘not knowing who’ means that that individual is not the same in all of the speaker’s epistemic or doxastic alternatives. In this paper, we will argue that the behaviour of Spanish *algún* presents challenges for both approaches. Our conclusion will be that we still lack a clear understanding of what the ignorance component conveys.

1 Epistemic Indefinites: Ignorance and Knowing Who.

Epistemic indefinites signal ignorance on the part of the speaker. This epistemic effect can be illustrated with the Spanish sentences in (1). Both (1-a), with the non-epistemic indefinite *un*, and (1-b), with the epistemic indefinite *algún*, convey that there is a linguistics student that María is dating. However, only (1-a) is consistent with the speaker knowing which student María is dating. Thus, the continuation “and I know who” is odd in (1-b).¹

- (1) a. María sale con un estudiante de lingüística, y yo sé quién es.
María goes out with UN student of linguistics and I know who is
'María is dating a linguistics student and I know who.'
- b. María sale con algún estudiante de lingüística, # y yo sé quién es.
María goes out with ALGÚN student of linguistics and I know who is
'María is dating some linguistics student and I know who.'

This much is clear. But what exactly counts as ‘knowing who’? Alonso-Ovalle and Menéndez Benito [4] noticed a contrast between *algún* and the epistemic indefinite *some*. Suppose L and P are visiting the Math department. They don’t know anything about the people working there, and they haven’t seen any of them before. They suddenly see an individual, who can be inferred to be a professor, frantically dancing lambada on his desk. In this scenario, P can felicitously utter (2) but not (3).

*Thanks to the members of the McGill Syntax/Semantics Research Group and the Göttingen English Linguistics Oberseminar for very useful feedback. Our names are listed in alphabetical order. Financial support: Social Sciences and Humanities Research Council of Canada (Insight Grant 435-2013-0103) and *Fonds de Recherche Société et Culture du Québec* (2013-NP-164823 and 2012-SE-144646.)

¹We will translate *algún* as *some*, although, as shown below, the two indefinites differ in their interpretation.

- (2) Look! Some professor is dancing on the table! [4, p.4]
- (3) # ¡Mira! ¡Algún profesor está bailando encima de la mesa!
 Look! ALGÚN professor is dancing on of the table
 'Look! Some professor is dancing on the table!' [4, p. 4]

There is a sense in which L and P know who the professor is. They can visually identify him, and they can point to him. This type of ‘knowing who’ blocks the use of *algún*, but not the use of *some*. Aloni and Port [3] note that the contrast also holds between German *irgendein* (which behaves like *some*) and Italian *un qualche* (which behaves like *algún*). In view of examples like this, we can conclude that different epistemic indefinites are sensitive to different types of knowledge. This conclusion raises questions such as: (i) what counts as ‘knowing who’ for different epistemic indefinites?, (ii) what is the range of cross-linguistic variation? and, (iii) how are the constraints on ‘knowing who’ encoded?

A substantial part of the literature on epistemic indefinites assumes that the epistemic effect can be characterized —across the board— in terms of variation across the speaker’s epistemic/doxastic alternatives (see, among others, Alonso-Ovalle and Menéndez-Benito [4, 5], Chierchia [6, 7], Fălăuș [9, 10], Giannakidou and Quer [16]). Under this view, which we will refer to as ‘the Variation Approach,’ epistemic indefinites convey that the individual satisfying the existential claim is not the same in all of the speaker’s epistemic/doxastic alternatives.² Recently, Aloni and Port ([3, 2]) have put forward a more fine-grained proposal, which we will dub ‘the Lack of Relevant Identification Approach’, that deals with the questions above by explicitly targeting the issue of what counts as ‘knowing who’ for different epistemic indefinites.

Our goal in this paper is rather modest. Using some hitherto unobserved data, we show that the epistemic effect of *algún* cannot be straightforwardly captured by either of the approaches above. On the one hand, the behaviour of *algún* presents a challenge for the Lack of Relevant Identification Approach. On the other, the Variation Approach is too coarse-grained to capture the full range of data. We tentatively suggest a way in which the Variation Approach could be refined to rise to the challenge, but we will leave an investigation of this idea for future research.

2 Ignorance as Lack of Relevant Identification.

We start by summarizing the core components of the proposal presented by Aloni and Port [3, 2].

Aloni and Port build on Aloni’s [1] observation that ‘knowing who’ is sensitive to context-dependent identification methods. Aloni [1] illustrates this with the context in (4).

- (4) In front of you lie two face-down cards. One is the ace of spades, the other is the ace of hearts. You know that the winning card is the ace of hearts, but you don’t know whether the ace of hearts it’s the card on the left or the card on the right. ([3],[1, p. 16])

Is (5) true in (4)? It depends on the method of identification chosen: (5) is true if cards are identified by their suit (identification by description) but false if they are identified by their position (identification by ostension).

- (5) You know which card is the winning card. ([3], cf. [1, p. 16])

²Variation approaches differ with respect to the source of the variation component. Two lines have been pursued: (i) the variation component is derived as a quantity implicature (e.g., [4, 5, 6, 9, 10, 7]); (ii) the variation component is hard-wired (for instance, in [16] this component is treated as a felicity condition.) In this paper we will focus on the content of the epistemic effect and stay away from the issue of how it is derived.

Building on this observation, Aloni and Port assume that the context provides a relevant way of ‘knowing who’ (a relevant identification method) and claim that epistemic indefinites signal that the speaker cannot identify the witness of the existential claim using that method. Assume, for instance, that being able to name individuals counts as knowing who they are in a particular context. In that context, we should be able to use the sentence in (6) if we cannot name the professor — even if we can point at him.

- (6) Look! Some professor is dancing on the table! [4, p.4]

In Aloni and Port’s approach, methods of identification are modelled as conceptual covers [1]. A conceptual cover CC is a set of individual concepts (functions from worlds to individuals) $\{i_1, i_2, \dots\}$ that jointly ‘cover’ the domain of quantification (in any w , each individual concept is true of one individual, and in any w each individual is picked out by one of these individual concepts). The use of an epistemic indefinite depends on conceptual covers in the following way.³ Suppose that there are two professors, Professor Smith and Professor Jones. A sentence like (7) could in principle be interpreted with respect to the two covers in (8). (7) signals that the speaker can identify the professor with respect to some cover CC (9), and that CC is not the cover made salient by the context.

- (7) Mary is dating some professor.
- (8) $\{\lambda w. \iota x. \text{TO-THE-RIGHT}_w(x), \lambda w. \iota x. \text{TO-THE-LEFT}_w(x)\}, \{\lambda w. \text{Smith}, \lambda w. \text{Jones}\}$
- (9) There is at least one i in CC such that for all w compatible with what the speaker believes, Mary is dating $i(w)$.

To account for the contrast between *algún* (2) and *some* (3), and for a parallel contrast between Italian *un qualche* and German *irgendein*, Aloni and Port assume the existence of the hierarchy of methods of identification in (10) and the principle in (11).

- (10) ostension >_{higher than} naming >_{higher than} description
- (11) In Romance, but not in Germanic, the identification method required by knowledge must be higher in order [in (10)] than the identification method required for epistemic indefinites.

Together, (10) and (11) predict that epistemic indefinites in Romance are incompatible with pointing (as ostension is the highest method in (10)). Thus, (2) is ruled out, but (3) will be acceptable if ostension is not the relevant identification method.

3 The Challenge of *Algún*

Giannakidou and Quer [16] argue that *algún* is incompatible with *all* the methods of identification considered by Aloni and Port [3]. They support their claim with the examples in (12) through (14).⁴ The example in (12) illustrates that *algún* is incompatible with ostension (as in Alonso-Ovalle and Menéndez-Benito’s [4] example). The example in (13) shows that *algún* is incompatible with naming the individual, and (14), that it is incompatible with identification

³This is a simplification. Aloni and Port’s theory is cast in a dynamic semantics with conceptual covers [1]. What we present here is an informal rendition of (part of) their proposal. The reader is referred to [3] for the details of the technical setup.

⁴They also discuss parallel facts for Greek *kapjios*.

by description. Giannakidou and Quer [16] conclude from this that Aloni and Port's account is not tenable.

- (12) Tengo que leer un artículo de algún profesor. # Es aquel señor de allí.
have:1s read an article of ALGÚN professor is that guy of there
'I have to read an article of some professor or other. It's that guy over there.' [16, p.140]
- (13) Tengo que quedar con algún profesor. # Se llama Bob Smith.
have:1s meet with ALGÚN professor SE is-named Bob Smith
'I have to meet some professor or other. # His name is Bob Smith.' [16, p.140]
- (14) Tengo que quedar con algún profesor. # Es el director del Dpt. de Filosofía.
have:1s meet with ALGÚN professor is the director of-the Dept of Philosophy
'I have to meet some professor or other. He is the director of the Ph. Dpt.' [16, p.140]

Note, however, that for this argument to be complete, we would need to provide relevant scenarios for the examples above: as we have seen, on Aloni and Port's account, epistemic indefinites signal that the speaker cannot identify the witness of the existential claim by *the contextually relevant method*. Thus, examples like (13) and (14) will only challenge Aloni and Port's account if they cannot be uttered in contexts where the salient identification method is not the one available to the speaker.⁵ That seems to be indeed the case. According to our intuitions, (13) is deviant in a context where ostension is relevant but not available to the speaker (e.g., in a situation where we are looking for the professor in a crowded room, and pointing at him would be the most effective way of finding him, but we cannot do so (see the contexts in Aloni and Port [2])), and (14) would still be ruled out in a scenario where naming is relevant but not available to the speaker (for instance, if we arrive at the Philosophy department, and we are looking at a series of doors with the professors' names on them, but we don't know the name of the professor that we are searching for).

While we agree with Giannakidou and Quer [16] that these judgments pose a problem for the Lack of Relevant Identification Approach, we think that their claim that *algún* is incompatible with *all* methods of identification is too strong. Surprisingly, *algún* is not always incompatible with pointing. Consider, as illustration, the following contrast. Suppose that P looks out of the window and she sees María kissing a boy. If the circumstances are as in (15), P cannot felicitously utter (17). However, if they are as in (16), P can felicitously utter (17) while pointing at the boy.

- (15) *Clear vision:* P hasn't seen the boy before, but she can see him very clearly now.
- (16) *Blurry vision:* María and the boy are far away. P can see that María is kissing a boy, but she cannot make out the boy's features.
- (17) ¡Mira! ¡María está besando a algún chico!
Look! María is kissing a ALGÚN boy!
'Look! María is kissing some boy.'

This type of contrast is problematic for the Lack of Relevant Identification Approach.⁶ Like the examples in (13) and (14) above, the contrast between (15) and (16) strongly suggests that *algún* is not sensitive to what method of identification is relevant in the context. It is not clear what method of identification would be required in (15) and (16), but whatever it

⁵Example (12) is actually not problematic for Aloni and Port, as they predict that ostension rules out epistemic indefinites in Romance. But see below.

⁶Slade [17] shows that the same pattern obtains for the Sinhala epistemic indefinite *wh-hari*.

might be, there is no reason to assume a difference between the two contexts. Additionally, the acceptability of (17) in (16) shows that identification by ostension does not necessarily rule out epistemic indefinites in Romance.

In what follows, we show why the contrast is also problematic for those approaches that characterize the ignorance effect in terms of variation of witnesses across the speaker's epistemic/doxastic alternatives.

4 The Challenge for the Variation Account.

As noted above, several accounts of epistemic indefinites assume that the epistemic effect can be captured by a condition that requires variation of the individual(s) satisfying the existential claim across the speaker's epistemic/doxastic alternatives. Alonso-Ovalle and Menéndez-Benito [5], for instance, model the epistemic effect of *algún* as in (18) below.

- (18) a. LF: *algún* (P)(Q)
 - b. $\exists w', w'' \in \mathcal{D}_w [\{x : P(w')(x) \& Q(w'(x))\} \neq \{x : P(w'')(x) \& Q(w'')(x)\}]$
- (\mathcal{D}_w is the set of worlds compatible with the speaker's evidence in w) [5, p. 38]

The formulation in (18-b) follows closely an unpublished suggestion that von Fintel [18] made for *some*. In turn, this suggestion follows von Fintel's own characterization of the ignorance component of *whatever* [19], which reformulates an earlier proposal by Veneeta Dayal [8].⁷

Heller and Wolter [11] argue that an account along these lines does not capture the ignorance effect of *whatever*. In what follows, we will summarise their argument and note that, in view of the examples presented in Section 3, this argument also applies to *algún*. Therefore, the formulation in (18-b) cannot fully capture the epistemic effect of this item.

The example in (19) illustrates the ignorance reading of *whatever* free relatives —the sentence conveys that the speaker does not know what Arlo is cooking.⁸

- (19) There's a lot of garlic in whatever (it is that) Arlo is cooking. [19, p. 27]

Building on Dayal's [8] analysis of *whatever*, von Fintel [19] models this ignorance component as a variation requirement. His analysis is summarized in (20).

- (20) a. LF: *whatever* (w)(F)(P)
- b. denotes: $\iota x.P(x)$
- c. presupposes: $\exists w', w'' \in F[\iota x.P(w')(x) \neq \iota x.P(w'')(x)]$ [19, p. 28]

In line with previous work, von Fintel [19] treats *whatever* free relatives as definite descriptions (20-b). Additionally, he proposes that *whatever* presupposes that the individual that the definite description picks out is not the same in all worlds in the modal base F (20-c). In cases like (19), F is the set of epistemic/doxastic alternatives of the speaker (or the hearer). Given this, the example in (19) will assert that that there is a lot of garlic in the thing that Arlo is cooking and presuppose that this thing *is not the same* in all epistemic/doxastic alternatives.

⁷For other, slightly different, ways of formulating the variation condition, see [6], [7], [9], and [16]. The epistemic component derived in [6, 7] and [9] requires that there be at least two (different) *actual* individuals d_1, d_2 in the domain of quantification and at least two (different) accessible worlds w_1, w_2 such that d_1 has property Q in w_1 and d_2 has the property Q in w_2 . The formulation poses the problem of how to make sure that the relevant accessible worlds share the same individuals. Space constraints prevent us from discussing the difference between these formulations and ours.

⁸Ignorance readings can also be anchored to the epistemic state of the hearer. See [19] for discussion.

Heller and Wolter [11] present a number of arguments against this view. One of them has to do with the possibility of identifying the individual that *whatever* picks out by ostension. They note that *whatever* can be used in situations where the speaker can point at the individual that *whatever* picks out in the actual world. They illustrate this with the context in (21).

- (21) “Suppose Becca enters the kitchen and sees Bob stirring a pot. Becca cannot say what dish Bob is cooking, but she can point out the stuff that he is stirring.” [11, p. 174]

In the situation in (21), Becca would be able to use a demonstrative pronoun, as in (22). If she can infer from the smell that one of the ingredients of the dish that Bob is cooking is onion, she will also be able to utter (23).

- (22) Bob is cooking that. [11, p. 174]
 (23) Whatever Bob is cooking uses onions. [11, p. 174], after [8] (ex. 27)

Heller and Wolter [11] note that this argues against an account based on strict identity across worlds. Demonstratives, like proper names, are assumed to be rigid designators (Kaplan [12]). Given that the speaker can use a demonstrative to refer to the thing that Bob is cooking, we can assume that the free relative in (23) picks out the same individual across the speaker’s epistemic/doxastic alternatives. If this is so, the variation condition is not met. Yet (23) is fine.

Similar remarks apply to the case of *algún*. In both the clear vision and the blurry vision contexts presented in Section 3, the speaker can point at the witness of the existential claim, and would therefore be able to use a demonstrative to refer to the individual. A condition like (18-b), based on strict identity, would thus predict *algún* to be infelicitous in both cases. As we have seen, this is contrary to fact.

But perhaps we should move away from strict identity and rely on similarity. In both contexts above, the individuals that María is kissing across the speaker’s doxastic/epistemic alternatives are similar in some way. In the clear vision scenario, these individuals share their physical appearance (although they can vary wildly with respect to many other properties, e.g., their histories or personalities). In the blurry vision scenario, they share, at the very least, the same physical location at the time of the kissing, relative to the speaker (but can also vary in many other respects, including now their physical appearance). Formulating the variation condition in terms of the types of similarity that *algún* allows for might give the Variation Approach a way to handle the contrast.

Adopting a Lewisian ontology gives us a way to do this. According to Lewis [13, 14, 15] individuals only exist in one world. Cross-world identity is modelled via counterpart relations, relations of comparative similarity. As an illustration, consider the modal sentence in (24). Suppose that Humphrey is an actual individual. Since individuals are world-bound, (24) cannot be paraphrased as “there is an accessible world *w* such that Humphrey wins at *w*”. This could never be true, since Humphrey is not even part of *w*. Instead, (24) should be paraphrased as “there is an accessible world *w* such that *some counterpart of Humphrey* wins at *w*.”

- (24) Humphrey might win. cf. [15, p.9 and ff.]

Counterpart relations are similarity relations. For an individual *x* to be counterpart of an individual *y*, *x* and *y* have to be substantially similar in some way. But individuals can resemble each other in many different ways. There are many possible similarity relations, and as a result, there are many possible counterpart relations.

The Variation Approach might be able to handle the contrast between the scenarios in (15) and (16) by formulating the variation condition in terms of counterpart relations. As noted

above, in both of those scenarios scenarios the individuals that María is kissing are similar in some way, and therefore can be considered counterparts under some similarity relation.⁹ One way to account for the contrast might be to claim that a sentence like (17) above (roughly) conveys (25), where *algún* would constraint what counts as a suitable counterpart relation R so as to rule out the counterpart relation in the clear vision scenario but allow the one in the blurry vision scenario.

- (25) There are at least two doxastic / epistemic alternatives of the speaker's, w_1 and w_2 , such that the boy that María is kissing in w_1 is not a counterpart of the boy that María is kissing in w_2 *under a suitable counterpart relation \mathcal{R}* .

This is where the challenge to the Variation Approach lies. The tasks ahead for a variation account would be (i) to specify what counterpart relations rule out the use of *algún*, and (ii) to provide an account of how exactly *algún* imposes those restrictions. We are not going to undertake these tasks here, but limit ourselves to noting that the first task is far from trivial. Let us illustrate why with a couple of examples.

In view of the examples in (15) and (16) one might think that what matters is the degree of perceptual acquaintance that the speaker has with the witness. But consider the following context. You and I are looking out the window in a hotel room. We arrived to the hotel at night, and we haven't yet seen any of the surroundings. The window overlooks a square with a statue. We can see the statue, but we cannot see its features clearly. All we can make out is that it is a statue, and that it represents a human figure. Suddenly, we see our friend María approach the statue and hug it. In this new blurry vision context, I would not be able to utter (26), even though I have the same degree of perceptual acquaintance with the witness of the existential claim as the speaker in (16) above.

- (26) María está abrazando (a) alguna estatua.
 María is hugging A ALGUNA statue
 'María is kissing some statue.'

One might also think that what matters is whether the counterparts share stable properties or not. Location is a (relatively) stable feature of a statue but not of a human being. But this cannot be it, either. For suppose that we attended a short meeting of a student committee yesterday morning. We know that, afterwards, María kissed the chair of that committee. Suppose, furthermore, that the individual in question held the property of being the chair of the committee for a very short period of time, only for the duration of the meeting. A speaker that has this information would not be able to utter (27).

- (27) Ayer María besó a algún estudiante.
 yesterday María kissed A ALGÚN student
 'Yesterday, María kissed some student.'

5 To Conclude: Are We Ignorant About Ignorance?

In view of the previous discussion, we conclude that we still do not have a firm grasp on the content of the ignorance effect triggered by *algún*. The 'clear vs. blurry vision' cases presented

⁹In the blurry vision scenario, the relevant individuals might be what Lewis [14, p. 379] calls *counterparts by acquaintance*: "Counterparts by acquaintance (...) are united by resemblance in the relation to a subject of attitudes."

in Section 3 cast doubts on the Lack of Relevant Identification Approach, according to which epistemic indefinites signal that the speaker cannot identify the witness of the existential claim by resorting to the contextually relevant method of identification. But these examples are also challenging for a Variation Approach that formulates variation in terms of similarity, since it is not immediately clear how to determine what similarity relations rule out the use of *algún*.

References

- [1] Maria Aloni. *Quantification under Conceptual Covers*. PhD thesis, University of Amsterdam, 2001.
- [2] Maria Aloni and Angelika Port. Epistemic indefinites and methods of identification. In Luis Alonso-Ovalle and Paula Menéndez-Benito, editors, *Epistemic Indefinites*. Oxford University Press, forthcoming.
- [3] Maria Aloni and Angelika Port. Epistemic indefinites crosslinguistically. In *Proceedings of NELS 41, 2010*, forthcoming.
- [4] Luis Alonso-Ovalle and Paula Menéndez-Benito. Some epistemic indefinites. In Makoto Kadowaki and Shigeto Kawahara, editors, *Proceedings of the North East Linguistic Society, 33*, pages 1–12. GLSA, Amherst, MA, 2003.
- [5] Luis Alonso-Ovalle and Paula Menéndez-Benito. Modal indefinites. *Natural Language Semantics*, 18(1):1–31, 2010.
- [6] Gennaro Chierchia. Broaden your views. Implicatures of domain widening and the ‘logicality’ of language. *Linguistic Inquiry*, 37(4):535–590, 2006.
- [7] Gennaro Chierchia. *Logic in Grammar*. Oxford University Press, Oxford, 2013.
- [8] Veneeta Dayal. Free relatives and ever: Identity and free choice readings. In Aaron Lawson, editor, *Proceedings of SALT VII*, pages 99–116, Ithaca, NY, 1997. CLC Publications.
- [9] Anamaria Fălăuș. *Polarity Items and Dependent Indefinites in Romanian*. PhD thesis, Université de Nantes, 2009.
- [10] Anamaria Fălăuș. Alternatives as sources of semantic dependency. In Nan Li and David Lutz, editors, *Semantics and Linguistic Theory (SALT) 20*, pages 406–427. eLanguage, 2011.
- [11] Daphna Heller and Lynsey Wolter. On identification and transworld identity in natural language: the case of -ever free relatives. *Linguistics and Philosophy*, 34(2):169–199, 2011.
- [12] David Kaplan. Demonstratives. In Joseph Almog, John Perry, and Howard Wettstein, editors, *Themes from Kaplan*, pages 481–583. Oxford University Press, 1989.
- [13] David Lewis. Counterpart theory and quantified modal logic. *Journal of Philosophical Logic*, 65:113–126, 1968.
- [14] David Lewis. Individuation by acquaintance and by stipulation. *The Philosophical Review*, 92:3–32, 1983.
- [15] David Lewis. *On the Plurality of Worlds*. Malden, MA, 1986.
- [16] Josep Quer and Anastasia Giannakidou. Exhaustive and non-exhaustive variation with free choice and referential vagueness: Evidence from Greek, Catalan, and Spanish. *Lingua*, 126:120 – 149, 2013.
- [17] Benjamin Slade. Sinhala epistemic indefinites with a certain *je ne sais quoi*. In Luis Alonso-Ovalle and Paula Menéndez Benito, editors, *Epistemic Indefinites*. Oxford University Press, forthcoming.
- [18] Kai von Fintel. Whatever. Class notes, 24.979 Topics in Semantics, Spring 1999, 1999.
- [19] Kai von Fintel. Whatever. In Brendan Jackson and Tanya Matthews, editors, *SALT X*, pages 27–40, Ithaca, NY, 2000. CLC Publications.

Correlating Cessation with Double Access

Daniel Altshuler¹ and Roger Schwarzschild²

¹ Heinrich-Heine-Universität, Düsseldorf

² Rutgers University

daltshul@gmail.com, rogerschwarzschild@gmail.com

Abstract

This paper investigates an inference (which we call *cessation*) that no state of the kind described currently holds. We propose that this inference occurs in a sentence ϕ when: (a) ϕ has stative verb in the past tense and (b) there is a present tense alternative to ϕ that shares a common reference time concept and is not vacuously false. We show how (a) and (b) correlate with the availability of the so-called *double access* reading, found in present-under-past reports and which we analyze by building on an analysis of tense proposed by Musan (1995). The novelty of our analysis is that the present tense in English is an amalgam of both a relative and a deictic present. More concretely, the English present poses presuppositional constraints on the reference time concept which demand truth at the local evaluation time and at or after the speech time. An important consequence of our analysis is the following conjecture: intuitions about so-called *simultaneous readings* in past-under-past reports are really intuitions about the absence of cessation.

1 Introduction

Suppose I run into my friend Sylvia in the mall one day, and she introduces me to her son, saying:

- (1) Here's my son Ralph, he was in the Peace Corps.

From (1), I learn of Ralph's Peace Corps experience. I further infer that Ralph is no longer in the Peace Corps. Altshuler and Schwarzschild (2013) called this: **cessation inference**, i.e. the inference that no state of the kind described currently holds. Although the term is new, the idea is not. Bello (1847: §692), for example, explains the use of the past tense in counterfactuals in terms of the tendency to deduce from a past tense utterance that the corresponding present tense is false.¹

Now consider the example below:

- (2) Jack told me that his mother is American and that his father was Dutch.

From (2), it is natural to infer that Jack's father is no longer alive. This inference begins with a chain of reasoning that starts with a cessation inference. From the past on *was Dutch* we infer that Jack's father is no longer Dutch. Assuming that being Dutch is for life, we deduce that Jack's father must

¹ More recent discussion of cessation inferences is found in, e.g. Musan (1995, 1997), Iatridou (2000: 248), Hogweg (2009: 189-90) and Magri (2011: §5). See also Tonhauser (2007), for a discussion of what she calls 'change of state': the inference that the described state has ended. As noted by Altshuler and Schwarzschild (2013: 48), it is important to keep these two types of inferences separate; cessation means that not only has that state ended, but also that no state of the kind described obtains currently.

have died. This chain of reasoning follows discussions of lifetime effects (see Thomas 2012 and references therein). However, notice that in this case the past tense sentence that triggers the cessation inference, *his father was Dutch*, is embedded. So if the cessation inference involves consideration of a present tense alternative, as we will propose (following Bello's lead), it would have to be an alternative in which you have a present tense clausal complement of a past tensed verb. Such a construction is known to give rise to the so-called *double access* reading, which is quite tricky to analyze. We will discuss double access in Section 3, with the ultimate goal of giving an analysis of the cessation inference in (2). We begin, however, with some more background on cessation inferences.

2 Background on cessation inferences

Musan (1995) treats cessation inferences as a quantity impudicature. She says that PRES- ϕ asymmetrically entails PAST- ϕ , so uttering the latter implicates the negation of the former (\neg PRES- ϕ). Musan (1997: 279) later modifies this view, reasoning that “if Gregory came into existence right now, at this very moment while I utter this sentence, then *Gregory is from America* would be judged true, but *Gregory was from America* would be judged false”. She nevertheless maintains the position that cessation inferences are quantity impudicatures. In turn, this view has been criticized by Thomas (2012), who considers (3) and (4) below and claims: “...the present tense sentence in [(4)] is not stronger than the past tense sentence in [(3)]. Rather, the two sentences are logically independent (pg. 27).” He concludes that a Gricean analysis is therefore not warranted: “...a consistent implementation of Musan’s analysis requires treating the notion of ‘more informative’ in the exploitation of the Maxim of Quantity as logical entailment rather than contextual entailment, therefore isolating the exploitation of the Maxim from contextual information and treating the Maxim more like a linguistic rule than a general principle of rational behavior (pg. 79).”²

- (3) John was a graduate student.
- (4) John is a graduate student.

Contra Thomas, we adopt Musan’s original (1995) claim that sentences like (3) and (4) are logically dependent, though we do not adopt her semantics of tense used to derive the dependence. Instead we follow Altshuler and Schwarzschild (2013) in assuming that the entailment from PRES- ϕ to PAST- ϕ follows from a general feature of the truth conditions of stative predictions. They argue that if a stative clause like *Greg be American* or *John be a graduate student* is true at a moment m , it is true at some moment m' preceding m . They assume time is dense, so there could be a lower limit on the truth of stative predictions without there being an earliest moment.

An observation that fuels the analysis of tense that we ultimately propose is Musan’s observation that the cessation inference is absent in the following discourse:

- (5) On that day, I was introduced to Gregory and Eva-Lotta. Gregory was from America, and Eva Lotta was from Switzerland (Musan 1997: 272).³

Musan explains the absence in terms of an implicit domain restriction on the past tense. When used in this discourse, *Gregory was from America* is true if there is a time t before now, t is within the time of introduction and Gregory-be-from-America is true at t . The present tense alternative maintains the implicit domain restriction and so it amounts to: there is a time t , t includes now, t is within the time of introduction and Gregory-be-from-America is true at t . Given that the time of introduction wholly

² For more insightful discussion of Musan’s (1997) analysis, see Magri (2009, 2011).

³ See Klein (1994: 4) for similar examples.

precedes the utterance time, these conditions could not be met regardless of Gregory's origins – so negating this alternative adds nothing and no contentful implicature is generated.⁴

We embrace this explanation of (5), differing from Musan in that we take tense domain restrictions to be intensional: properties of times, not times themselves. We call these restrictions **reference time concepts (RefTimeConcepts)**. (6) illustrates the intuition that leads to this move:

- (6) Everyone was unusually friendly at the 6th Annual Rowers Meeting. Jack thought that the punch was spiked. Jill thought that the brownies were loaded. In fact, the air was artificially oxygenated.

Every instance of *be* in this discourse is ‘about’ the same time. Intuitively the following descriptive restriction applies to each past time:

- (7) $\lambda w \lambda t. t$ is during the 6th Annual Rower's Meeting in *w*

This restrictor, we claim, is embedded under the propositional attitude verbs in (6). The actual interval of time that the concept picks out will depend on Jack's and Jill's belief states.

These, then, are the ingredients for a cessation inference:

- (8) Ingredients for a cessation inference

- a. A sentence ϕ with a stative verb in the past tense.
- b. A present tensed alternative to ϕ that shares a common RefTimeConcept and is not vacuously false.

We now turn to show how (8) correlates with the so-called *double access* reading. Crucial to our analysis will be the idea that the English present tense poses presuppositional constraints on the RefTimeConcept, making the tense an amalgam of both a relative and a deictic present.

3 Double Access Reading

Suppose that back at the mall, I ask Sylvia where her friend, Mary, is. She replies: “Mary is at home today”. Later that day, when I'm at the beach and asked for Mary's whereabouts, I can truthfully say:

- (9) Sylvia said that Mary is at home.

Taking our cue from Abusch's (1997) discussion of present complements of *believed*, we observe that (9) is true on the so-called *double access* reading because two conditions are met.⁵ To describe those conditions, we will need to refer to the time and world at which Sylvia replied to me at the mall. We symbolize those as $w@$ and t_{mall} respectively. Below, in (10)-(11), we first give each condition in descriptive terms, and then in more analytical terms, presupposing a theory in which *say* quantifies over world-time pairs compatible with what was said.

- (10) Relative Present Condition

- a. If Sylvia's utterance was true, then Mary is at home in $w@$ at t_{mall} .
- b. For any $\langle w, t \rangle$ compatible with what Sylvia said: Mary is at home in *w* at *t*.

⁴ Cf. Magri (2009, 2011) and Thomas (2012) who propose that the present tensed counterpart is not relevant (rather than not being a viable alternative to its past tensed counterpart).

⁵ The name ‘double access’ reflects the fact that the truth-conditions of (9) can be divided into two key components. Discussions of this reading go back to (at least): Smith (1978), Comrie (1985), Enç (1987) and Ogihara (1989, 1995).

(11) Deictic Present Condition

- a. If Sylvia's utterance was true, then Mary is at home throughout an interval that includes the time at which (9) is uttered.
- b. There is a description δ such that when δ is evaluated in $w@$ at t_{mall} it picks out an interval that includes the time at which (9) is uttered. For any $\langle w, t \rangle$ compatible with what Sylvia said: δ evaluated at $\langle w, t \rangle$ picks out an interval throughout which Mary is at home in w .

We call (10) ‘Relative Present Condition’ because it makes the present tense look like a relative present – a sort of Priorian present tense (Prior 1967) that picks out those times compatible with what the subject said. Following Absch (1997), we refer to such times as *the local evaluation time*. In order to appreciate its force, imagine that prior to the election, Clinton said: “Obama will be the winner in November”. After the election in November it would be odd to report that as “Clinton said that Obama is the winner”.

We call (11) ‘Deictic Present Condition’ because it makes the present tense look like a deictic present – a tense that in any context picks out the utterance time of that context. Note that (11)b requires the utterance time to be part of an interval that is picked out by a description δ . We assume that this description is the RefTimeConcept that comes with any tense (recall the discussion in the previous section). In (9), the RefTimeConcept would presumably correspond to *today*, which was part of Sylvia’s actual utterance in the mall (i.e. “Mary is at home today”). This would explain why (9) is true if uttered on the same day as t_{mall} but not if uttered on the following day.

In sum we propose that the present tense in English is an amalgam of both a relative and a deictic present. More concretely, we propose that the English present demands truth at the local evaluation time (relative tense component) and at or after the speech time (deictic tense component). In a simple present tense clause like (4) above, the local evaluation time is the speech time so the two components cannot be told apart. The two components also cannot be distinguished when the present is embedded under *will*:

(12) Jack will arrive with a child who is crying.

The local evaluation time for *is* in (12) is a future time introduced by *will*. Since a future time is at or after the speech time, again the two components are indistinguishable. However, when the present is embedded under a past attitude verb, which controls the local evaluation time, the two components come apart.

We demonstrate this by providing the meaning of the complement of *said* in (9) above, namely *Mary is at home*. This meaning is given in (13) below. Note that s^* and w^* denote the speech-time and the speech-world of (9) respectively.

$$(13) \lambda t_0 \lambda w_0: t_0 \in \text{RTC}_1(w_0, t_0) \wedge [\exists t' t' \geq s^*] t' \in \text{RTC}_1(w^*, t_{\text{mall}}). \forall t (t \in \text{RTC}_1(w_0, t_0) \rightarrow \text{be.at.home}(w_0, t, \text{mary}))$$

The formula above treats the aforementioned relative and deictic components of the present tense as presuppositional constraints on the RefTimeConcept: RTC_1 must include the local evaluation time t_0 , as well as some time that is at or after s^* . The present tense also contributes a universal statement to the assertion, namely that Mary’s state of being at home in w_0 holds throughout the time interval t described by RTC_1 in w_0 at t_0 .⁶

The meaning in (13) allows us to mimic Abusch’s (1997) analysis of double access via a new route. Before discussing some possible advantages of our proposal, we note a possible disadvantage. Notice that the presupposition in (13) includes the formula:

⁶ Following the standard terminology, we take t to be the reference time (Reichenbach 1947) or the topic time (Klein 1994).

- (14) $[\exists t' t' \geq s^*] t' \in \text{RTC}_1(w^*, t_{\text{mall}})$

The question that arises for our analysis is: how does t_{mall} get into the truth conditions? It turns out that if that index is filled in by the speech time, the present becomes or can be just a simple relative present. For the most part, the data does not support this option for English. Therefore, we have to stipulate that t_{mall} gets into the truth conditions via binding by the higher tense:

- (15) Sylvia PAST λt told me that Mary PRES_{w@, t} is at home.

We leave it for further research as to how to best motivate (15), and offer the following truth conditions of (9)⁷ that assume it:

- (16) a. $\lambda w_0. \exists t (t < s^* \wedge t \in \text{RTC}_1(w_0, s^*) \wedge \text{say}(\text{sylvia}, w_0, t, \phi))$, where $\phi = (16)\text{b}$
b. $\lambda t_0 \lambda w_0. t_0 \in \text{RTC}_1(w_0, t_0) \wedge [\exists t' t' \geq s^*] t' \in \text{RTC}_1(w^*, t). \forall t'' (t'' \in \text{RTC}_1(w_0, t_0) \rightarrow \text{be.at.home}(w_0, t'', \text{mary}))$

Given Sylvia's original utterance at the mall, namely "Mary is at home today", we assume that RTC₁ in the formula above is the *today-function*: it assigns to any w, t , the set of times t' that are on the same day as t . Given the universal statement in the assertion of (16)b, it must have been compatible with what Sylvia said at the mall that Mary continued to be at home throughout the day – in fact, the whole day – not just after Sylvia's utterance at the mall, but also at s^* , i.e. the speech time of (9), when the speaker of (9) is on the beach. This is the hallmark of the double access reading.

Let us now turn to show why the presupposition in (16)b is satisfied. The first conjunct of (16)b is satisfied since t_0 occurs during the day in which t_0 occurs. The second conjunct of (16)b is satisfied as long as s^* (i.e. the speech time of (9), when the speaker of (9) is on the beach) is no later than the end of the day when t , the time of Sylvia's utterance at the mall, occurs. The past on *said* already tells us that Sylvia's utterance at the mall precedes s^* . Therefore, s^* has to be between Sylvia's utterance at the mall and the end of the day in which Sylvia's utterance at the mall was made.⁸

We end this section by briefly noting ways in which our analysis differs from Abusch's. On our account the intuition in (10)a comes from the fact that the English present tense is, in part, a relative present. Abusch captures the intuition with the Upper Limit Constraint (ULC), a blanket constraint that forbids locating an event after the local evaluation time. Abusch needs the ULC to constrain de-re pasts – these are examples where a past time is introduced in an extensional context and then is re-used in an intensional context.⁹ Instead of de-re pasts, we build on Altshuler (2008) and think of such cases in terms of sharing RefTimeConcepts analogous to (6), discussed in Section 2. We also think the ULC is too strong. The interaction at the Air Berlin baggage counter below uses a true de-re present and it violates the ULC:

- (17) a. Customer: I believe you have my bags.
b. Employee: Who said I have your bags?
c. Customer: The stewardess told me you have my bags.
d. Employee: When did she tell you that?
e. Customer: On the flight.

⁷ As discussed in section 2, we assume a meaning of the past tense that is essentially what Musan proposes, with the caveat that we take tense domain restrictions to be intensional.

⁸ Cf. Heim (1994: 157), where presupposition failures are discussed in light of examples similar to (9) uttered in difference contexts. Space constraints prevent us from showing that our analysis can duplicate Heim's results.

⁹ For example in *The defendant was at home at the time of the crime...the jurors clearly believed that he was in the laboratory building*, the tense embedded under *believed* is anaphoric to the time introduced in the first sentence, the time of the crime (see Abusch 1997: 2-3).

In a typical de-re fashion, the customer uses the present tensed verb *have* in (17)c to speak about a time that is present from his and the employee's perspective, but would have been future from the stewardess' perspective. The bag-having is future relative to the local evaluation time set by *told* – violating the ULC.

In order to account for the use of the de-re present in (17)c, our analysis would assume de-re movement such that the present tense would be evaluated relative to the speech time. In this way, the relative tense component of the present tense would not be in play. Therefore, unlike Abusch, who would post de-re movement for (9), we would posit de-re movement only in cases such as (17)c; what's crucial to account for (9), we believe, is the sharing of RefTimeConcepts.

Let's return now to the deictic condition in (11)b. It is part of the genius of an Abuschian account that the description δ connects the actual context of the speech-report in (9) with the content of Sylvia's utterance. But for this to work, δ must be part of what Sylvia said. In the case examined above, δ was explicitly in the utterance – identified as *today*. But Sylvia might have simply said: “Mary is at home” which would still warrant the report in (9). What would δ be in this case? Note that in Abusch's treatment of double access in belief contexts, the description comes in as an acquaintance relation for a de-re present tense and δ is part of the belief. That is unobjectionable. The problem comes when we try to extend her analysis to speech-reports.¹⁰

This problem does not arise on our account because δ is the RefTimeConcept that comes with any tense. Thus, if Sylvia's original utterance were simply “Mary is at home”, our analysis of (9) would posit a present tense that includes an implicit δ . That is, a RefTimeConcept that describes some salient time interval (which may be accommodated).

The final problem for Abusch's analysis that we mention is that there is no account of what counts as a proper acquaintance relation. As such, some choices—which are allowed by her analysis—lead to incorrect results (see Heim 1994: fn28).¹¹ This problem translates in our account to the problem of constraining the choice of a RefTimeConcept. In other words, our analysis differs from Abusch in that we owe an account of RefTimeConcept choice, while Abusch owes an explanation of what counts as a proper acquaintance relation.¹²

4 Cessation in embedded contexts

We now return to the cessation implicature discussed at the outset and repeated below:

- (2) Jack told me that his mother is American and that his father was Dutch.

The first conjunct embeds a present tense under a past tensed verb. To be felicitous, the RefTimeConcept on *is* must describe an interval that includes the utterance time of (2). Assuming that the RefTimeConcept is shared across the conjuncts, the RefTimeConcept for the second conjunct also

¹⁰ For more discussion of the difference between speech-reports and belief-reports, see Forbes (1997), Brasoveanu and Farkas (2007), and references therein.

¹¹ Thanks to Corien Bary for discussing this point.

¹² Thus far, the choice of RefTimeConcept was aided by considering the temporal adverb in the original utterance prompting an indirect report (recall Sylvia's use of *today* that led to the indirect report in (9)). Moreover, in (6), we simply assumed that the RefTimeConcept for the second and third sentences was provided by the first sentence. This assumption is reasonable if we further assume that RefTimeConcept choice—like many other cases of anaphora resolution—often reduces to satisfaction of constraints imposed by rhetorical relations (e.g. Hobbs 1979, Lascarides and Asher 1993, Kehler 2002). In particular, it seems reasonable to say that the interpretation of the tenses in the second and third sentences of (6) is dependent on the interpretation of the tense of the first sentence because both the second and third sentences stand in a parallel relationship (i.e. they are connected via PARALLEL) insofar as they both relate to first sentence via OCCASION. While this analysis may suffice for (6), it is worth noting that it is unlikely that rhetorical relations could solve the problems mentioned by Heim (1994: fn28).

meets this requirement and so a present tense could have felicitously been used there. Alas, it was not used; the weaker, past tense was used. Assuming a cooperative speaker, it must be that the present could not be used because Jack's original utterance would not have supported it. And assuming Jack was being cooperative, this would mean that "my father is Dutch" would have been false and that is because his father has passed away.

Compare (2) with what happens in (18):

- (18) I met a musician last night. He had a cool accent. He said his father was Dutch and that affected his speech.

Here a cessation implicature is less likely to arise from the past tensed *was Dutch*. We conjecture that the first sentence in (18) sets the RefTimeConcept for subsequent verb phrases. In particular, it restricts tenses to times during my encounter with the musician. This means that the RefTimeConcept on the past tense in *was Dutch* does not satisfy the condition needed for a felicitous present tense alternative. So this works essentially like Musan's example in (5), discussed in Section 2.

5 ‘Backshifted’ and ‘simultaneous’ readings

In our semantics for *say* and *tell*, we quantify over times compatible with what the subject said. We called those times the *local evaluation time*. Examples (2) and (18) above both embed the sentence *his father was Dutch*. In (2), we understand the Dutchness to completely precede the local evaluation time and in (18), it overlaps the local evaluation time. The first pattern is usually dubbed a “backshifted reading”, the second a “simultaneous reading” (Enç 1987: 635). We suggest that the intuition behind the application of these terms is the perception of a cessation implicature in the backshifted case and the non-perception of a cessation inference in the simultaneous case. We think this could explain why the simultaneous reading is harder to find in languages that have a purely relative present tense. Recall that what blocked the cessation inference in (18) was precisely the deictic component of the present tense in English.

As an illustration, consider the Hebrew past-under-past report in (19):

- (19) *ha-more xašav še avi haya acbani.*

The teacher think-PAST that Avi PAST.be anxious
‘The teacher thought that Avi had been anxious’

As the gloss suggests, when a past tensed copula is used in the embedded clause, the impression one gets is that the teacher had a thought he might have expressed by saying: “You were anxious”, insinuating cessation of anxiety. We suggest that such is the case because the present tense in Hebrew is a purely relative tense (Sharvit 2003), which means that—when it is embedded under a past tense—it is compatible with a RefTimeConcept that describes a past interval. Consequently, the Hebrew present is a viable alternative to an embedded past in reports such as (19), thereby triggering cessation and leading to the intuition that (19) lacks a “simultaneous reading”. In English past-under-past report such as (20), however, the embedded present is not a viable alternative to the embedded past because it has a deictic component and, therefore, no cessation is triggered, leading to the intuition that (20) has a “simultaneous reading.”

- (20) The teacher thought that Avi was anxious.

References

- Abusch, D. (1997). "Sequence of Tense and Temporal De Re," *Linguistics & Philosophy* 20: 1-50.
- Altshuler, D. (2008). "Narrative effects in Russian indirect reports and what they reveal about the meaning of the past tense," in T. Friedman, S. Ito (eds.), *Proceedings of Semantics and Linguistic Theory 18*, 19-36. Cornell, CLC Publications.
- Alshuler, D. and R. Schwarzschild (2013). "Moment of Change, Cessation Implicatures and Simultaneous Readings," in E. Chemla, V. Homer and G. Winterstein (eds.), *Proceedings of Sinn und Bedeutung 17*, pp. 45-62. Paris, ENS.
- Bello, A. (1847). *Grammatica de la Lengua Castellana destinada al uso de los Americanos* 1972 Ediciones del Ministerio de Educacion, Caracas Venezuela.
- Brasoveanu, A. and D. Farkas (2007). "Say Reports, Assertion Events and Meaning Dimensions", in G. Alboiu et al. (eds.), *Pitar Moș: A Building with a View. Papers in Honour of Alexandra Cornilescu*, pp. 23-43. Bucharest, Editura Universității din București.
- Comrie, B. (1985). *Tense*. Cambridge – Cambridge University Press.
- Enç, M. (1987). "Anchoring Conditions for Tense", *Linguistic Inquiry* 18(4): 633-657.
- Forbes, G. (1997). "Belief Reports and Speech Reports", in W. Künne et al. (eds.), *Direct Reference, Indexicality and Propositional Attitudes*, 313-320. CSLI. Stanford.
- Heim, I. (1994). "Comments on Abusch's Theory of Tense," in Hans Kamp (ed.) *Ellipsis, Tense and Questions*, pp 143-170. University of Amsterdam.
- Hobbs, J. (1979). "Coherence and coreference", *Cognitive Science* 3, 67–90.
- Hogeweg, L. (2009). "What's so unreal about the Past: Past tense and counterfactuals," in A. Tsangalidis and R. Facchinetti (eds.), *Studies on English Modality in Honour of Frank Palmer*, Bern, Peter Lang.
- Iatridou S. (2000). "The grammatical ingredients of counterfactuality". *Linguistic Inquiry* 31(2), 231–270.
- Kehler, A. (2002). *Coherence, Reference and the Theory of Grammar*. CSLI. Menlo Park, CA.
- Klein, W. (1994). *Time in Language*. Routledge – London.
- Lascarides, A. and N. Asher. (1993). "Temporal interpretation, discourse relations and commonsense entailment", *Linguistics & Philosophy* 16, 437–93.
- Magri, G. (2009) *A Theory of Individual-Level Predicates Based on Blind Mandatory Implicatures. Constraint Promotion for Optimality Theory*. Doctoral dissertation. MIT.
- Magri, G. (2011). "Another argument for embedded scalar implicatures based on oddness in downward entailing environments", *Semantics & Pragmatics* 4(6), 1-51.
- Musan, R. (1995) . *On the temporal interpretation of noun phrases*. Doctoral dissertation, MIT.
- Musan, R. (1997) "Tense, Predicates, and Lifetime Effects," *Natural Language Semantics* 5.3:271–301.
- Ogihara, T. (1989). *Temporal Reference in English and Japanese*. Doctoral dissertation. University of Texas at Austin.
- Ogihara, T. (1995). "Double-Access sentences and Reference to States", *Natural Language Semantics* 3: 177-210.
- Prior, A. (1967). *Past, Present, and Future*. Oxford, Clarendon Press.
- Reichenbach, H. (1947). *Elements of Symbolic Logic*. New York, Garland.
- Sharvit, Y. (2003). "Embedded Tense and Universal Grammar", *Linguistic Inquiry* 34(4), 669-681.
- Smith, C. (1978). "The Syntax and Interpretation of Temporal Expressions in English", *Linguistics and Philosophy* 2(1): 43-99.
- Thomas, G. (2012) *Temporal Implicatures*. Doctoral dissertation. MIT.
- Tonhauser, J. (2007). "Nominal Tense? The meaning of Guaraní nominal temporal markers", *Language* 84(4), 332–342.

Modification of DPs by epistemic modal adverbs

Elizabeth Bogal-Allbritten*

University of Massachusetts Amherst, Massachusetts, U.S.A.
eba@linguist.umass.edu

Abstract

I argue that when epistemic modal adverbs form constituents with DPs, the modal takes scope only over the DP. In contrast with sentences with higher adjoined adverbs, sentences with modal-modified DPs carry actuality entailments. Composition of type $\langle st, st \rangle$ modals and type $\langle se \rangle$ DP intensions is mediated by shifting the DP into a property-type expression by a variant of Partee's IDENT. Precedent comes from work on Concealed Questions (Frana 2006, Schwager 2008) and Grosu's (2003, 2007) analysis of Transparent Free Relatives.

1 Introduction

Epistemic modal adverbs are standardly treated as type $\langle st, st \rangle$ propositional operators. Recently, however, it has been observed that epistemic modal adverbs can apparently modify subpropositional material, including numerals (Zaroukian 2011) and adjectives (Cinque 2010).

- (1) a. **Numeral-modifying:** Mary met with possibly 15 students.
b. **Adjective-modifying:** They found a probably precancerous mole on John's back.

Much less has been said about the modification of DPs by epistemic modal adverbs (2). While Ernst (1984) discusses their syntax — arguing that modal adverbs form constituents with DPs — there has not yet been discussion of how the intension of a DP (3a) can be modified by an epistemic modal adverb with a familiar type $\langle st, st \rangle$ meaning (3b) (Kratzer 1981).

- (2) **DP-modifying:**
- a. Mary ate [DP possibly [DP the most expensive pizza in Amherst]].
 - b. Mary drank [DP perhaps [DP {an / the} American bourbon]].
 - c. Mary is meeting with [DP probably [DP a nurse practitioner]].
- (3) a. $\llbracket \text{the most expensive pizza in Amherst} \rrbracket = \lambda w' \iota y [\text{most expensive pizza in Amherst}(y, w')]$
b. $\llbracket \text{possibly} \rrbracket = \lambda p_{st} \lambda w [\exists w'' \in \text{EPI-MB}(w)[p(w'')]]$
c. $\llbracket \text{IDENTIFY} \rrbracket = \lambda \mathcal{X}_{se} \lambda z_e \lambda w' [z = \mathcal{X}(w')] \quad (\text{Frana 2010})$

This paper puts forth such an analysis. I maintain familiar entries for modal adverbs. I argue that adverbs compose with DPs that are shifted into properties via IDENTIFY (3c), an intensional form of Partee's (1986) IDENT (Frana 2006, 2010). The modal and shifted DP then compose through Function Composition.

Under this account, modal adverbs that form constituents with DPs only take scope over that DP. This prediction is verified by sentence pairs like (4). When the modal adverb takes scope over the verb *ate* as in (4a), the sentence is felicitous in a context where it is unknown whether Mary ate anything. By contrast, epistemic uncertainty introduced by the DP-modifying modal

*I thank Rajesh Bhatt, Seth Cable, Ilaria Frana, Angelika Kratzer, and Barbara Partee for their help and comments. Any remaining mistakes are mine. This work was supported by a NSF Graduate Research Fellowship.

in (4b) is restricted to the identity of the object eaten by Mary: in at least one of the speaker's epistemic alternatives *w*', what Mary ate in *w*' is *the most expensive pizza in Amherst in w'*. It must be the case, however, that Mary ate something.

- (4) Context: Mary visited Amherst yesterday and planned to eat a pizza lunch at Athena's, which you know makes the most expensive pizza in Amherst. You know that Mary anticipated having to skip lunch, however, so perhaps she didn't eat anything. You say,
 - a. Mary [VP possibly [VP ate the most expensive pizza in Amherst.]]]
 - b. #Mary ate [DP possibly [DP the most expensive pizza in Amherst.]]]

The paper proceeds as follows. Section 2 presents empirical evidence for the syntax and truth conditions of modal-modified DPs. Section 3 presents the formal analysis, building in part on analyses of Concealed Questions (Frana 2006, 2010; Schwager 2008) and Transparent Free Relatives (Grosu 2003, 2007). Section 4 concludes and presents directions for further work.

2 Syntactic and Semantic Scope of Modal Adverbs

Against a parenthetical analysis: Although this paper focuses on modification of DPs by epistemic modal adverbs, this position can also be occupied by other typically sentential adverbial expressions, including speaker-oriented adverbs as in (5a) and clause embedding fragments argued to behave like sentential adverbs (5b) (Jackendoff 1972).¹

- (5) a. Mary ate {possibly/regrettably/unfortunately} the most expensive pizza in Amherst.
- b. Mary ate {I believe/I think} the most expensive pizza in Amherst.

Since all of the expressions that I argue can modify DPs can also function as sentence-modifying expressions, we must rule out an alternative analysis of sentences like (5) in which the adverb is a sentential modifier that is linearized in sentence medial position (6):

- (6) Possibly Mary ate the most expensive pizza in Amherst. → Mary ate, possibly, the most expensive pizza in Amherst.

Linearization of *possibly* as in (6) would occur on analogy to other adverbs that function as parenthetical expressions. When linearized in sentence medial position, parenthetical adverbs are pronounced with obligatory comma intonation (prosodic breaks) around the adverb. Furthermore, the semantic scope of parenthetical adverbs does not depend on their linear position in the sentence (Ernst 2002). Regardless of its linear position, the adverb *happily* in (7) takes scope over the entire sentence (Mary's missing her sister's birthday was a happy occurrence).

- (7) Happily Mary missed her sister's birthday. → Mary missed, happily, her sister's birthday.

Ernst (1984) observes that the modal adverbs in sentences like (8) lack comma intonation:

- (8) a. Mary ate possibly the most expensive pizza in Amherst.
- b. Mary walked towards possibly a Canadian doctor.

¹Non-epistemic modals cannot modify DPs, e.g. (i). I leave explanation of this restriction to future work.

(i) *Mary accidentally ate {obligatorily / deliberately} the worst pizza at the cooking competition.
Cannot mean: Mary accidentally ate something which was deliberately / obligatorily the worst pizza at the cooking competition.

The next subsection shows that unlike parenthetical adverbs, the linear position of prosodically integrated modal adverbs determines their scope. These two pieces of evidence point away from a parenthetical analysis of the adverbs in (8).

Scope of modal adverbs: Further evidence against a parenthetical analysis of DP-modifying modal adverbs comes from their fixed scope. While parenthetical adverbs are interpreted as having sentential scope regardless of their linear position, non-parenthetical adverbs have fixed scope (Jackendoff 1972). For example, *happily* in (9a) can only modify the VP (Mary missed her sister's birthday in a happy manner) and cannot be interpreted as having scope over the entire sentence. Given their fixed scope, non-parenthetical adverbs are ungrammatical when they cannot be interpreted in situ, as Ernst (2002) argues is the case for sentences like (9b,c).

- (9) a. Mary happily missed her sister's birthday.
- b. *Mary missed happily her sister's birthday.
- c. *Mary looked forward to happily her sister's birthday.

The grammaticality of examples like (8) demonstrates that non-parenthetical modal adverbs can be interpreted in object positions. When interpreted in such positions, modal adverbs only take scope over the DP. Examples in this section demonstrate that, as a result of their low scope, sentences with DP-modifying adverbs carry an actuality entailment. I use ‘actuality entailment’ to mean that the (a) sentences are only licit if there was an event of Mary eating something (10). By contrast, the (b) sentences, where adverbs adjoin higher in the structure (e.g. at VP), are felicitous in both types of scenarios.²

- (10) Context: Mary visited Amherst yesterday and she ate at Athena's pizza. You can't remember whether Athena's or Antonio's has the most expensive pizza. You say,
 - a. Mary ate [DP possibly [DP the most expensive pizza in Amherst]].
 - b. Mary [VP possibly [VP ate the most expensive pizza in Amherst]].
- (11) Context: Mary visited Amherst yesterday and planned to eat a pizza lunch at Athena's, which you know makes the most expensive pizza in Amherst. You know that Mary anticipated having to skip lunch, however, so perhaps she didn't eat anything.
 - a. #Mary ate [DP possibly [DP the most expensive pizza in Amherst]].
 - b. Mary [VP possibly [VP ate the most expensive pizza in Amherst]].

Note that the sentences above contained DPs containing superlative-marked adjectives. One may be tempted to argue that the adverb only composes with the adjective in these cases, particularly since Cinque (2010) argues that modal adverbs can in general compose with adjectives (e.g. ‘There was a probably precancerous mole on John's back’). I argue that such an analysis cannot be maintained. Modal adverbs are still licit in examples like (12) in which there is no adjective for the modal to modify. As before, the (a) sentences with modal-modified DPs again carry actuality entailments while (b) sentences do not:

- (12) Context: You know that Mary felt unwell today and went to the health center. She wasn't sure if she'd be seeing a doctor or a nurse. I ask you where Mary is. You say,
 - a. Mary is meeting with [DP possibly [DP a doctor]] at the health center.
 - b. Mary is [VP possibly [VP meeting with a doctor at the health center]].

²All sentences considered contain verbs with transparent object positions. I set aside the question of whether actuality entailments arise when modal-modified DPs are taken as object by verbs with opaque object positions.

- (13) Context: You know that Mary felt unwell today but could not decide whether to see a doctor at the health center or go home and rest. Later on we see that Mary's desk is empty and I ask you where Mary is. You say,
- a. #Mary is meeting with [DP possibly [DP a doctor]] at the health center.
 - b. Mary is [VP possibly [VP meeting with a doctor]] at the health center.

Examples like (14) further demonstrate that even when a DP does contain an adjective, the modal can still take scope over the entire DP. The speaker's uncertainty in (14) is not about whether what Mary ate was the best, but rather what type of food it was.³

- (14) Context: Mary went on part of a tour of New York restaurants known for serving the best versions of a particular dish (hamburger, pizza, spaghetti). You know she ate at just one restaurant, but you don't know which. I ask you what Mary did. You say,
 Mary ate possibly the best pizza in New York.

This section has shown that because the verb is always outside of the scope of a DP-modifying modal adverb, sentences with such adverbs always carry actuality entailments.

3 Analysis

Composition of the modal and DP: I argued above that unlike other sentential adverbs (e.g. *happily*), modal adverbs can be interpreted when they only scope over a DP.

- (15) a. *Mary missed happily her sister's birthday.
 b. Mary ate possibly the most expensive pizza in Amherst.

Assuming that modal adverbs are of type $\langle st, st \rangle$ (16a) regardless of their syntactic attachment site, the puzzle is how they can compose with a type $\langle se \rangle$ DP intension (16b) (*MEPA = the most expensive pizza in Amherst*).⁴

- (16) a. $\llbracket \text{possibly} \rrbracket = \lambda p_{st} \lambda w [\exists w'' \in \text{EPI-MB}(w)[p(w'')]]$
 b. $\llbracket \text{the most expensive pizza in Amherst} \rrbracket = \lambda w' \iota y [\text{MEPA}(y, w')]$

I argue that composition of the modal adverb and a DP is made possible by first type shifting the denotation of the definite description into a property via IDENTIFY:

- (17) a. $\llbracket \text{IDENTIFY} \rrbracket = \lambda \mathcal{X}_{se} \lambda z_e \lambda w' [z = \mathcal{X}(w')] \quad \langle se, est \rangle$
 b. $\llbracket \text{IDENTIFY the most expensive pizza in Amherst} \rrbracket = \lambda z_e \lambda w' [z = \iota y [\text{MEPA}(y, w')]]$

Precedent for this approach to modal-modified DPs comes from two sources. First, precedent for the type shifter comes from Frana (2006, 2010) — who introduces the name IDENTIFY — and Schwager (2008), who use a similar type shifter in their accounts of Concealed Questions (18), which can also be analyzed as type $\langle se \rangle$ DP intensions composing with type $\langle st, st \rangle$ expressions (factive verbs):

- (18) Mary knows [IDENTIFY [DP the capital of Italy]].

³Prosodic prominence — perhaps Focus marking — on *pizza* in (14) helps to bring out the intended reading.

⁴More accurately, modal adverbs compose with the intension of a DP. Regardless of where they occur in the sentence, modals always require access to intensions.

Second, this analysis of modal-modified DPs makes them similar in semantics to Transparent Free Relatives under Grosu's (2003, 2007) treatment:

- (19) a. John was attacked by [TFR what seemed to be *a student*].
 b. John is a [TFR what Mary would call *unique*] individual. (adapt. Grosu 2003)

Grosu (2003, 2007) analyzed TFRs as structurally similar to standard free relatives (e.g. 'Mary ate what John cooked').⁵ TFRs are always headed by bare *what* and characteristically contain an intensional operator. The core semantic content of the TFR is expressed by the pivot (italicized in (19)) which can be a DP (19a) or property (19b). When the pivot is a DP as in (19a), it is taken as argument by a copula with an equative-specificational semantics.⁶ The entry of IDENTIFY in (17) only differs from (20)— one of Romero's (2005) entries for the specificational copula — in the order of composition with its intensional and extensional arguments. Higgins (1973) notes that order of composition is a general point of flexibility in specificational sentences.

- (20) $\llbracket be_{SPEC} \rrbracket = \lambda z_e \lambda \mathcal{X}_{se} \lambda w' [z = \mathcal{X}(w')]$ (Romero 2005)

TFRs have a number of similarities with modal-modified DPs. In particular, the intensional operator in a TFR only takes scope within the clause containing the pivot; it does not take scope over the matrix verb. As a result, sentences with TFRs — like sentences with modal-modified DPs — carry actuality entailments. I illustrate this in (21) with a TFR based on a modal-modified DP familiar from discussion above:

- (21) Context from (11): #Mary ate what is possibly the most expensive pizza in Amherst.

Once IDENTIFY is admitted into our account of modal-modified DPs, composition between the modal and type shifted DP could be handled in two ways. First, we could maintain as much of the TFR structure as possible and argue that modal-modified DPs are reduced relative counterparts to TFRs. Following Bhatt's (2006) Direct Predication analysis of reduced relatives, a modal-modified DP could have the syntax in (22). The structure is identical to a TFR except that we replace *what* with semantically vacuous PRO (which is then abstracted over) and the specificational copula with IDENTIFY.

- (22) $\lambda y [\text{possibly} [\text{SC } \text{PRO}_y \text{ IDENTIFY } \text{the most expensive pizza in Amherst }]]$

Alternatively, we could argue that the type shifted DP can compose with the modal adverb via Function Composition (23).

$$\begin{aligned}
 (23) \quad & \text{Shifted DP: } f = \lambda z \lambda w' [z = \iota x [\text{MEPA}(x, w')]] && \langle e, st \rangle \\
 & \text{Modal: } g = \lambda p_{st} \lambda w [\exists w'' \in \text{EPI-MB}(w) [p(w'')]] && \langle st, st \rangle \\
 & \text{Output: } (g \circ f) && \langle e, st \rangle \\
 & = \lambda y (\lambda p_{st} \lambda w [\exists w'' \in \text{EPI-MB}(w) [p(w'')]]) (\lambda w' [y = \iota x [\text{MEPA}(x, w')]]) \\
 & = \lambda y \lambda w [\exists w'' \in \text{EPI-MB}(w) [y = \iota x [\text{MEPA}(x, w'')]]]
 \end{aligned}$$

Both analyses maintain a type $\langle st, st \rangle$ entry for modal adverbs and do not require us to treat sentential adverbs capable of modifying DPs as lexically ambiguous in any way. I leave to further work comparison of the two analyses. In discussion below, I will assume the Function Composition analysis since it is representationally simpler.

⁵For an alternative treatment of TFRs, see van Riemsdijk (2000).

⁶Equivalently, the DP could first be shifted into a property via Partee's (1986) IDENT and then compose with a predicative copula (Rothstein 2001).

Composition of modal-modified DPs and verbs: Both the reduced relative and Function Composition analyses produce modified DPs of type $\langle e, st \rangle$. There are several options for composing this expression with the rest of the clause. First, we could have the expression move above the verb and subject, leaving behind a trace of type e (24a). The predicate produced by lambda abstraction could compose with the modified DP through Predicate Modification followed by Existential Closure.⁷

- (24) a. [[possibly IDENTIFY the most expensive pizza in Amherst] λx Mary ate x]
- b. $\llbracket(24a)\rrbracket = \lambda z \lambda w' [\lambda y \lambda w [\exists w'' \in EPI-MB(w)[y = \iota x[\text{MEPA}(x, w'')]]](z)(w') \ \& \ PM$
 $[\lambda x \lambda s [\text{Mary ate } x \text{ in } s]](z)(w')$
- c. $\exists z \lambda w' [\exists w'' \in EPI-MB(w')[z = \iota x[\text{MEPA}(x, w'')]]] \ \& \ \text{Mary ate } z \text{ in } w'$ EC
 ‘There exists a z such that Mary ate z in w' and in some world w'' epistemically accessible from w' , z is the most expensive pizza in Amherst.’

Second, we could treat modal-modified DPs as specific indefinites, having them compose with a type $\langle est, e \rangle$ choice function f that returns some member of the set denoted by the modified DP. The returned entity composes in situ with the verb and the choice function is existentially quantified over (as in (25); Reinhart 1997) or given by the context (Kratzer 1998).

- (25) $\exists f [\text{Mary ate } f(\lambda z \lambda w[\exists w'' \in EPI-MB(w)[z = \iota x[\text{MEPA}(x, w'')]])]$

Third, the modified DP could compose with a definiteness (or maximalization) operator on analogy with free relatives (Jacobson 1995).

- (26) $\llbracket \text{DEF} \rrbracket (\lambda y \lambda w [\exists w'' \in EPI-MB(w)[y = \iota x[\text{MEPA}(x, w'')]])$
 $= \iota z \lambda w' [\exists w'' \in EPI-MB(w')[y = \iota x[\text{MEPA}(x, w'')]]]$

Of these three approaches, the first two make the right predictions while the third is untenable. The following example shows that modal-modified DPs can appear in contexts where there is not a single unique referent for the modal-modified description.⁸ This is expected given the existential quantification in the first two approaches but unexpected under the third.

- (27) Context: You went to a wine tasting yesterday. You arrived late and missed the descriptions of each wine and only got to try one of the samples on the table. You know there was to be one American, one Chilean, and one Argentinian wine. The wine you drink is white; you know that the US produces a lot of white wine. You say,
 I drank possibly the American wine.

The only requirement imposed by a definite determiner within the modified DP in (27) is that in each accessible alternative w' , there be a unique referent for the description within that alternative w' . If context (27) were minimally changed such that the speaker knew there to be two American wines, the target sentence could no longer be uttered.⁹

⁷This operation could also be modeled as a case of composition through Restrict, a mode of composition that Chung and Ladusaw (2001) propose for the composition of predicates with properties. Composition through Restrict does not change the predicate’s degree of saturation so Existential Closure must still follow.

⁸This is another point of similarity between modal-modified DPs and TFRs. Sentence (27) could have contained TFRs rather than a modal-modified DP. Grosu (2003) pursues a strategy for composition of TFRs with verbs similar to the first strategy outlined above, arguing that they are lifted from properties into existentially quantified expressions which then raise to compose.

⁹It is, of course, not generally impossible for modal adverbs to occur within the scope of definite determiners. When they do, however, the uniqueness requirement of the definite determiner returns and the modal appears to only take scope over the adjective, e.g., *Mary drank the possibly American wine* is only felicitous if the context contains a unique wine describable as ‘possibly American.’

4 Summary and Discussion

I have demonstrated that when epistemic modal adverbs form syntactic constituents with DPs, the modal takes scope only over the DP: the world variable of the verb is not bound by the modal. As a result, sentences with modal-modified DPs — but not sentences with higher adjoined adverbs — carry actuality entailments. The puzzle was: how can we maintain a type ⟨st,st⟩ entry for modal adverbs while allowing them to compose with type ⟨se⟩ DP intensions? Drawing on proposals for similar puzzles arising for Transparent Free Relatives and Concealed Questions, I argued that DPs are shifted into property type prior to composition with adverbs.

One direction for further work is comparison of modal-modified DPs with Collins Conjunctions (Collins 1988, Schein 1997). Strings like (28) are ambiguous between the DP-modifying reading addressed in this paper and the Collins Conjunction reading. In the Collins Conjunction reading, it appears that the adverb takes scope over an unpronounced instance of *brought*. Unlike sentences with DP-modifying modals, Collins Conjunction sentences lack an actuality entailment for the modal-modified individual. The structural difference underlying the ambiguity seems related to the adverb's height: DP-modifying modals are lower than Collins Conjunction modals. This is supported by sentences which contain adverbs in both positions (29).

- (28) Mary brought a cake and [possibly the most expensive pizza in Amherst] to the party.
- a. **DP-modifying:** Mary brought two things: a cake and something which is possibly the most expensive pizza in Amherst.
 - b. **Collins Conjunction:** Mary brought a cake and possibly also brought the most expensive pizza in Amherst.
- (29) Mary brought a cake and possibly [perhaps the most expensive pizza in Amherst].

There is, however, a similarity between adverbs in both positions: neither position permits non-epistemic modals (fn. 2 above, Vicente 2013). What is the source of this restriction?

A second direction to explore is the binding possibilities of DP-modifying modal adverbs. While world (or, situation) variables of DPs can be bound non-locally, world variables of adverbial quantifiers on the clausal spine must be bound by the closest operator (Percus 2000). The world variable of *always* in (30) can only be bound by *think* (*w2*), and not by *w1*. Were the latter possible, (30) is predicted to be true if my brother (or who Mary believes to be my brother) won every actual game round.

- (30) $\lambda_1 w1 [\text{Mary thinks } [\lambda_2 \text{my brother}(w2/w1) \text{ always}(w2/*w1) \text{ won the game}]]$.

What happens when an adverbial quantifier modifies a DP? While the modal certainly can be locally bound, it appears that it can alternatively be bound by the same expression that non-locally binds the DP's situation variable, e.g. *w1*:

- (31) Context: Mary had pizza served from a box with Athena's logo on it. I know that Athena's is among the most expensive in town. Mary is visiting and does not know this. Unknown to Mary, her pizza was actually from Don's, which is the cheapest in town. I report Mary's belief, saying:
- a. Mary thinks she ate possibly the most expensive pizza in Amherst.
 - b. $\lambda_1 w1 [\text{Mary thinks } [\lambda_2 \text{she ate } [\text{possibly the most expensive pizza in Amherst}(w1)]]]$

As discussion of modal modification of DPs increases, it will be interesting to continue to compare the behavior of DP-modifying adverbs with modals along the clausal spine.

Finally, as noted in the introduction, DPs are not the only subpropositional constituents to

be modified by epistemic modal adverbs. Although the modification of type ⟨e,st⟩ adjectives (1b) is expected under my analysis, it is not clear how it can extend to numerals (1a). Will it be necessary to treat numeral-modifying modals as approximators with an added modal component (Zaroukian 2011)? Or is a single, familiar semantics for modal adverbs tenable?

References

- [1] R. Bhatt. *Covert Modality in Non-finite Contexts*. Mouton de Gruyter, Berlin, 2006.
- [2] S. Chung and W. Ladusaw. *Restriction and Saturation*. MIT Press, 2001.
- [3] G. Cinque. *The Syntax of Adjectives*. MIT Press, Cambridge, MA, 2010.
- [4] C. Collins. Conjunction adverbs. Ms, MIT, 1988.
- [5] T. Ernst. *Towards an Integrated Theory of Adverb Position in English*. PhD thesis, Indiana University, 1984.
- [6] T. Ernst. *The Syntax of Adjuncts*. Cambridge University Press, Cambridge, MA, 2002.
- [7] I. Frana. The *de re* analysis of concealed questions. In C. T. et al., editor, *Proceedings of SALT XVI*, 2006.
- [8] I. Frana. Copular questions and concealed questions. In V. S. M. Prinzhorn and S. Zobel, editors, *Proceedings of Sinn und Bedeutung 14*, 2010.
- [9] A. Grosu. A unified theory of ‘standard’ and ‘transparent’ free relatives. *Natural Language and Linguistic Theory*, 2003.
- [10] A. Grosu. ‘Direct’ vs. ‘indirect’ approaches to transparent free relatives. In *Pitar Moş: A Building with a View*. University of Bucharest Press, 2007.
- [11] R. Higgins. *The Pseudo-Cleft Construction in English*. PhD thesis, MIT, 1973.
- [12] R. S. Jackendoff. *Semantic Interpretation in Generative Grammar*. MIT Press, Cambridge, MA, 1972.
- [13] P. Jacobson. On the quantificational force of English free relatives. In A. K. E. Bach, E. Jelinek and B. Partee, editors, *Quantification in Natural Languages*. Kluwer, Dordrecht, 1995.
- [14] A. Kratzer. The notional category of modality. In H. Ekmeyer and H. Rieser, editors, *Words, Worlds, and Contexts*, pages 38–74. de Gruyter, 1981.
- [15] A. Kratzer. Scope or pseudo-scope: Are there wide-scope indefinites? In *Events in Grammar*. Dordrecht, 1998.
- [16] B. Partee. Noun phrase interpretation and type-shifting principles. In P. Portner and B. Partee, editors, *Formal Semantics: The Essential Readings*. Blackwell, 1986.
- [17] O. Percus. Constraints on some other variables in syntax. *Natural Language Semantics*, 8:173–229, 2000.
- [18] T. Reinhart. Quantifier scope: How labor is divided between QR and choice functions. *Linguistics & Philosophy*, 20(335–397), 1997.
- [19] M. Romero. Concealed questions and specifical subjects. *Linguistics & Philosophy*, 28:687–737, 2005.
- [20] S. Rothstein. *Predicates and their Subjects*. Kluwer, Dordrecht, 2001.
- [21] B. Schein. Conjunction reduction redux. Ms, University of Southern California, 1997.
- [22] M. Schwager. Keeping prices low: an answer to a concealed question. In *Proceedings of Sinn und Bedeutung 12*, 2008.
- [23] H. van Riemsdijk. Free relatives inside out: Transparent free relatives as grafts. In B. Rozdowksa, editor, *PASE Papers in Language Studies*. University of Wrocław, 2000.
- [24] E. Zaroukian. Divergent approximators. In I. Reich and E. Horch, editors, *Proceedings of Sinn und Bedeutung 15*, 2011.

Witness Sets, Polarity Reversal and the Processing of Quantified Sentences

Oliver Bott¹, Udo Klein² and Fabian Schlotterbeck¹

¹ Tübingen University, Tübingen, Germany

{oliver.bott,fabian.schlotterbeck}@uni-tuebingen.de

² Bielefeld University, Bielefeld, Germany

udo.klein@uni-bielefeld.de

Experimental results show that the monotonicity of a quantifier (Q) affects how it is processed [4, 7, 6]. Q 's are upward entailing (UE) if they permit inferences to supersets, i.e. from $Q A B$ follows $Q A B'$ for any $B \subseteq B'$ (e.g. *every*, *at least five*). If from $Q A B$ it follows that $Q A B'$ for any $B' \subseteq B$, the quantifier Q is downward entailing (DE) (e.g. *no*, *at most five*). As compared to UE Q 's, DE Q 's are more difficult to verify [4, 7] and draw inferences from [6]. Recent attempts at deriving predictions about the processing of Q 's are built on the notion of semantic automata [12], and have been presented e.g. in [11] and [10]. E.g., [11] improve on results by [8] by showing not only that the computational distinction between Q 's recognized by finite-automata and push-down automata is psychologically relevant, but also that there are differences in the time required for verifying statements involving Q 's even among the class of quantifiers recognized by finite state automata. However, since these approaches employ essentially the same kind of semantic automata for both DE and UE Q 's, they cannot explain why DE Q 's are more difficult to process than UE Q 's. To explain this, we formulate a quantification theory which predicts that the (expansion) operation employed in processing DE Q 's is more difficult than that for processing UE Q 's, because it involves (i) inferences from negative information and (ii) polarity reversal. The predictions of our account were tested in two experiments investigating the online comprehension and verification of simply (Exp. 1) and doubly quantified sentences (Exp. 2) with UE vs DE Q 's.

1 An algorithmic theory of quantification

In the following, we aim to provide an algorithmic theory of meaning in the sense of [9]: The meaning of an expression is the algorithm which computes its denotation. Here is the basic idea of what will be worked out below. In the case of a quantificational sentence $Q_1 \text{ } boys \text{ } tickled \text{ } Q_2 \text{ } girls$, with $Q_1 \text{ } boys$ scoping over $Q_2 \text{ } girls$, the interpretation consists in specifying an algorithm which outputs whether S is true given any model, and the verification consists in applying this algorithm to a specific model.

Given a binary relation R (e.g. *tickles*), we first construct a polarity relation R^* , such that $\langle a, b, + \rangle \in R^*$ if $\langle a, b \rangle \in R$; and $\langle a, b, - \rangle \in R^*$ otherwise. We now specify an algorithm which derives $\langle Q_1 \text{ } boys, Q_2 \text{ } girls, + \rangle$ if and only if our input sentence is true (relative to a certain scoping). The algorithm for the verification of the doubly quantified sentence above (relative to the subject wide scope reading) consists of the following steps:

- (1)
 - a. For every boy x who *tickles* some girl, determine which *girls* he *tickles*. If the set of girls tickled by x is a witness set of $Q_2 \text{ } girls$, add the positive information $\langle x, Q_2 \text{ } girls, + \rangle$. Otherwise, add the negative information that $\langle x, Q_2 \text{ } girls, - \rangle$.
 - b. For every boy x who tickles no girl, add $\langle x, Q_2 \text{ } girls, + \rangle$ if $Q_2 \text{ } girls$ has the empty set as a witness set. Otherwise, add the negative information $\langle x, Q_2 \text{ } girls, - \rangle$.

- (2) a. If for some boy x we have $\langle x, Q_2 \text{ girls}, + \rangle$, and the set of boys y with $\langle y, Q_2 \text{ girls}, + \rangle$ is a witness set of $Q_1 \text{ boys}$, then add $\langle Q_1 \text{ boys}, Q_2 \text{ girls}, + \rangle$, else $\langle Q_1 \text{ boys}, Q_2 \text{ girls}, - \rangle$.
b. If for every boy x we have $\langle x, Q_2 \text{ girls}, - \rangle$, and the empty set is a witness set of $Q_1 \text{ boys}$ then add $\langle Q_1 \text{ boys}, Q_2 \text{ girls}, + \rangle$. Otherwise, add $\langle Q_1 \text{ boys}, Q_2 \text{ girls}, - \rangle$

$Q_1 \text{ boys tickled } Q_2 \text{ girls}$ is true iff $\langle Q_1 \text{ boys}, Q_2 \text{ girls}, + \rangle$ can be added. Importantly, if neither $Q_1 \text{ boys}$ nor $Q_2 \text{ girls}$ have the empty set as a witness set, then the general algorithm can be simplified as follows:

- (3) For every boy x who *tickles* some girl, determine which *girls* he *tickles*. If the set of girls tickled by x is a witness set of $Q_2 \text{ girls}$, add the positive information $\langle x, Q_2 \text{ girls}, + \rangle$.
(4) If for some boy x we have $\langle x, Q_2 \text{ girls}, + \rangle$, and the set of boys y with $\langle y, Q_2 \text{ girls}, + \rangle$ is a witness set of $Q_1 \text{ boys}$ then add $\langle Q_1 \text{ boys}, Q_2 \text{ girls}, + \rangle$.

In other words, if neither quantifier is an *empty-set quantifier*, the negative information can safely be ignored in every model, and the algorithm can be restricted to positive information. We will refer to this simplified procedure as *simple expansion (s-exp)*. If one (or both) quantifiers are empty-set quantifiers, and the antecedent of (2b) holds, then negative information becomes relevant, since step (2b) allows the addition of positive information based on negative information, and therefore the more complex algorithm (**c-exp**) is required. The differentiation between **s-exp** and **c-exp** sets our theory apart from existing algorithmic proposals of quantification, as for instance [10]’s automata theory building on [12].

2 Quantification theory

First, we assume that simple NL determiners denote unary functions from restrictor sets to pairs consisting of the restrictor set itself and a set of subsets of the restrictor set, called the set of witness sets, i.e. as functions $q : \mathcal{P}(E) \mapsto \mathcal{P}(E) \times \mathcal{P}(\mathcal{P}(E))$.

Definition 1 (w-function, w-quantifier, witness sets): Let q be a function from the set $\mathcal{P}(E)$ of subsets of the domain of entities E into $\mathcal{P}(E) \times \mathcal{P}(\mathcal{P}(E))$. Then q is called a *w-function*¹ if for any $A \subseteq E$ there is a $W \subseteq \mathcal{P}(A)$, such that $q(A) = \langle A, W \rangle$. If q is a *w-function* and $A \subseteq E$, then $q(A) = \langle A, W \rangle$ is called a *w-quantifier*, and W the set of witness sets of q at A . If a *w-quantifier* $q(A) = \langle A, W \rangle$ is such that $\emptyset \in W$, it is called an *empty-set quantifier*.

To illustrate, for any subset A of the domain of individuals E , we have:

$$(5) \quad \begin{aligned} \llbracket \text{some} \rrbracket(A) &= \langle A, \{X : X \subseteq A \wedge |X| \geq 1\} \rangle \\ \llbracket \text{most} \rrbracket(A) &= \langle A, \{X : X \subseteq A \wedge |A \cap X| > |A - X|\} \rangle \end{aligned}$$

Note that these w-quantifiers naturally correspond to the standard generalized quantifier denotations:

$$(6) \quad \begin{aligned} \llbracket \text{some} \rrbracket_E(A) &= \{B : B \subseteq E \wedge |A \cap B| \geq 1\} \\ \llbracket \text{most} \rrbracket_E(A) &= \{B : B \subseteq E \wedge |A \cap (A \cap B)| > |A - (A \cap B)|\} \end{aligned}$$

Importantly, note that for most type $\langle 1, 1 \rangle$ quantifiers there is no corresponding w-quantifier:

$$(7) \quad \begin{aligned} \text{EQUI}_E(A) &= \{B : B \subseteq E \wedge |A| = |B|\} \\ \text{TOTAL}_E(A) &= \{B : B \subseteq E \wedge E = A\} \end{aligned}$$

¹Mnemonic for witness-set function.

In fact, a type $\langle 1, 1 \rangle$ quantifier Q has a corresponding w-quantifier if and only if Q satisfies both conservativity (CONS) and extension (EXT).²

In order to explicitly encode positive and negative information we define the polarity relation P^* of an n -ary predicate: if $\langle a, b \rangle \in P$ then $\langle a, b, + \rangle \in P^*$, and if $\langle a, b \rangle \notin P$ then $\langle a, b, - \rangle \in P^*$.

Definition 2 (n-ary predicate, polarity relation): A subset P of $(E \cup \mathcal{P}(E) \times \mathcal{P}(\mathcal{P}(E)))^n$, with $n \geq 1$, is called an n -ary predicate. For any n -ary predicate P and $\pi_i(\sigma)$ the i -th element of an n -ary tuple σ (for any $1 \leq i \leq n$), let $P^* = \{\langle \pi_1(\sigma), \dots, \pi_n(\sigma), + \rangle : \sigma \in P\} \cup \{\langle \pi_1(\sigma), \dots, \pi_n(\sigma), - \rangle : \sigma \notin P\}$ be the polarity relation of P .

To illustrate, if $E = \{a, b, c\}$, $A = \{a, b\}$, and $P = \{\langle a, a \rangle, \langle a, b \rangle, \langle c, c \rangle\}$ is a binary predicate, then $P^* = \{\langle a, a, + \rangle, \langle a, b, + \rangle, \langle c, c, + \rangle\} \cup \{\langle a, c, - \rangle, \langle b, a, - \rangle, \langle b, b, - \rangle, \langle b, c, - \rangle, \langle c, a, - \rangle, \langle c, b, - \rangle\}$ is polarity relation of P .

Next, we define the set of i -fillers $[\sigma]_i^{P^*}$ of a tuple $\sigma \in P^*$ as the set of elements in the i -th position of an i -variant $\tau \in P^*$ of σ (τ is an i -variant of σ iff τ and σ differ at most at the i -th position).

Definition 3 (i-filters): Let P^* be the polarity relation of some n -ary predicate P and let $\sigma \sim_i \tau$ hold iff σ and τ differ at most at the i -th element. For every $\sigma \in P^*$ and any integer $1 \leq i \leq n$, let $[\sigma]_i^{P^*} = \{\pi_i(\tau) : \tau \in P^* \wedge \tau \sim_i \sigma\}$ be the i -fillers of σ in P^* . Let $[\sigma]_i^{P^*, A} = [\sigma]_i^{P^*} \cap A$ be the set of those i -fillers of σ in P^* which are also in the set A (called restrictor set).

For example, if $P^* = \{\langle a, a, + \rangle, \langle a, b, + \rangle, \langle b, a, - \rangle, \langle b, b, + \rangle\}$, then $[\langle a, a, + \rangle]_1^{P^*} = \{a\}, [\langle a, a, + \rangle]_2^{P^*} = \{a, b\}$ and $[\langle b, a, - \rangle]_1^{P^*} = \{b\}, [\langle b, a, - \rangle]_2^{P^*} = \{a\}$. Finally, we formulate two expansion operations which, when applied to a quantifier $q(A)$ and a polarity relation P^* , add those tuples representing negative or positive information involving $q(A)$ at position i in P^* .

Definition 4 (simple and complex i-expansion): Let $q : \mathcal{P}(E) \mapsto \mathcal{P}(E) \times \mathcal{P}(\mathcal{P}(E))$ be a function such that $q(A) = \langle A, W \rangle$, where $W \subseteq \mathcal{P}(A)$, and P be an n -ary predicate over E ($n \geq 1$). For any $\sigma \in P^*$ and any $1 \leq i \leq n$, let $\sigma[i/x]$ be the result of replacing the i -th element of σ by x . Then the simple expansion of P^* by $q(A)$ at position i , written as $\mathbf{s-exp}_i(q(A), P^*)$ is the smallest set Q such that $P^* \subseteq Q$ and clause 1 holds. The complex expansion of P^* by $q(A)$ at position i , written as $\mathbf{c-exp}_i(q(A), P^*)$ is the smallest set Q such that $P^* \subseteq Q$ and clauses 1-4 below hold:

1. $\sigma^+ \in P^* \wedge [\sigma^+]_i^{P^*, A} \in W \rightarrow \sigma^+[i/q(A)] \in Q$ (positive \triangleright positive)
2. $\sigma^+ \in P^* \wedge [\sigma^+]_i^{P^*, A} \notin W \rightarrow \sigma^-[i/q(A)] \in Q$ (positive \triangleright negative)
3. $\sigma^- \in P^* \wedge [\sigma^-]_i^{P^*, A} = A \wedge \emptyset \in W \rightarrow \sigma^+[i/q(A)] \in Q$ (negative \triangleright positive)
4. $\sigma^- \in P^* \wedge [\sigma^-]_i^{P^*, A} = A \wedge \emptyset \notin W \rightarrow \sigma^-[i/q(A)] \in Q$ (negative \triangleright negative)

To illustrate, consider the evaluation of *Every boy (Q₁) tickled exactly three girls (Q₂)* by means of **s-exp** in a model with $B = \{b_1, b_2\}$, $G = \{g_1, g_2, g_3, g_4\}$ and $T = \{\langle b_1, g_1 \rangle, \langle b_1, g_2 \rangle, \langle b_1, g_3 \rangle, \langle b_2, g_1 \rangle, \langle b_2, g_2 \rangle, \langle b_2, g_3 \rangle, \langle g_1, g_2 \rangle, \langle g_1, g_3 \rangle, \langle g_1, g_4 \rangle\}$. By the first clause of **s-exp**₂ we can add $\langle b_1, \llbracket Q_2 \rrbracket, + \rangle, \langle b_2, \llbracket Q_2 \rrbracket, + \rangle$ and $\langle g_1, \llbracket Q_2 \rrbracket, + \rangle$, and by clause 1 of **s-exp**₁ we add $\langle \llbracket Q_1 \rrbracket, \llbracket Q_2 \rrbracket, + \rangle$, since $\{b_1, b_2, g_1\} \cap B$ is a witness set of $\llbracket \text{every boy} \rrbracket$. The subject-wide-scope reading of $Q_1 V Q_2$ is true in \mathcal{M} iff $\langle \llbracket Q_1 \rrbracket, \llbracket Q_2 \rrbracket, + \rangle \in \mathbf{c-exp}_1(\llbracket Q_1 \rrbracket, \mathbf{c-exp}_2(\llbracket Q_2 \rrbracket, \llbracket V \rrbracket^*))$. Clearly, if $\langle \llbracket Q_1 \rrbracket, \llbracket Q_2 \rrbracket, + \rangle \in \mathbf{s-exp}_1(\llbracket Q_1 \rrbracket, \mathbf{s-exp}_2(\llbracket Q_2 \rrbracket, \llbracket V \rrbracket^*))$, then $\langle \llbracket Q_1 \rrbracket, \llbracket Q_2 \rrbracket, + \rangle \in \mathbf{c-exp}_1(\llbracket Q_1 \rrbracket, \mathbf{c-exp}_2(\llbracket Q_2 \rrbracket, \llbracket V \rrbracket^*))$. Importantly, the truth evaluation of this sentence in a model requires neither inferences based

²A quantifier Q satisfies CONS iff for all domains E and for all $A, B \subseteq E$, it holds that $Q_E(A, B)$ iff $Q_E(A, A \cap B)$, and EXT iff iff for all domains E and all $A, B \subseteq E$, it holds that $Q_E(A, B)$ iff $Q_A(A, B)$.

on negative information nor polarity reversal. On the other hand, whether *Less than two boys* (Q_1) *tickled exactly three girls* (Q_2) can be shown to be true by **s-exp** depends on the situation. If $B = \{b_1, b_2\}$, $G = \{g_1, g_2, g_3, g_4\}$ and $T = \{\langle b_1, g_1 \rangle, \langle b_1, g_2 \rangle, \langle b_1, g_3 \rangle\}$, then $\langle b_1, \llbracket Q_2 \rrbracket, + \rangle$ is added by **s-exp**₂, and by **s-exp**₁ we add $\langle \llbracket Q_1 \rrbracket, \llbracket Q_2 \rrbracket, + \rangle$, since $\{b_1\}$ is a witness set of *less than two boys*. But in the situation where $B = \{b_1, b_2\}$, $G = \{g_1, g_2, g_3, g_4\}$ and $T = \{\langle g_1, g_2 \rangle\}$ **c-exp** is required: By clause 4 of **c-exp**₂ we add the negative information that $\langle b_1, \llbracket Q_2 \rrbracket, - \rangle$ and $\langle b_2, \llbracket Q_2 \rrbracket, - \rangle$, and by clause 3 of **c-exp**₁ we add the positive information $\langle \llbracket Q_1 \rrbracket, \llbracket Q_2 \rrbracket, + \rangle$ based on the negative information added before (polarity reversal). An important consequence of this theory is the following proposition:

Proposition 1: *Let $q_1(A_1) = \langle A_1, W_1 \rangle$ and $q_2(A_2) = \langle A_2, W_2 \rangle$ be w-quantifiers with $\emptyset \notin W_1$ and $\emptyset \notin W_2$, and let P^* be the polarity relation of some predicate $P \subseteq E^n$ ³. Then $\langle q_1(A_1), q_2(A_2), + \rangle \in \mathbf{s-exp}_1(q_1(A_1), \mathbf{s-exp}_2(q_2(A_2), P^*))$ if and only if $\langle q_1(A_1), q_2(A_2), + \rangle \in \mathbf{c-exp}_1(q_1(A_1), \mathbf{c-exp}_2(q_2(A_2), P^*))$.*

That is, if a quantified statement contains no empty-set quantifiers, then $\langle q_1(A_1), q_2(A_2), + \rangle$ can be added by **s-exp** if and only if it can be added by **c-exp** – in other words, the evaluation of such a statement’s truth requires only inferences of positive information from positive information. If, on the other hand, a quantified statement does contain an empty-set quantifier, then it cannot be evaluated for truth by **s-exp** in every model, as shown above.

Based on our theory we derive the following predictions about the processing of quantified sentences. **Prediction 1:** Given that negation and polarity reversal increase processing difficulty [2], we predict first that quantified statements which have to be evaluated by **c-exp** are more difficult to process than statements which can be evaluated by **s-exp**. Non-empty-set quantifiers (e.g. UE quantifiers like *at least/more than n* or *every*⁴) can be evaluated by **s-exp** in all models/situations. On the other hand, the evaluation of empty-set quantifiers (all DE quantifiers, among others) depends on the model: in some models **s-exp** suffices, in others **c-exp** is necessary. We predict increased processing difficulty whenever **s-exp** does not suffice. In particular, processing difficulty should be increased if a rule has to be applied that involves polarity reversal (clauses 2 and 3) or is based on negative information (clauses 3 and 4). The most difficult cases should be applications of clause 3 because this clause involves polarity reversal and inference from negative information. **Prediction 2:** If the comprehension of a quantified sentence involves the specification of the simplest algorithm for checking this sentence *in any model*, we predict not only differences in ease of verification but also differences in ease of comprehension: sentences involving no empty-set Q should be read faster than sentences involving empty set Q ’s, because the former requires the specification of an algorithm involving **s-exp**, whereas the latter involves the complex expansion **c-exp**.

3 Experiments

Experiment 1: 72 German participants read simply quantified intransitive sentences of the type Q dots are blue manipulating Q : (a) *mindestens ein...* (at least one, non-empty-set), (b) *höchstens ein...* (at most one, empty-set) and (c) *weniger als zwei* (less than two, empty-set). After reading the sentence, a picture was presented that showed either zero objects (0-model), one object (1-model), or three objects (3-model) of the relevant color among other objects of a different color.

³Note that this proposition also holds for empty predicates.

⁴Assuming existential import *every* can also be considered a non-empty-set quantifier since it presupposes a non-empty restriction [3]. Note that without this presupposition the empty set is a witness set of *every*(A) in case A is empty.

All objects (total number between 4 to 7 varying across items) in the display were elements of the restrictor set. 27 items were constructed in a 3×3 within design and distributed to nine lists using a latin square. The descriptive statistics are presented in Table 1.

	error rates			judgment times		
	0-model	1-model	3-model	0-model	1-model	3-model
<i>at least one</i>	0%	0%	4.9%	1511ms	1300ms	1453ms
<i>fewer than two</i>	10.3%	2.0%	0%	2076ms	1744ms	1755ms
<i>at most one</i>	22.5%	1.0%	1.0%	1939ms	1385ms	1478ms

Table 1: Mean judgments and judgment times in Exp. 1.

The analysis of error rates revealed that, except for the two empty-set quantifiers with 0-models, all of the conditions had error rates well below 5%. The 0-models led to 10.3% errors for *less than* and to even 22.6% errors for *at most*⁵. In a logit mixed effects model analysis, the observed differences led to a significant *quantifier* \times *model* interaction ($z = 2.42$). The increased error rates for empty-set Q 's when the predicate is empty are fully expected since this case requires the application of the most difficult clause of **c-exp**, namely clause 3, which involves polarity reversal and inference from negative information. If the predicate is not empty, then there is an $a \in E$ with $\langle a, + \rangle \in P^*$, so clause 1 or 2 can be applied. The verification times⁶ also showed a significant interaction between *quantifier* and *model* (ANOVAs: $p_1 < .01$; $p_2 < .05$) which lends further support to our theory (prediction 1). Thus, the evaluation of 0-models proves to be a source of quantificational complexity in empty-set quantifiers even in simply quantified intransitive sentences.

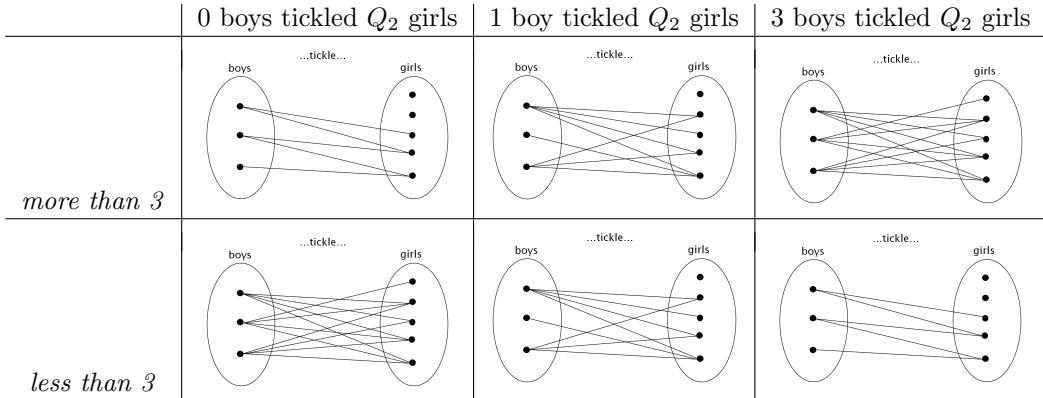


Figure 1: Sample models in Exp. 2 for Q_1 boys tickled Q_2 girls. Note: diagrams corresponding to 0- and 3-models had to be swapped for $Q_2=more\ than$ and $Q_2=fewer\ than$, but the set of diagrams was kept constant across sentence conditions.

Experiment 2: 72 participants read German sentences of the form Q_1 boys tickled Q_2 girls which have independently been shown to exhibit surface scope only [cf. 1]. Q_1 was either one of the Aristotelian quantifiers *jeder* (each, non-empty-set) and *kein* (no, empty-set) or one of the superlative quantifiers *mindestens ein* (at least one, non-empty-set) and *höchstens*

⁵Being well aware of the fact that these “errors” may be (at least partly) caused by pragmatic factors.

⁶All judgment times and reading times in Exp. 1/2 were corrected for outliers by excluding RTs from the analysis that were below 100ms or above a participant’s mean RT plus 2.5sd’s.

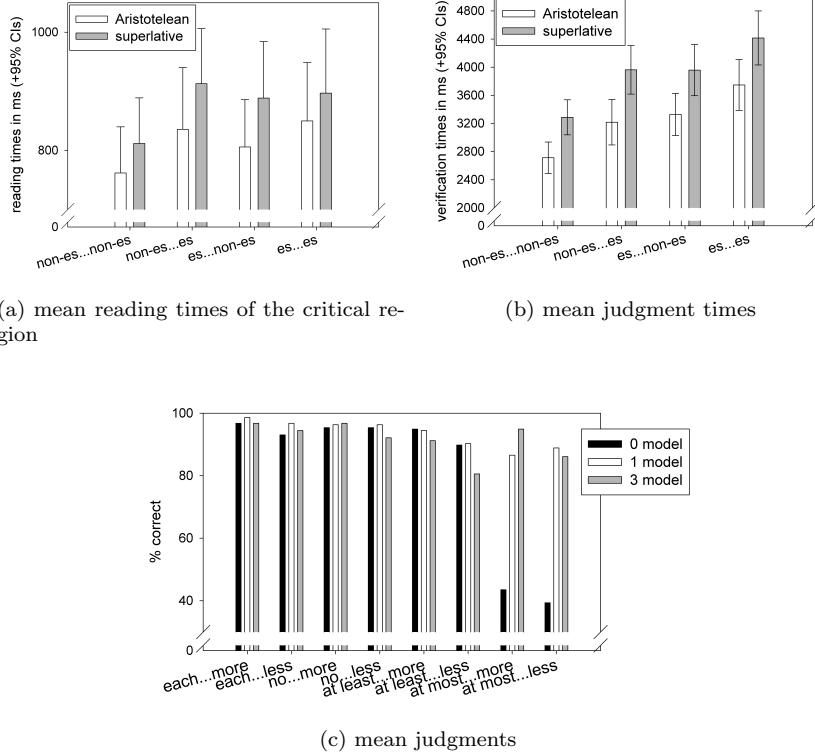


Figure 2: Results of Experiment 2. *es* = empty-set Q ; *non-es* = non-empty-set Q .

ein (*at most one*, empty-set). Q_2 was either *mehr als zwei/drei* (*more than two/three*, non-empty-set) or *weniger als zwei/drei* (*less than two/three*, empty-set). (8) is a sample sentence, asterisks indicate segmentation for self-paced reading. The underlined noun phrase of the second quantifier marks the critical region since this is the earliest point at which a verification algorithm can be fully specified.

- (8) Mindestens ein * Junge * kitzelte * mehr * als drei * Mädchen.
 At least one * boy * tickled * more * than three * girls.

Participants read these sentences in a self-paced reading experiment with moving window presentation. After each sentence, a set diagram appeared on the screen of the types shown in Figure 1 and they had to provide a truth value judgment. As in Exp. 1 there were three types of diagrams: 0-, 1- and 3-set diagrams which showed a) no, b) one or c) three boys tickling Q_2 girls, respectively. Accordingly, the experiment employed a factorial 2 (*empty-set* Q_1 vs. *non-empty-set* Q_1) $\times 2$ (*type of* Q_1 : Aristotelian vs. superlative) $\times 2$ (*empty-set* Q_2 vs. *non-empty-set* Q_2) $\times 3$ (*diagram*) within design. Diagrams were constructed in such a way that across sentence conditions always the same set of pictures was used for the 0- vs 1- vs 3-set diagram. Since 0- and 3-models had to be swapped for empty-set vs. non-empty-set- Q_2 s, judgment times had to be collapsed over diagram types to allow for comparison. 72 experimental items plus 78 fillers were constructed and distributed to 24 lists in a latin square. Across the experiment, 50% of the sentences were true. Sentence-picture-pairs were presented in randomized order.

	each ↑	each↓	at least ↑	at least↓	no↑	no↓	at most↑	at most↓
0	1(1/1/1)	4(2/2/2)	1(1,1,1)	4(2/2/2)	3(2/2/2)	3(2/2/2)	3(2/2/2)	3(2/2/2)
1	1(1/1/1)	2(1/2/2)	1(1,1,1)	1(1/2/2)	2(1/2/2)	2(1/2/2)	1(1/2/2)	1(1/2/2)
3	1(1/1/1)	1(1/1/1)	1(1,1,1)	1(1/1/1)	2(1/1/1)	2(1/1/1)	2(1/1/1)	2(1/1/1)

Table 2: Clauses of **c-exp** that had to be applied in Exp. 2. The arrows indicate whether Q_2 was an empty-set quantifier (\downarrow) or not (\uparrow). The clauses for **c-exp**₂ are given in parentheses after the clause numbers for **c-exp**₁. Conditions in boldface indicate quantifier combinations that can be evaluated by **s-exp** in any model. 0 = 0-model, 1 = 1-model and 3 = 3-model.

We analyzed reading times as well as verification latencies by computing repeated measures ANOVAs. Error rates of the subsequent verification stage were analyzed in a logit mixed effects model. The full set of results is presented in Figure 2. Up to the critical noun phrase the only significant effect was the main effect of *empty set* Q_2 ($p_{1/2} < .01$) at regions *fewer/more* and *than n*. Unsurprisingly, *fewer than n* took longer to read than *more than n*. Effects due to the semantic properties of the first quantifier only emerged at the critical noun phrase of the second quantifier phrase. Conditions with two non-empty-set quantifiers were read faster than the conditions involving empty-set quantifiers (ANOVAs: *empty set* $Q_1 \times *empty set* Q_2 , $p_1 < .05$; $p_2 = .09$) in line with our second prediction that an algorithm based on **s-exp** is inherently less complex than an algorithm based on **c-exp**.$

Both empty-set subject and empty-set object quantifiers led to a slowdown of judgment times as reflected by significant main effects of *empty set* Q_1 ($p_{1/2} < .01$) and *empty set* Q_2 ($p_{1/2} < .01$) but no reliable interaction. Note that this is slightly different from what we would expect on the basis of prediction 1. According to our theory, one empty-set quantifier may already suffice to trigger **c-exp**. Table 2 summarizes which clauses had to be applied in each of the 24 conditions. As can be seen from the table, we would have expected an interaction between *witness sets* Q_1 and *empty set* Q_2 because for an empty-set Q_1 the same clauses were required for empty-set and non-empty-set Q_2 . To test whether this unexpected difference was reliable we computed ANOVAs for the subset of conditions with an empty-set Q_1 . This analysis revealed a main effect of *empty set* Q_2 ($p_{1/2} < 0.1$). Thus, the judgment times provide partial evidence against our theory. As it seems, the combination of two empty-set quantifiers adds processing difficulty beyond what would be expected on the basis of the complexity of the proposed verification algorithms.

Consistent with recent proposals from the semantic literature (e.g., [5]) superlative quantifiers were more difficult to process than Aristotelian quantifiers consistently across the reading and verification stages (main effects of *type of* Q_1 , all $p_{1/2} < .01$) suggesting that they are inherently more complex. In fact, complexity due to the *type of* Q_1 and *empty set* Q_1/Q_2 added up. This was indicated by extraordinarily high error rates of sentences with a superlative empty-set Q_1 . Accordingly, a logit mixed effects model analysis of the superlative conditions revealed a reliable two-way interaction between *empty set* Q_1 and *diagram type* ($z = 8.98$). As in Exp. 1, errors were limited to those models which required the evaluation of a empty-set quantifier with an empty predicate, further confirming prediction 1. This time, polarity reversal led to performance even below chance level, in particular in sentences with the two empty-set quantifiers *at most* and *less than* ($\chi^2(1) = 5.06$; $p < .05$).

The results of Exp. 2 are for the most part consistent with our quantification theory. The observed pattern of reading times supports the distinction between **s-exp** and **c-exp** and the verification data lend support to the predicted difficulty of empty-set quantifiers when having to evaluate a situation in which the scope of the subject quantifier phrase consists of the empty set. Only the judgment times are not fully compatible with our theory.

4 Conclusions

We have presented an algorithmic theory of quantifier interpretation which predicts that the operation employed in processing empty-set quantifiers is more difficult than that for processing non-empty-set quantifiers. This prediction was mainly confirmed in two experiments that investigated the online comprehension and verification of simply and doubly quantified sentences.

In order to further disentangle complexity effects that are due to the presence of empty-set quantifiers from complexity effects due to the monotonicity of the quantifiers, we plan to extend this line of research and investigate whether non-monotonic Q 's which are not empty-set Q 's (e.g. *exactly one boy or exactly three boys*) are easier to process than non-monotonic empty-set Q 's (e.g. *no boy or exactly three boys*), a prediction which - to our knowledge - is unique to the present account.

References

- [1] O. Bott and F. Schlotterbeck. Incremental truth value judgments. In B. Stolterfoht and S. Featherston, editors, *Empirical approaches to linguistic theory*, pages 3–28. Mouton de Gruyter, Berlin, 2012.
- [2] H. Clark and W. Chase. On the process of comparing sentences against pictures. *Cognitive Psychology*, 3:472–517, 1972.
- [3] B. Geurts. Existential import. In I. Comorovski and K. von Heusinger, editors, *Existence: Semantics and syntax*, pages 253–271, Dordrecht, 2007. Springer.
- [4] B. Geurts, N. Katsos, J. Moons, and L. Noordman. Scalar quantifiers: logic, acquisition, and processing. *Language and Cognitive Processes*, 2010.
- [5] B. Geurts and R. Nouwen. ”At least” et al.: the semantics of scalar modifiers. *Language*, 83(3):533–559, 2007.
- [6] B. Geurts and F. van der Slik. Monotonicity and processing load. *Journal of Semantics*, 22:97–117, 2005.
- [7] J. Koster-Moeller, J. Varvoutis, and M. Hackl. Verification procedures for modified numeral quantifiers. In N. Abner and J. Bishop, editors, *Proceedings of the 27th West Coast Conference on Formal Linguistics*, Somerville, MA, 2008. Cascadilla Proceedings Project.
- [8] C. T. Mcmillan, R. Clark, P. Moore, C. Devita, and M. Grossman. Neural basis for generalized quantifiers comprehension. *Neuropsychologia*, 43:1729–1737, 2005.
- [9] Y. Moschovakis. Sense and denotation as algorithm and value. In J. Oikkonen and J. Väänänen, editors, *Lecture Notes in Logic*, vol. 2, pages 210–249. Association for Symbolic Logic, A.K. Peters, Ltd, Natick, MA, 1994.
- [10] S. Steinert-Threlkeld and T. F. I. III. Iterating semantic automata. *Linguistics and Philosophy*, 36:151–173, 2013.
- [11] J. Szymanik and M. Zajenkowski. Comprehension of simple quantifiers: Empirical evaluation of a computational model. *Cognitive Science*, 34(3):521–532, 2010.
- [12] J. van Benthem. *Essays in Logical Semantics*. Reidel, 1986.

Universal quantification as iterated conjunction

Dylan Bumford*

New York University
dbumford@gmail.com

Abstract

I analyze distributive universal quantifiers like ‘each’ and ‘every’ in terms of iterated dynamic update. I argue that this minor adjustment to standard dynamic setups has at least two empirical advantages. First, because information flows forward through the universal computation, anaphoric elements can assume “quantifier-internal” interpretations [1]. Second, because conjunction is usually analyzed as relation composition over input and output structures, the emerging representations are in a sort of disjunctive normal form that facilitates “functional” readings of indefinites. Following [13], I suggest that these two phenomena are closely related, and argue that the current approach which generates the two readings via the same compositional mechanism is simpler, more general, and more empirically adequate than the alternatives.

1 Introduction

In many dynamic frameworks for natural language semantics, sentences denote relations over some sort of data structure. For instance, several prominent fragments in the compositional wake of DRT define sentence meanings in terms of constraints on pairs of (sets of) assignment functions, including DPL [6], CDRT [9], and PCDRT [1]. In most of these frameworks, sentential conjunction is modeled as relation composition: the meaning of $\phi \wedge \psi$ is a relation between input and output structures i and j just in case there is an intermediate structure k such that k is a possible output of ϕ evaluated at i , and j is a possible output of ψ evaluated at k . This composition is sometimes called dynamic conjunction because the intermediate structure k acts as an emissary between the conjuncts, transmitting anaphoric and truth-conditional information along the discourse gradient from ϕ to ψ .

In what follows, I will argue that distributive universal quantification consists in iterated applications of this basic sequencing operation, just as the classical universal quantifier \forall generalizes static Boolean conjunction. A sentence like (1a) will be analyzed in terms of the sequence in (1b), assuming the relevant students are Mary, John, and Bill.

- (1) a. Every student read a book
- b. $\llbracket \text{Mary read a book} \rrbracket ; \llbracket \text{John read a book} \rrbracket ; \llbracket \text{Bill read a book} \rrbracket$

Given (i) the well-known algebraic connection between Boolean conjunction and universal quantification, (ii) the common treatment of sentential conjunction as sequential update (i.e. relation composition) in dynamic semantics, and (iii) the cross-linguistic tendency for conjunctive coordinations to be order-sensitive [14], I take it that this analysis is on plausible theoretic ground. More importantly, it delivers a couple of empirical patterns that have caused a good deal of grief for compositional semantics. First, the iterated update strategy paves the way for a uniform treatment of “internal” uses of adjectives of comparison, as in (2a). Second, when the nuclear scope contains a source of nondeterminism — typically a disjunction or indefinite DP — the strategy generates effectively functional discourse representations, of the form sketched

*Thanks to Chris Barker, Lucas Champollion, Simon Charlow, and Anna Szabolcsi for discussion and advice.

in (2b). And in any framework that identifies semantic nondeterminism as the *vis viva* behind exceptional scope and exceptional binding [4], the iterated update strategy also provides an explanation for the surprising pair-list reading of sentences like (2c).

- (2) a. Every year I buy {a different, a new, another, a faster} computer
- b. Each guest brought a certain dish
 $= \exists f : \text{guest} \rightarrow \text{dish}. \forall x \in \text{guest}. \text{brought}(fx) x$
- c. If each guest brought a certain dish, the party was surely a success [11]: (23)
 $= \exists f : \text{guest} \rightarrow \text{dish}. (\forall x \in \text{guest}. \text{brought}(fx) x) \Rightarrow \text{success the.party}$

To formalize the iterated update approach, I will present a fragment in the style of Dekker's [5] Predicate Logic with Anaphora. Because of the way discourse referents accumulate monotonically over the course of a computation, the stacks of PLA make it particularly clear how the relevant functional objects take shape in the dynamic contrail of iterated conjunction. The next section lays out the relevant parts of this PLA-ish system, and Section 3 focuses the technique on the empirical issues highlighted in (2). Section 4 discusses the coverage of the procedure, and compares it to some of the alternative analyses on the market.

2 Predicate Logic with Anaphora, e.g.

As presented in [5], PLA is an update logic in which formulas are interpreted with respect to the usual model and variable assignment parameters, as well as sets of stacks that record anaphoric possibilities. To bring out the resemblance to other compositional dynamic treatments, I recast PLA in familiar Montagovian terms.¹

To this end, sentences denote relations over stacks, or equivalently, functions from stacks to sets of stacks, of type $p \equiv sst$. Dynamically charged items modify their input stacks. Predicates, by and large, simply perform tests on the values of bound variables.

(3)	<i>Item</i>	<i>Type</i>	<i>Denotation</i>
	John	$(ep)p$	$\lambda Qs. Q j(s \cdot j)$
	book	ep	$\lambda xs. \{s \mid \text{book } x\}$
	read	$((ep)p)ep$	$\lambda Qys. Q(\lambda xs'. \{s' \mid \text{read } x y\}) s$
	a	$(ep)(ep)p$	$\lambda PQs. \bigcup \{Q x (s' \cdot x) \mid s' \in P x s\}$

Here, the name 'John' denotes a dynamic generalized quantifier. It passes the entity corresponding to John into its continuation, along with an updated discourse stack, now listing j as its most recent (final) element. 'Book' tests an individual for book-hood; if it passes, the input stack is the only thing returned unmodified, but if it fails, the input stack is discarded and the empty set returned instead. The transitive verb 'read' has been argument raised for simplicity. Like 'book', it fails quietly when its arguments do not stand in the appropriate relation, and returns just its input when they do.

These three values are *deterministic*, in the sense that when fed an input, they return at most one output. The indefinite article, on the hand, is in general *nondeterministic*. At a single input, the indefinite generates at least as many updated stacks as there are witnesses of its restrictor

¹This re-coding of PLA follows that in [4]. The notational conventions are, I hope, mostly standard: Model-theoretic entities/relations/etc. are given in sans-serif. Variables are in math-italics x, y, z , etc., with the usual types ($x-z$ of type e ; P, Q of type ep ; s, s' for stacks). Function application is left associative; type descriptions are right associative; parentheses are omitted, except where necessary for grouping. If $s = [s_0, \dots, s_n]$, and $i \geq 0$, then s_i identifies the i th element from the beginning of s ; if $i < 0$, then s_i picks out the i th element from the end of s . $s \cdot x = [s_0, \dots, s_n, x]$; and $s \cdot s' = [s_0, \dots, s_n, s'_0, \dots, s'_{n'}]$. The interpretation brackets $\llbracket \cdot \rrbracket$ map each item onto its denotation, where $\llbracket X \dots Z \rrbracket = \llbracket X \rrbracket \dots \llbracket Z \rrbracket$. $\lambda x_0 \dots x_n. M$ is short for $\lambda x_0 \dots \lambda x_n. M$.

(more if the restrictor contains indefinites of its own). Then each of these updated stacks is threaded through the nuclear scope, where it is potentially subjected to further manipulation, as in (4c). In this fashion, the dynamic effects of the restrictor and nuclear scope are crossed, so that the final denotation of (4d) is a function from an input stack to a set of extending output stacks, one for each pair in `man` × `book`.

- (4) a. $\llbracket \text{a book} \rrbracket = \lambda Qs. \bigcup \{Qx(s' \cdot x) \mid s' \in \{s \mid \text{book } x\}\} = \lambda Qs. \bigcup \{Qx(s \cdot x) \mid \text{book } x\}$
- b. $\llbracket \text{read [a book]} \rrbracket = \lambda ys. \bigcup \{\{s \cdot x \mid \text{read } xy\} \mid \text{book } x\} = \lambda ys. \{s \cdot x \mid \text{book } x \wedge \text{read } xy\}$
- c. $\llbracket \text{John [read [a book]]} \rrbracket = \lambda s. \{s \cdot j \cdot x \mid \text{book } x \wedge \text{read } xj\}$
- d. $\llbracket \text{A man [read [a book]]} \rrbracket = \lambda s. \{s \cdot y \cdot x \mid \text{book } x \wedge \text{man } y \wedge \text{read } xy\}$

This last point is important. The indefinite article multiplies nondeterminism in its restrictor and nuclear scope by essentially *composing* the latter with the former. In fact, as indicated in Section 1, this is the general strategy for conjoining propositional meanings in dynamic systems. Thus the generic entry for sentential ‘and’ looks very much like that for ‘a’, but with a simpler type.² Both operators are in a sense prepared for nondeterministic arguments, and both respond by distributing the respective possibilites pointwise. So if there are several ways of making *A* true (in the sense that there are several witnesses for some nondeterministic source inside of it), and there are several ways of making *B* true, then saying ‘*A* and *B*’ amounts to saying that *some* combination of an *A* verifier and a *B* verifier is true. In other words, relation composition has disjunctive normal form in its bones; it turns ‘(A_0 or … or A_i) and (B_0 or … or B_j)’ into ‘($A_0 \wedge B_0$) or ($A_0 \wedge B_1$) or … or ($A_i \wedge B_j$)’.

(5)	<i>Item</i>	<i>Type</i>	<i>Denotation</i>
	and	ppp	$\lambda pq. p ; q \equiv \lambda pqs. \bigcup \{q s' \mid s' \in p s\}$
	every	$(ep)(ep)p$	$\lambda PQs. ; \{ \lambda s'. Qx(s' \cdot x) \mid Pxs \neq \emptyset \} s,$ where $; \{A_0, A_1, \dots, A_n\} \equiv A_0 ; A_1 ; \dots ; A_n$

The semantics of ‘every’, as promised, folds the compositional procedure embodied by ‘and’ over a set of propositions built by mapping the restrictor over the nuclear scope.^{3,4} ‘Every’ and ‘a’ have much in common here, as do ‘and’ and ‘or’ ($= \lambda pqs. \bigcup \{p s, q s\}$), of which they are generalizations. Both determiners can be seen as building a Cartesian matrix of possible computational threads from the crossed witness sets of their restrictors and scopes, but they differ in what they do with that matrix. The indefinite article does almost nothing, except reduce the dimensionality of the resulting set of sets, in accordance with its type. The universal determiner, however, *sequences* the rows of that matrix, each of which corresponds to a dynamic proposition with a single point of nondeterminism (in the simplest case).

Take the sentence in (4d) for example, and its universal counterpart in (6) (which is α -equivalent to the $\lambda s. ; \{\dots\} s$ entry given in (5)).

²Note that this is just the usual $\llbracket \phi \text{ and } \psi \rrbracket = \lambda ss''. \exists s'. s[\phi]s' \wedge s'[\psi]s''$ dynamic denotation for ‘and’ translated into set-theoretic terms.

³Technically, because $;$ is not in general commutative, the semantics must make a choice about the order in which the elements of the restrictor should be evaluated. As we will see in Section 3, this choice usually does not make any difference for the resulting truth conditions, so I assume it is random. When it does make a difference, I assume it is determined by context and world knowledge.

⁴Though this entry for ‘every’ will pass information from one restrictor element to another, it will not pass along information from the restrictor to the scope. This simplification makes it easier to think about functional readings and internal ‘different’, but it means that donkey anaphora out of DP is out of the question. [3] provides a version of this universal in a monadic semantics that feeds information through both semantic dimensions, though the empirical facts concerning ‘different’ donkey sentences are entirely uncharted.

$$(6) \llbracket \text{Every man [read [a book]]} \rrbracket = ;\{\lambda s'. (\lambda ys. \{s \cdot x \mid \text{book } x \wedge \text{read } x y\}) y (s' \cdot y) \mid \text{man } y\} \\ = ;\{\lambda s'. \{s' \cdot y \cdot x \mid \text{book } x \wedge \text{read } x y\} \mid \text{man } y\}$$

In both sentences, we are dealing with the set of propositions $\{\text{John read a book}, \text{Dave read a book}, \text{Fred read a book}, \dots\}$, one for each man. The indefinite article is content to merely pass in the input stack to each alternative *in parallel*, and to collect the results with a big \bigcup . But the universal runs the sentences off *in sequence*, as sketched in (7). This is just the semantic reflex of taking ‘a’ to generalize ‘or’ and ‘every’ to generalize ‘and’.

But then we are in the situation mentioned earlier in which we are composing several nondeterministic propositions. The result is again in a kind of disjunctive normal form. It is true iff a suitable output stack can be found, i.e. iff each man can be paired with a book that he read. If any of the men in fact did not read any books, then the particular proposition that he is responsible for (e.g. ‘Fred read a book’) will be equivalent to $\lambda s'. \{\}$, the dynamic analog of falsity. The emptiness will inevitably swallow the entire update.

$$(7) \quad (\lambda s'. \{s' \cdot m_0 \cdot x \mid \text{book } x \wedge \text{read } x m_0\}) ; \quad (\lambda s'. \{s' \cdot m_1 \cdot x \mid \text{book } x \wedge \text{read } x m_1\}) ; \\ \vdots \quad \rightsquigarrow \lambda s. \left\{ \begin{array}{l} s \cdot m_0 \cdot b_{00} \cdot m_1 \cdot b_{10} \cdot \dots \cdot m_n \cdot b_{n0} \\ s \cdot m_0 \cdot b_{00} \cdot m_1 \cdot b_{10} \cdot \dots \cdot m_n \cdot b_{n1} \\ \vdots \\ s \cdot m_0 \cdot b_{00} \cdot m_1 \cdot b_{10} \cdot \dots \cdot m_n \cdot b_{ni} \\ \vdots \\ s \cdot m_0 \cdot b_{0i} \cdot m_1 \cdot b_{1j} \cdot \dots \cdot m_n \cdot b_{nk} \end{array} \right\}$$

3 Applications

3.1 Internal Adjectives

The application of iterated conjunction to internal adjectives is immediate. The reason ‘different’ or ‘new’ or ‘faster’ can seem to straddle multiple elements of a universal restrictor is that the information from early evaluations feeds forward to later evaluations. Consider (2a) again. At each year, we iterate again through the VP, selecting as we go a witnessing computer for each potential branch of the update. Once a particular computer has been selected, it is recorded on that branch’s stack. At the next iteration, every update to the branch will be able to see which computers have already been selected to witness the previous years’ updates. This means that all ‘different’, e.g., has to do is restrict its complement’s extension to those entities not already on the relevant stack. Likewise, ‘faster’ simply restricts its complement to those entities that are faster than the next fastest entity on the input stack.

Item	Type	Denotation
different	(ep)ep	$\lambda Pxs. \{s' \mid s' \in P x s \wedge x \notin s'\}$
faster	(ep)ep	$\lambda Pys. \{s' \mid s' \in P x s \wedge \text{speed } x > \max\{\text{speed } u \mid u \in s \wedge P u s \neq \emptyset\}\}$

(9) presents an internally modified analog of (1a). As the derivation unfolds, one by one the assembled propositions contribute a student and a book to each of the potential update stacks. Crucially, on any given thread of the update, the new book must be different from the books already witnessing the VP on behalf of the other individuals. So if we are at, say, John’s loop through the VP and we cannot find a stack to contribute to (either because John didn’t read any books, or because he read a subset of the books that Mary did), then the entire update fails. If, however, there is any pairing of books to students in such a way that no book is repeated, then the update succeeds. Ultimately, all such injective pairings are retained as potential outputs of the sentence.

- (9) a. $\llbracket \text{a [different book]} \rrbracket = \lambda Qs. \bigcup \{Qx (s \cdot x) \mid \text{book } x \wedge x \notin s\}$
 b. $\llbracket \text{John [read [a [different book]]]} \rrbracket = \lambda s. \{s \cdot j \cdot x \mid \text{book } x \wedge x \notin s \wedge \text{read } x j\}$
 c. $\llbracket \text{Every student read a different book} \rrbracket =$
 $\llbracket \text{Mary read a different book} \rrbracket ; \llbracket \text{John read a different book} \rrbracket ; \llbracket \text{Bill read a different book} \rrbracket$
- $$(\lambda s'. \{s' \cdot m \cdot x \mid \text{book } x \wedge x \notin s' \wedge \text{read } x m\}) ;$$
- $$(\lambda s'. \{s' \cdot j \cdot x \mid \text{book } x \wedge x \notin s' \wedge \text{read } x j\}) ; \rightsquigarrow \lambda s. \left\{ \begin{array}{l|l} s \cdot m \cdot x & x, y, z \in \text{book}, \\ \cdot j \cdot y & \langle m, x \rangle, \langle j, y \rangle, \langle f, z \rangle \in \text{read}, \\ \cdot f \cdot z & x \notin s, \quad y \notin s \cdot m \cdot x, \\ \dots & z \notin s \cdot m \cdot x \cdot j \cdot y \end{array} \right.$$

The same procedure generates internal readings for morphologically comparative adjectives, as well as adjectives that depend on an implicit ordering source, like ‘new’. The only difference is that since these adjectives denote *asymmetric* relations, the resulting truth conditions depend on the order in which the elements are evaluated. For this reason, internal comparatives are only natural when their distributor’s restrictor is either inherently ordered — usually by time — or when some ordering of the restrictor elements is recoverable from the context. For instance, (10) sketches the derivation of the comparative version of (2a). The only stacks that survive the update are those that can be incrementally updated with a sequence of computers that I bought in consecutive years, each faster than the ones preceding it.

- (10) a. $\llbracket \llbracket \text{In 2010}, \text{I [bought [a [faster computer]]]} \rrbracket \rrbracket =$
 $\lambda s. \{s \cdot 2010 \cdot x \mid \text{comp } x \wedge x > \max \{\text{speed } u \mid \text{comp } u \wedge u \in s\}\}$
 b. $\llbracket \text{Every year, I bought a faster computer} \rrbracket =$
 $\llbracket \text{In 2009, I bought a faster computer} \rrbracket ; \llbracket \text{In 2010, I bought a faster computer} \rrbracket ; \dots$
 c. $\lambda s. \left\{ \begin{array}{l|l} s \cdot 2009 \cdot x & x, y, z, \dots \in \text{comp}, \quad \langle j, x, 2009 \rangle, \langle j, y, 2010 \rangle, \langle j, z, 2011 \rangle, \dots \in \text{bought}, \\ \cdot 2010 \cdot y & x > \max \{\text{speed } u \mid \text{comp } u \wedge u \in s\}, \\ \cdot 2011 \cdot z & y > \max \{\text{speed } u \mid \text{comp } u \wedge u \in s \cdot 2009 \cdot x\}, \\ \dots & z > \max \{\text{speed } u \mid \text{comp } u \wedge u \in s \cdot 2009 \cdot x \cdot 2010 \cdot y\}, \dots \end{array} \right\}$

3.2 Functional Readings

Returning to the example in (6), we said that ‘Every man read a book’ is true just in case there is a pairing of men with books such that each man read the book he is paired with. These truth conditions are equivalent to the usual $\forall \gg \exists$ variety, and are in fact exactly what we’d get by Skolemizing the indefinite and existentially closing over the resulting parameterized function, as in the gloss of (2b). But we got to this point simply by sequencing a number of nondeterministic propositions and letting the nondeterminism bubble through the conjunctions. When this happens, we are left with a fundamentally *disjunctive* kind of meaning, the sort of meaning that takes a single input stack to a (potential) multiplicity of output stacks. Each of these output stacks will correspond to a function $f : \text{man} \rightarrow \text{book}$ mapping each man to one of the books he read.

Following [10, 8, 11], and especially [4], I will assume that nondeterminism is at the heart of exceptional scope; indefinite DPs and disjunctions of any category can outscope any chunk of their compositional context, while quantificational DPs and conjunctions of any category are effectively clause-bounded.⁵ If this is right, then we already have an account of the “functional

⁵The unconvincing reader is referred to [4], where the point is given theoretical teeth in a continuations-based semantics where scope-taking possibilities follow directly from semantic types. [3] implements the strategy here in that continuized grammar.

readings” of sentences like (2c).⁶ What’s really taking scope here are the alternative ways of satisfying the universal proposition. (2b), for example, corresponds to a meaning along the lines of “EITHER Mary brought turkey, John potatoes, and Fred casserole, OR Mary brought turkey, John casserole, and Fred potatoes, OR …”, for each possible pairing of individuals and dishes. The functional reading of (2c) then corresponds to the wide scope interpretation of this disjunction with respect to the conditional it is in: “EITHER ‘Mary brought turkey, John potatoes, and Fred casserole’ OR ‘Mary brought turkey, John casserole, and Fred potatoes’ OR … is such that if it came true, then the party was surely a success”.

It is hard to be more precise about this without an explicit theory of scope-taking, but that would take us a far afield. Nevertheless, these are exactly the truth conditions predicted by the Skolem-function gloss typically assigned to such readings, exemplified by the translation in (2c). But again, they are generated without any explicit representation of functional variables as part of the meaning of either DP.

That said, we may occasionally want to make use of the implicit functions generated by iteration, *qua* functions. For example, the little discourse in (11a) intuitively sets up a functional relationship between students and books, and then elaborates on that relationship. The second sentence makes anaphoric reference to whichever alternative happens to witness the first sentence, and when it retrieves that alternative, it retrieves it *as a function*.

One way to view this sort of functionalization on the fly is just as an iterated version of well known cases of pluralization on the fly, as in ‘A man walked in, and then a woman walked in. They sat down together.’ To give an example of how this might play out, the entry for ‘other’ in (11b) deploys a functionalizing operator \mathcal{F} that replaces a (contextually-determined) subportion of the input stack with a function f by pairing off the relevant alternating entities, so that $[m, x, j, y, f, z]$ becomes $f \equiv \{\langle m, x \rangle, \langle j, y \rangle, \langle f, z \rangle\}$. Then ‘other’ returns a dynamic property true of entities that are distinct from what the function assigns to the most recent entry on the stack (which will be the individual added by the current pass through the local distributor’s restrictor). In this sense, ‘other’ behaves exactly like a quantificationally-subordinated pronoun, in that it retrieves an anaphoric *ee*-type dependency from the discourse, and uses it to fix the reference of some bound pronominal element. (11c) gives some idea of what this looks like in action, where $\bullet f = [x_0, f(x_0), x_1, f(x_1), \dots]$, for x_0, x_1, \dots in the domain of f .

(11) a. Every student read a book. Then every student read another book.

$$\begin{aligned} b. \llbracket \text{other} \rrbracket &= \lambda Pxs. \{s' \mid s' \in Px(\mathcal{F}s) \wedge x \neq s'_f(s'_{-1})\} \\ c. \lambda s. \left\{ s \cdot \bullet f \mid \begin{array}{l} f : \text{student} \rightarrow \text{book} \\ f \subseteq \text{read} \end{array} \right\} ; \lambda s. \left\{ s \cdot \bullet g \mid \begin{array}{l} g : \text{student} \rightarrow \text{book}, \quad g \subseteq \text{read} \\ g(m) \neq f((s \cdot m)_{-1}), \\ g(j) \neq f((s \cdot m \cdot g(m) \cdot j)_{-1}), \\ g(f) \neq f((s \cdot m \cdot g(m) \cdot j \cdot g(j) \cdot f)_{-1}) \end{array} \right\} \end{aligned}$$

4 Discussion and Comparison

Most accounts of functional readings have appealed to Skolemized choice-functional representations of indefinite determiners (e.g., [12, 11]). To this end, an indefinite DP like ‘a certain dish’ is represented in the semantics as $f x \text{dish}$, where the f variable, ranging over type $e(et)e$ (paramaterized) choice functions, will eventually be existentially or contextually bound, and the x variable will eventually be bound by the universal.

⁶I follow [7, 12, 11, 13] in calling the relevant truth conditions of ‘every’ \gg ‘a’ sentences “functional readings”, though as several of those authors have pointed out, there is good reason to keep the notion of a functional reading separate from that of the arbitrary pair-list interpretations described here. We return to the issue briefly in Section 4.

However, as both [12] and [11] point out, the choice-functional analysis dramatically over-generates in non-upward-entailing contexts. (12), for example, is falsely predicted to mean that no student read every book. They both conclude, then, that choice function variables only ever refer to “natural functions”, which rules out the arbitrary potential functional witnesses that render the logical form of (12) so weak.

- (12) No student read a book that I recommended
 $\neq \exists f. \neg \exists x \in \text{student}. \text{read } x (f \ x \ \text{book}.!.\text{recommended})$

Yet, as [13] argues, this puts too strong a bind on the potential functions that may emerge from ‘every’ \gg ‘a’ configurations, and doesn’t really rule out the improper readings of other quantifiers in the same positions. [11] makes a similar point regarding the difference between pair-list and “natural functional” answers to questions. (13) can be answered by naming a function that intensionally maps guests to dishes, or by specifying the contents of such a mapping in an arbitrary way. But the same question with any non-distributive-universal quantifier (14) can only be answered with the former.

- (13) Which dish did every guest bring? (14) Which dish most/several/no guests bring?
a. His favorite a. Their favorite
b. Al pasta; Bill salad; Carl pudding b. # Al pasta; Bill salad; Carl pudding

[13] observed that the same contrast is evident in conditionals like (2c). The sentences in (15a) and (15b) lack any hint of the arbitrary pair-list assignment of guests to dishes that underlies the truth conditions of (2c). That is, (15a) doesn’t mean that for some pairing of guests to dishes, if any two guests brought their dish, the party probably went well; and (15b) doesn’t mean that if no guest brought their assigned dish, the party was a flop. If they have any functional reading at all, it depends on ‘a certain dish’ being interpreted as some sort of functional definite description, à la ‘the one his mother suggested’. What’s more, no restriction on “naturalness” will suffice to rule out the functional readings of (15), since precisely the same functions verify the available reading of (2c).

- (15) a. If two guests brought a certain dish, the party was probably a success
b. If no guest brought a certain dish, the party was surely a failure

[2] presents a number of other problems for generalized choice-functional analyses of exceptionally scoping indefinites. The authors suggest instead that functional readings instantiate a kind of quantificational subordination to a contextually salient association. However, like the choice-functional approaches, this doesn’t explain the general ability of ‘every’/‘each’ to support such readings, as in (2c), and the general inability of all other quantifiers to do so, as in (15). What’s more, the existential force of the quantification over functional witnesses can take intermediate scope beneath other operators, which is unexpected if the witness is due to some kind of contextual anaphora. In fact, (16) has a reading on which the pair-list that witnesses the embedded ‘hope’ clause donkey-binds the propositional anaphor in the consequent: whenever Mary has a specific wishlist about the dishes people will bring, *whatever it happens to be*, I have the same wishlist. It’s hard to see how this could be derived from a deictic dependency.

- (16) Whenever Mary hopes that everyone brings a certain dish, I hope so too

The lesson from functional readings seems to be that something special happens when plain indefinites are within the nuclear scope of universal distributors. Then, and only then, can arbitrary nondeterministic pair-list associations emerge between the two quantificational restrictors. As a result, all of the attempts to locate the expressive force of suc readings in the ability of the indefinite determiner to go proxy for a Skolemized choice function are bound to overgenerate, because they pay no attention to the special role of the universal.

In contrast, theories of internal ‘different’ typically rely on a souped-up universal because it is well-established that singular comparatively modified NPs can only be interpreted “internally” in the scope of a genuinely distributive universal quantifier. So [1], for instance, develops a dynamic semantics in which ‘every’ temporarily generates two independent update streams, which are simultaneously universally quantified over. ‘Different’ bridges the anaphoric gap between the quantificational channels, in the same way that it might compare items across sentences. This is very much in the same spirit as what I’ve proposed here, but with iterated conjunction, there’s no need for dual quantification. The universal quantifier just plows through its restrictor, accumulating referents as it goes.

5 Conclusion

That “internal” and “functional” readings have the same very particular syntactic distribution should be a clue that they are derived from the same underlying mechanism. While polyadic quantification generates appropriate truth conditions for internal uses of comparative adjectives (at least, for symmetric relations like ‘same’ and ‘different’), it doesn’t give us any leg up on the emergence of exceptionally scoping/binding functional witnesses. And while Skolemized choice functions generate appropriate truth conditions for ‘every’ \gg ‘a’ configurations, they generate inappropriate truth conditions for all other configurations.

The analysis presented here, on the other hand, derives both kinds of readings as by-products of iterated conjunction, in any semantics that treats conjunction as relation composition. This, as [4] stresses, is standard in dynamic setups. As a result, the approach is both more general and more conservative than the choice-functional and polyadic solutions currently in the literature.

References

- [1] Adrian Brasoveanu. Sentence-internal *different* as quantifier-internal anaphora. *Linguistics and Philosophy*, 34(2):93–168, 2011.
- [2] Adrian Brasoveanu and Donka Farkas. How indefinites choose their scope. *Linguistics and Philosophy*, 34(1):1–55, 2011.
- [3] Dylan Bumford. Incremental quantification. Unpublished ms, NYU, 2013.
- [4] Simon Charlow. *Sub-clausal dynamics*. PhD dissertation, New York University, 2013.
- [5] Paul Dekker. Predicate logic with anaphora. In *Proceedings of SALT 4*, pages 79–95, 1994.
- [6] Jeroen Groenendijk and Martin Stokhof. Dynamic predicate logic. *Linguistics and Philosophy*, 14:39–100, 1991.
- [7] Jaakko Hintikka. The semantics of *a certain*. *Linguistic Inquiry*, 17(2):331–336, 1986.
- [8] Angelika Kratzer and Junko Shimoyama. Indeterminate pronouns. In *Proceedings of the Third Tokyo Conference on Psycholinguistics*, pages 1–25, Tokyo, 2002.
- [9] Reinhard Muskens. Combining Montague semantics and Discourse Representation. *Linguistics and Philosophy*, 19(2):143–186, 1996.
- [10] Mats Rooth and Barbara Partee. Conjunction, type ambiguity and wide scope *or*. In *Proceedings of WCCFL 1*, pages 353–362, 1982.
- [11] Philippe Schlenker. Scopal independence. *Journal of Semantics*, 23(3):281–314, 2006.
- [12] Bernhard Schwarz. Two kinds of long-distance indefinites. Unpublished ms, UT Austin, 2001.
- [13] Mike Solomon. True distributivity and the functional interpretation of indefinites. <http://semanticsarchive.net/Archive/zkxN2M4M/solomon-functional-indefinites.pdf>, 2011.
- [14] Anna Szabolcsi. What do quantifier particles do? <http://ling.auf.net/lingbuzz/001857>, 2013.

‘Than’ = ‘More’ + Exhaustivity: Evidence from Circassian*

Lisa Bylinina¹ and Yury Lander²

¹ Meertens Instituut, Amsterdam, Netherlands

bylinina@gmail.com

² National Research University Higher School of Economics, Moscow, Russia

yulander@yandex.ru

Abstract

This paper discusses new facts on comparative constructions in Circassian languages that contribute to a theoretical debate about the semantics of comparatives. We argue that Circassian comparatives provide the direct evidence for the combination of two recent insights into comparative semantics: the theory of two loci of degree quantification in such constructions [1, 9] and theories postulating exhaustivity or maximisation at the edge of the standard clause [2].

1 Introduction

The goal of this work is twofold: first, it brings in new facts on comparative constructions in Circassian languages (a branch of Northwest Caucasian); second, it contributes to a theoretical debate about the semantics of comparative constructions. We will argue that Circassian comparatives provide the direct evidence that has been missing so far for the combination of two recent insights into comparative semantics: the theory of ‘two loci of degree quantification’ in such constructions [1, 9] and theories postulating ‘exhaustivity’ or ‘maximisation’ at the edge of the standard clause, hypothetically associated with the standard morpheme *than* and its analogues in other languages [2].

As these insights have been introduced independently of each other, an extra step will be needed to glue the two analyses together. The proposal presented here solves this task and sheds light on the morphological make-up of Circassian comparatives that otherwise would have remained a mystery.

Let us start with the theoretical background on the semantics of the comparative constructions, followed by a short introduction to Circassian languages and a formulation of a challenge they present to the standard theories of comparison. Section 2 presents the ingredients we will need to make sense of the Circassian data. There will be two ingredients: the idea of two sources (or two loci) of degree quantification in the comparative construction – and the idea of a maximisation or exhaustivity operation at the edge of the standard clause in the comparative construction. In section 3, we apply these ideas to Circassian data, developing our analysis of comparatives in Circassian languages.

1.1 Comparatives: Standard analyses

In a comparative construction in (1), *-er* is a comparative marker, or a comparative morpheme; *than* is a standard marker, or a standard morpheme; finally, *than Mary (is)* is a standard phrase:

*The authors thank Rick Nouwen, Roger Schwarzschild, and the anonymous reviewers of Amsterdam Colloquium for discussion.

- (1) John is taller than Mary (is).

Semantically, (1) conveys that the degree that John reaches on the scale of height exceeds the degree that Mary reaches on the same scale:

- (2) $\llbracket \text{John is taller than Mary (is)} \rrbracket = \text{John's height} > \text{Mary's height}$

Several theories of how this interpretation is achieved compositionally has been formulated. The now standard analysis of comparative constructions gives the comparative morpheme the semantics of a quantifier over degrees, as in (3), see [4, 11, 5] a.m.o. – the denotation also known as a ‘2-place MORE’:

- (3) $\llbracket \text{MORE}_2 \rrbracket_{(dt(dt,t))} = \lambda D_{(dt)} \lambda D'_{(dt)} \cdot \mathbf{max}(D') > \mathbf{max}(D)$

The comparative morpheme in (3) takes two sets of degrees and returns ‘true’ if the maximal point of the second set exceeds the maximal point of the first set. How does this 2-place MORE get interpreted? It depends on a set of assumptions about the meanings of other elements in the comparative construction – in particular, the semantics of the standard phrase. In English, *than* is known to be able to take a clausal complement (there are debates on whether it has to, but this is irrelevant at this point), and some part of this clause gets elided, as in (4-a). Practically all the existing analyses of clausal standards of comparison involve degree abstraction in the standard clause, for which the elided material is necessary. Under these analyses, the standard phrase denotes a set of degrees, as in (4-b):

- (4) a. John is taller than Mary is *d-tall*.
 b. $\llbracket \text{than Mary is} \rrbracket = \lambda d. \text{Mary is } d\text{-tall}$

This set of degrees denoted by the standard phrase is supplied as the first argument of a 2-place MORE. The resulting semantic type of this combination is $\langle dt, t \rangle$. The 2-place MORE cannot stay in situ to get its interpretation because of the type mismatch between the adjective (we assume the type of gradable adjectives is $\langle d, et \rangle$) and the comparative morpheme. The DegP (MORE + standard phrase) has to undergo QR leaving the trace of type d behind. This trace of type d can combine felicitously with the gradable adjective with a type $\langle et \rangle$ result, combining later with the subject, the result being a truth value.

The movement of DegP is accompanied by lambda-abstraction over the degree variable, so that the resulting semantic type of the matrix clause would be $\langle dt \rangle$ – quite like the semantic type of the standard clause. Two arguments of type $\langle dt \rangle$ is exactly what MORE can combine with.

DegP movement is also motivated by ellipsis resolution in the *than*-clause, which has been argued to be the case of antecedent-contained deletion. If the DegP with the *than*-clause remains in situ, this would result in infinite regress in the ellipsis site, see [5] a.m.o.

A sentence *John is taller than Mary (is)* would then have the following LF and semantics:

- (5) a. **LF:** $[-\text{er} [\langle dt \rangle \text{ than } 1 [\text{Mary is } t_1\text{-tall}]] [\langle dt \rangle 2 [\text{John is } t_2\text{-tall}]]]$
 b. **Semantics:** $\mathbf{max}(\lambda d. \text{John is } d\text{-tall}) > \mathbf{max}(\lambda d. \text{Mary is } d\text{-tall})$

The driving force of the comparative interpretation in this analysis is the comparative morpheme. The standard morpheme THAN is treated as meaningless and is practically disregarded in the derivation.

An alternative entry for MORE is often suggested, for the cases when the standard is a DP rather than a clause. This alternative entry, often called a 3-place MORE, can be interpreted

‘directly’ to get the same result as (5-b), see [3] a.o.:

- (6) $\llbracket \text{MORE}_3 \rrbracket = \lambda s_e \lambda g_{(d,et)} \lambda x_e. \mathbf{max}(\lambda d. g(d)(x)) > \mathbf{max}(\lambda d. g(d)(s))$
- (7)
 - a. John is taller than Mary.
 - b. **LF:** [John [[-er [than Mary]] $[\lambda d \lambda x. x \text{ is } d\text{-tall}]]]$
 - c. **Semantics:** [John is taller than Mary] = $\llbracket \text{MORE}_3 \rrbracket (\text{Mary})(\lambda d \lambda x. x \text{ is } d\text{-tall})(\text{John}) = \mathbf{max}(\lambda d. \text{John is } d\text{-tall}) > \mathbf{max}(\lambda d. \text{Mary is } d\text{-tall})$

Note that this analysis too treats standard markers (e.g. English *than*) as semantically empty [4, 11].

This assumption of semantic transparency of the standard marker has been questioned recently, based on quite intricate reasoning and indirect evidence. We will discuss two recent insights into the semantics of the standard phrase and its role in the comparative construction: 1) standard markers can introduce degree quantification, quite like comparative morphemes [1, 9]; 2) the edge of standard phrase involves a maximality, or exhaustivity, operator [2].

However convincing the argumentation for these two innovations, what is immediately worrying is the absence of direct evidence for them – in particular, no language has been observed that uses one and the same morpheme both as a comparative and a standard marker, nor has a language been attested with an explicit maximal informativity operator. We discuss novel data from Circassian, which fills both gaps.

Before we discuss these ideas in more detail and apply them to the Circassian comparative, let us give you some background on Circassian languages.

1.2 Introduction to Circassian and Circassian comparatives

The two Circassian languages, namely Adyghe and Kabardian, constitute a branch of the Northwest Caucasian (alias Abkhaz-Adyghe) family. While being originally spoken in the Northwest Caucasus, the Circassian languages are now scattered not only in a few Russian districts, but also in Turkey and some other countries of the Middle East. Adyghe and Kabardian are left-branching ergative polysynthetic languages with very complicated morphology but without productive incorporation. All arguments are cross-referenced in their heads. Further, there are two core cases, namely absolute and oblique, marking primarily non-absolute arguments, including the transitive agent. Overt marking of these cases depends partly on specificity and the lexical type of a noun, cf. [8] a.o.

The data presented here are mainly from the Temirgoi dialect of Adyghe as it is represented in Republic of Adygea, Russian Federation, as well as from Standard Adyghe based on the same dialect. Some data come from the Besleney dialect of Kabardian. The source of the data is original fieldwork by the authors, unless indicated otherwise. While comparatives in Adyghe and Kabardian may differ in details, the relevant facts seem to hold for both languages, with differences stated where needed.

A typical Circassian comparative sentence has a comparative morpheme appearing in front of the gradable adjective and a postposition introducing the standard phrase. These are *nah* ≈ *more* and *nah(r)jə* ≈ *than* respectively in the varieties under discussion ((9) = (28) from [10]). The standard marker *nah(r)jə* assigns Oblique case to the standard DP:

- (8) mə b^weg^wə-r [mwe-drje b^weg^wə-m **nahrjə**] [nah č'əh] BESLENEY KABARDIAN
this road-ABS that-other road-OBL THAN MORE long
'This road is longer than the other one'

- (9) a-r [se-š' **nahjə**] [nah ?^wəšə-^v] ADYGHE
 that-ABS I-OBL THAN MORE smart-PST
 'He was smarter than me'

The constituency marked in (8) and (9) is supported by constraints on word order found in Circassian (the following examples come from Adyghe). Generally, the order between the subject, the standard phrase and the MORE + gradable predicate constituent is free in Circassian. Consider one example where the standard phrase is not adjacent to the MORE + gradable predicate constituent:

- (10) [dnjepre **nahjə**] volye [nah č'əh]
 Dnieper THAN Volga MORE long
 'Volga is longer than Dnieper'

There is evidence that MORE can form a lexical complex with the gradable predicate (see example (14), where this is the case), but we won't concentrate on these facts. Nothing hinges on the morphological status of *nah* in what follows.

Furthermore, an adjunct cannot intervene between the standard marker *nah(r)jə* and the standard DP, as well as between the comparative morpheme *nah* and the gradable predicate, but it can appear in other intermediate positions (data from [10]):

- (11) (njewəš') a-r (njewəš') d-e-pč'e-je-š't (njewəš') [nah (*njewəš')]
 tomorrow that-ABS tomorrow LOC-DYN-jump-UP-FUT tomorrow MORE tomorrow
 Łag-ew] (njewəš') [se-š' (*njewəš') **nahjə**] (njewəš')
 high-ADV tomorrow I-OBL tomorrow THAN
 'Tommorow he will jump higher than me'

Another indirect indication that *nah(r)jə* is related to the standard syntactically and the comparative morpheme is related to the subject (via combination with the gradable predicate) has to do with 'agreement' (the term is used here atheoretically, more accurate term would be 'cross-reference'). Both *nah(r)jə* and *nah* can bear cross-reference morphology, but only *nah(r)jə* cross-references the standard (= (22) from [10]):

- (12) te [t-jo-^wəne^wə-me **a-nahjə**] [**tə-nah-de^wə**]
 we 1PL.IO-3SG.A-neighbour-OBL.PL 3PL.IO-TAN 1PL.ABS-MORE-good
 'We are better than our neighbours'

Finally, semantically, Circassian comparatives don't generally bear evaluativity (from a comparative like 'John is taller than Bill' it doesn't follow that either John or Bill is tall) and can contain differential measure phrases (marked with instrumental case):

- (13) a-r d-e-pč'e-je-š't **mjetr-jə-š'-č'e** [nah Łag-ew] [se-š' **nahjə**]
 that-ABS LOC-DYN-jump-UP-FUT meter-LNK-3-INSTN MORE high-ADV I-OBL THAN
 'He will jump three meters higher than me'

This suggests that the construction doesn't involve any 'emphasis' (although evaluative/'emphatic' readings of comparatives are possible) and the standard marker and the comparative morpheme together with other components of the construction constitute a familiar comparative that semantically amounts to a statement that a certain individual (subject) exceeds some other individual (standard) on a particular scale (with an option to specify the difference between the two).

The last core fact about the Circassian comparative construction is that the standard marker *nah*-(*r*)jø is optional in Besleney Kabardian, but much less so in Adyghe. The constructions with omitted *nah*-(*r*)jø is judged in Adyghe as colloquial (data from [10]):

- (14) he-r **nah-ə-ṣʷə** [ç'etwə-m ?(nahjə)]
 dog-ABS MORE-ALNK-kind cat-OBL THAN
 'The dog is more kind than the cat'

Why are Circassian comparatives interesting? We believe that it's the fact that the standard marker *nah*(*r*)jø is transparently decomposed into *nah* + particle (*r*)jø. This decomposition is supported, first, by the wide use of (*r*)jø in non-degree constructions – and, second, by the optionality of *r* both in the standard marker and in the other uses of the particle in Besleney Kabardian (although conditioned morphonologically). (*r*)jø has a number of uses in Circassian. Most prominently, it is a scalar additive particle similar to English ‘even’, but it also forms NPIs, free choice items and universal quantifiers in combination with different elements, such as wh-words. One example where a wh-element + (*r*)jø occur in a free relative clause is below:

- (15) **xet-jə** ap-ew sə-z-?ʷəč'e-re-m s-jə-wəč'ə-š't
 who-ADD first-ADV 1SG.ABS-REL-meet-DYN-OBL 1SG.ABS-3SG.ERG-kill-FUT
 'Whoever finds me will kill me' (Genesis 4:14)

We will have more to say on (*r*)jø towards the end of the paper. Let's now formulate our goal.

Under the assumption that *nah* as a comparative morpheme and *nah* as part of the standard marker is one and the same element with one and the same semantics, Circassian seems to not fit easily in the classic analyses of comparatives outlined in the previous section. These classic analyses encode the comparative semantics in the comparative morpheme, while the standard markers are assumed to be semantically empty. Something else is needed to make sense of Circassian data.

The task of a compositional analysis of the Circassian comparative has two subtasks: 1) explaining the appearance of the same element *nah* both as a comparative marker and as part of the standard marker; 2) explaining the role of (*r*)jø in the semantics of standard of comparison.

Before moving on to solving this task, let's put in place the theoretical ingredients needed for our analysis.

2 Towards an analysis: Theoretical ingredients

2.1 Two sources of degree quantification in comparative constructions

Several new insights into the role of standard markers and semantic effects at the edge of standard phrase have been formulated recently.

Unlike in the classic analyses outlined in the Introduction, some authors argue that standard markers are not semantically empty, but can perform degree quantification along with the comparative morpheme. The general idea that the standard marker performs degree quantification has been proposed several times recently for different reasons in different forms for different types of languages [1, 9] We sketch two implementations.

The theory developed in [1] involves slightly different entries for *more* and *than*, but the roles of the two are very similar, as, under this analysis, both encode the ‘exceed’ relation between two degrees (MORE looks more like a version of a 3-place MORE discussed in the introduction, THAN is the same as the classic 2-place MORE):

- (16) a. $\llbracket \text{MORE} \rrbracket = \lambda g_{(d,et)} \lambda s_d \lambda x_e. \mathbf{max}(\lambda d. g(d)(x)) > s$
b. $\llbracket \text{THAN} \rrbracket = \lambda S_{(d,t)} \lambda T_{(d,t)}. \mathbf{max}(T) > \mathbf{max}(S)$

According to the analysis in [1], a simple comparative sentence would actually contain two instances of MORE and one instance of THAN. The authors assume that MORE in the standard clause is required by identity conditions on ellipsis, and the instance of MORE in the standard clause goes unpronounced. Thus the comparative semantics gets introduced three times. Omitting the details of the derivation, we sketch how the resulting familiar comparative semantics is built up under this analysis:

- (17) a. Rod A is longer than Rod B is.
LF: $\llbracket \text{THAN } [s \lambda d'. \text{Rod B } \text{MORE long } d'] \rrbracket [T \lambda d. \text{Rod A } \text{MORE long } d]$
b. $\llbracket T \rrbracket = \lambda d. \llbracket \text{MORE} \rrbracket ([\text{long}](d)([\text{Rod A}])) = \lambda d. \text{long}(\text{Rod A}) > d$
c. $\llbracket S \rrbracket = \lambda d'. \llbracket \text{MORE} \rrbracket ([\text{long}](d')([\text{Rod B}])) = \lambda d'. \text{long}(\text{Rod B}) > d'$
d. $\llbracket \text{THAN} \rrbracket ([\text{S}]) ([\text{T}]) = 1 \text{ iff } \max(\lambda d. \text{long}(\text{Rod A}) > d) > \max(\lambda d'. \text{long}(\text{Rod B}) > d')$

A different ‘two loci of degree quantification’ theory [9] is motivated by comparative constructions in languages like Hebrew, where the comparative morpheme is optional, and when it is absent, the standard marker is enough for the comparative interpretation. Under this analysis, both THAN and MORE are of the same type $\langle \langle d, t \rangle, \langle d, t \rangle, t \rangle$. It is a degree quantifier that states that there is a degree such that it falls within both intervals in combines with:

- (18) $\llbracket \text{MORE} \rrbracket = \llbracket \text{THAN} \rrbracket = \lambda S_{(d,t)} \lambda T_{(d,t)}. \exists d [d \in S \& d \in T]$

For this semantics to work, [9] assumes a silent negation in the standard clause, so the two intervals in (19) are 1) the set of degrees of strength that Yoni does **not** meet, and 2) the set of degrees that Miri does meet:

- (19) a. Miri xazaka mi-Yoni
Miri strong.FEM THAN-Yoni
‘Miri is stronger than Yoni’
b. $\llbracket \text{Miri xazaka mi-Yoni} \rrbracket = \llbracket \text{THAN} \rrbracket (\lambda d. \text{Yoni is NOT } d\text{-strong})(\lambda d'. \text{Miri is } d'\text{-strong})$

When both the standard marker and the comparative morpheme are present, the standard phrase, presumably, acts as a degree quantifier domain adverbial.

The semantics we will develop departs from both analyses presented here, but the core idea remains the same.

2.2 Maximization/exhaustivity in comparative constructions

Independently and based on quite different data – scope of quantifiers in the standard phrase – Beck [2] (building on [6] a.o.) motivates the necessity of a silent ‘maximal informativity’ operator m_inf at the edge of a standard clause. A slightly modified version of m_inf :

- (20) $m_inf(p_{\langle \langle d, t \rangle, t \rangle}) = \lambda D. p(D) \& \neg \exists D' [p(D') \& D \neq D' \& [p(D') \rightarrow p(D)]]$
The maximally informative intervals out of set of intervals p are the set of intervals D s.t. there is no other interval D' in p s.t. $p(D')$ entails $p(D)$ (i.e. if D is in p then so is D').

The resulting set of intervals can be a singleton (in a simple case) or it can contain more than one interval in certain cases. To get from this resulting set of intervals to one interval, [2] defines

an extra $\mathbf{max}_>(p)$ operation that picks the maximum element out of this set relative to the $>$ relation on intervals or degrees. Thus the combination $m_inf + max_>$ at the edge of *than*-clause yield an interval $\langle dt \rangle$, which is a familiar semantic type for standards of comparison.

The motivation for these operations is quite complex and we won't reproduce it here, but they have analogues outside of the degree domain, e.g. at the edge of free relative clauses [7].

3 The analysis of Circassian comparatives

Following the idea about two sources of degree quantification in comparatives, we assume that *nah* as a comparative marker and as part of a standard marker has the same semantics of a degree quantifier. (21) is equivalent to the classic denotation for a 2-place MORE:

$$(21) \quad [\![nah]\!] = \lambda S_{\langle dt \rangle} \lambda T_{\langle dt \rangle} . \mathbf{max}(T) > \mathbf{max}(S)$$

Let's first build the standard phrase semantics using (21). The first argument of *nah* is an interval, so the standard phrase should host degree abstraction for this analysis to work. In Circassian, the standard syntactically is a DP rather than a clause. We treat this as a purely syntactic requirement of the standard marker that doesn't have to have semantic consequences: the standard phrase in Circassian is still a degree interval semantically.

Different ways to achieve this are possible. Here we assume a covert/elided gradable predicate as part of the standard phrase: $\lambda d. \text{that } d\text{-long road}$. Degree abstraction inside a standard phrase is visible in Circassian when the DP contains a relative clause. In this case, one can observe 'degree relativisation' as part of verbal morphology in the relative clause: **ze-re-š'ə-tə-be-m nahjə** (**REL**-manner-locative-stand-past-obl than) 'than how it was'. (22), we propose, is the denotation of the standard phrase before the particle combines with it – a set of intervals:

$$(22) \quad [\![\text{that road } \mathbf{nah}]\!] = \lambda T_{\langle dt \rangle} . \mathbf{max}(T) > \mathbf{max}(\lambda d. \text{that } d\text{-long road}) = \lambda T_{\langle dt \rangle} . \text{LENGTH}(\text{that road}) \in T$$

We propose that $(r)jə$ at the edge of Circassian standard clause combines the semantics of m_inf and $\mathbf{max}_>$ from [2] and thus is equivalent to successive application of these two operators, taking as input a set of intervals and giving one largest maximally informative interval as output:

$$(23) \quad [\![(r)jə]\!] (p_{\langle \langle d, t \rangle \rangle t}) = \mathbf{max}(\lambda D. p(D) \& \neg \exists D' [p(D') \& D \neq D' \& [p(D') \rightarrow p(D)]])$$

After $(r)jə$ combines with (22), the result will be the largest maximally informative interval containing the *length of that road*:

$$(24) \quad \begin{aligned} [\![\text{that road } \mathbf{nah}-(r)jə]\!] &= \mathbf{max}_>-\mathbf{m_inf}(\lambda T_{\langle dt \rangle} . \text{LENGTH}(\text{that road}) \in T) \\ &= \mathbf{max}(\lambda T_{\langle dt \rangle} . \text{LENGTH}(\text{that road}) \in T \& [\neg \exists T'. T' \subset T \& \text{LENGTH}(\text{that road}) \in T']) \end{aligned}$$

(24) has the right type to be an argument of *nah* again and is basically equivalent to the interval $\lambda d. \text{LENGTH}(\text{that road}) \geq d$. It now can be combined with the matrix clause:

$$(25) \quad \begin{aligned} [\![\text{That road } \mathbf{nah}-(r)jə \text{ this road is } \mathbf{nah} \text{ long}]\!] &= [\![\mathbf{nah}]\!] ([\![\text{that road } \mathbf{nah}-(r)jə]\!]) ([\![\text{this road is } d'\text{-long}]\!]) \\ &= [\lambda S_{\langle dt \rangle} \lambda T_{\langle dt \rangle} . \mathbf{max}(T) > \mathbf{max}(S)] (\lambda d. \text{LENGTH}(\text{that road}) \geq d) (\lambda d'. \text{LENGTH}(\text{this road}) \geq d') = \mathbf{max}(\lambda d'. \text{LENGTH}(\text{this road}) \geq d') > \mathbf{max}(\lambda d. \text{LENGTH}(\text{that road}) \geq d) \end{aligned}$$

Finally, we need to analyse comparatives without (overt) *nah*-(r)jə, as in (14). Tests (omitted here) show that a bare standard DP is a direct argument of *nah* on the gradable predicate (=

MORE), which means that we are dealing with a 3-place version of *nah* when *nah-(r)jø* is absent. We assume a type-shifted version of *nah* taking a gradable predicate and two individuals:

- (26) a. $\llbracket nah_3 \rrbracket = \lambda y_e \lambda G_{\langle det \rangle} \lambda x_e. \llbracket nah \rrbracket (\lambda d. G(d)(y))(\lambda d'. G(d')(x))$
b. $\llbracket nah_3 \rrbracket = \lambda y_e \lambda G_{\langle det \rangle} \lambda x_e. \mathbf{max}(\lambda d'. G(d')(x)) > \mathbf{max}(\lambda d. G(d)(y))$

This will derive the right result for sentences with only one instance of *nah*, but it doesn't derive the slightly degraded status of such sentences in Adyghe. We propose that the status of the 3-place *nah* in Besleney Kabardian and in Adyghe is different – in the former, it is a freely available shift or a systematic lexical ambiguity of *nah*, while in the latter it is a marked option.

4 Discussion

We discussed Circassian data that fit some recent proposals concerning the semantics of comparatives – and, at the same time, provide support for such proposals. However, there are issues with the analysis we develop here.

First, to make the analysis fully decompositional, we need to clarify the relation between *(r)jø* in comparatives and its other uses. Its use as ‘even’ is ‘non-truth-conditional’ in the sense that its only contribution is presuppositional, while *(r)jø* in comparatives contributes to the truth-conditional meaning. *(r)jø* in combination with wh-elements (*xet-jø* ‘any-/everybody’) has a different kind of meaning altogether. All these uses are intuitively related, but not identical. We leave this to future work.

Second, a careful look is needed at how Circassian standard DPs get to have type $\langle dt \rangle$. In some cases postulating extra structure inside the standard DP (as we did in (22)) is problematic.

References

- [1] Peter Alrenga, Christopher Kennedy, and Jason Merchant. A new standard of comparison. In Proceedings of WCCFL 30, 2012.
- [2] Sigrid Beck. Quantifiers in ‘than’-clauses. Semantics and Pragmatics, 3:1–72, 2010.
- [3] Rajesh Bhatt and Shoichi Takahashi. When to reduce and when not to: Crosslinguistic variation in phrasal comparatives. In Generative Linguistics in the Old World (GLOW) 28, Newcastle, 2008.
- [4] M.J. Cresswell. The semantics of degree. In Barbara Partee, editor, Montague Grammar, pages 261–292. Academic Press, New York, 1977.
- [5] Irene Heim. Degree operators and scope. In Wolfgang Sternefeld, editor, Audiatur Vox Sapientiae. A Festschrift for Arnim von Stechow, pages 214–239. Akademie-Verlag, Berlin, 2001.
- [6] Irene Heim. Remarks on comparative clauses as generalized quantifiers. Ms., MIT, 2006.
- [7] Pauline Jacobson. On the quantificational force of free relatives. In Emmon Bach, Eloise Jelinek, Angelika Kratzer, and Barbara Partee, editors, Quantification in Natural Languages. Kluwer Academic Publishers, Dordrecht, 1995.
- [8] M.A. Kumakhov, K. Vamling, and Z. Kumakhova. Ergative case in the Circassian languages. Lund University Department of Linguistics Working Papers, 45:93–111, 1996.
- [9] Roger Schwarzschild. Comparative markers and standard markers. In Michael Y. Erlewine and Yasutada Sudo, editors, Proceedings of the MIT Workshop on Comparatives (to appear), 2011.
- [10] M.V. Vill. Sravnitel'naya stepen v temirgoevskom dialekte adygeyskogo yazyka (Comparison in Temirgoy dialect of Adyghe). Fieldwork report, Moscow, 2013.
- [11] Arnim von Stechow. Comparing semantic theories of comparison. Journal of Semantics, (3):1–77, 1984.

Man and Woman: the Last Obstacle for Boolean Coordination

Lucas Champollion*

NYU Department of Linguistics
10 Washington Place
New York, NY 10003
United States of America
champollion@nyu.edu

Abstract

The word *and* can be used both intersectively and collectively. A major theme in research on coordination has been the quest for a lexical entry that unifies these uses, either based on boolean intersection or based on collective formation. Focusing on English noun-noun coordination, this paper argues for the boolean option. This immediately delivers the intersective behavior of *and*, as in *liar and cheat*; as for its collective behavior, as in *man and woman*, it falls out of the interaction of *and* with a series of independently motivated type shifters, mainly taken from Winter (2001). Such coordinations are interpreted collectively because the two nouns are interpreted in the same way as the DPs in *a man and a woman*.

1 Introduction

The word *and* can be used both intersectively, as in the sentences in (1), and collectively, as in the sentences in (2). This paper focuses on noun-noun conjunction in English, which also shows both intersective and collective behavior. For example, sentence (1b) is about a person in the intersection of the sets denoted by the predicates *liar* and *cheat*, while sentence (2b) is about a collective entity formed by a man and a woman.

- | | | |
|-----|--|----------------------------------|
| (1) | a. John lies and cheats.
b. That liar and cheat can not be trusted. | (intersective)
(intersective) |
| (2) | a. John and Mary met in the park last night.
b. A man and woman met in the park last night. | (collective)
(collective) |

A major theme in research on coordination has been the quest for a lexical entry that unifies such uses, either based on boolean intersection, as for example by Gazdar (1980), or based on “non-boolean” set/sum formation, as for example by Heycock and Zamparelli (2005). Many authors also assume that *and* is lexically ambiguous between the two uses (e.g., Link, 1984).

The purpose of this paper is to argue for the boolean option, that is, for the idea that *and* invariably denotes generalized intersection. This immediately delivers the intersective behavior of *and*, as in (1). For example, the coordination in (1b) is predicate intersection:

$$(3) \quad [\text{liar and cheat}] = \lambda x. \text{liar}(x) \wedge \text{cheat}(x)$$

As for the collective behavior of coordination, as in (2), I will show that it emerges as a consequence of the interaction of *and* with a series of independently motivated type shifters. Coordinations like *man and woman* are interpreted collectively because the two nouns are interpreted in the same way as the DPs in *a man and a woman*.

*I thank Chris Barker, Dylan Bumford, Simon Charlow, Anna Szabolcsi, and an anonymous reviewer.

2 Analysis: Boolean *And* plus Type Shifters

I assume that *and* always has the meaning in (4), suggested among others by Gazdar (1980). Following Winter (2001), I will refer to this as the “boolean assumption”. Roughly, Gazdar’s entry says that a conjunction of sentences S_1 and S_2 is true whenever both of the conjuncts are true, and a conjunction of subsentential constituents C_1 and C_2 denotes their intersection.

$$(4) \quad [\text{and}_{\text{bool}}] = \sqcap_{(\tau, \tau\tau)} =_{\text{def}} \begin{cases} \wedge_{(t, tt)} & \text{if } \tau = t \\ \lambda X_\tau \lambda Y_\tau \lambda Z_{\sigma_1}. X(Z) \sqcap_{(\sigma_2, \sigma_2\sigma_2)} Y(Z) & \text{if } \tau = \sigma_1\sigma_2 \end{cases}$$

I build on Winter (2001), who does not consider noun-noun conjunction but who shows that the boolean option is viable in related cases, such as the collectivity effect in the VP of (2a). Winter’s account relies on the insight that one can use the minimizer in (5) to “distill” any intersection or union of principal ultrafilters into a set of sets of their generators.

$$(5) \quad \mathbf{GQ \ minimizer:} \min =_{\text{def}} \lambda Q_{\tau t} \lambda A_\tau. A \in Q \wedge \forall B \in Q [B \subseteq A \rightarrow B = A]$$

For example, the intersection of the generalized quantifiers that are obtained by Montague-lifting the two constants corresponding to *John* and *Mary* is $\{P \mid j \in P\} \cap \{P \mid m \in P\}$, the set of all properties that hold both of John and of Mary. The result of applying the minimizer in (5) to this set is $\{\{j, m\}\}$. Following Winter, I represent plural individuals as sets, so this is the property of being the plural individual consisting of John and Mary. This is one of the two properties involved in sentence (2a); the other is denoted by the VP and is a property of sets who met in the park last night. The meaning of (2a) can then be obtained by combining these two properties via the silent existential-closure type-shifter E defined in (6), whose meaning is the same as the meaning of the determiner *a*: it states that the intersection of the two properties it combines with is not empty. As Winter discusses in detail, one can also think of E as a generalization of independently-needed choice-functional operators. The fact that choice-functional operators are generally taken to apply to nouns makes it a natural assumption to apply E to nouns as well, as I will do below.

$$(6) \quad \mathbf{Existential-closure type-shifter:} E =_{\text{def}} \lambda P_{\tau t} \lambda Q_{\tau t}. P \cap Q \neq \emptyset$$

Given these assumptions, Winter analyses the subject of sentence (2a) as in (7), a property which is true of any set that contains the plural individual consisting of John and Mary. This gives the right truth conditions once it combines with the VP, as shown here:

$$(7) \quad [E(\min(\lambda P.P(\text{john}) \cap \lambda P.P(\text{mary})))] = \lambda C_{(\text{et}, t)}. \{j, m\} \in C$$

Turning now to noun-noun coordination, which Winter (2001) does not discuss, I assume that the coordination *man and woman* in (2b) involves the two silent operators just presented, but in a different order. Specifically, I assume that E may apply to nominal predicates without affecting their syntactic category. So it can apply to a nominal like *man* and return another nominal, which I assume it does here (on both sides of the conjunction). The standard assumption is that coordination does not affect syntactic categories, so the result of coordinating E(*man*) with E(*woman*) is again a nominal. At this stage, the denotation of *man and woman* is the same as the denotation of the noun phrase *a man and a woman*, although their syntactic categories differ. The denotation results from intersecting the generalized quantifier $\lambda P. \exists x. \text{man}(x) \wedge P(x)$ with the generalized quantifier $\lambda P. \exists x. \text{woman}(x) \wedge P(x)$. The result is the following:

$$(8) \quad [E(\text{man}) \text{ and } E(\text{woman})] = \lambda P_{et} \exists x \exists y. \text{man}(x) \wedge \text{woman}(y) \wedge P(x) \wedge P(y)$$

With Winter, I assume that a nominal predicate of type $\langle et, t \rangle$ must first be “distilled” before it can be used further, in this case as the restrictor of the (overt) determiner a . As in the previous case, this is achieved by the minimization operator. Conceptually, here as before, the input to this operator is a generalized quantifier over ordinary individuals, and its output is a predicate over collective individuals. In this case, assuming (as I will throughout the paper for convenience) that the set of men and the set of women are disjoint, the output is the predicate that holds of any man-woman pair:

$$(9) \quad [\min(E(\text{man}) \cap E(\text{woman}))] = \lambda P_{et} \exists x \exists y. \text{man}(x) \wedge \text{woman}(y) \wedge P = \{x, y\}$$

From here on, I will abbreviate this collective predicate as **mw-pair**.

As mentioned, I model collective predicates as set predicates of type $\langle et, t \rangle$. Ordinary determiners expect their restrictor and their nuclear scope to be of type $\langle et \rangle$. To combine with $\langle et \rangle$ -type predicates, I assume following Winter that determiners are adjusted via the determiner fitter $dfit$, defined as follows:

$$(10) \quad \textbf{Determiner fitter: } dfit =_{def} \lambda D_{\langle et, \langle et, t \rangle \rangle} \lambda A_{\langle et, t \rangle} \lambda B_{\langle et, t \rangle}. D(\bigcup A)(\bigcup(A \cap B))$$

Winter motivates this operator by sentences like (11), in which the collective predicate *met* is an argument of a quantificational determiner.

$$(11) \quad \text{No students met.}$$

To see how determiner fitting works, note first that the plural morpheme on *students* is modeled by Winter by a “predicate distributivity” (*pdist*) operator, whose function is similar to the well-known * and D operators in the literature on plurals (e.g. Link (1998)) but which also prepares ordinary $\langle et \rangle$ -type predicates for determiners that have been adjusted for collective predicates via determiner fitting. The *pdist* operator is defined as follows:

$$(12) \quad \textbf{Predicate distributivity: } pdist =_{def} \lambda P_{et} \lambda P'_{et}. P' \neq \emptyset \wedge P' \subseteq P$$

Using the operators (10) and (12), Winter analyzes sentence (11) in terms of the meanings of singular *no* and *student*. Its meaning is predicted to be “No student is a member of a set of students that met”.

$$\begin{aligned} (13) \quad & [\text{dfit}(\text{no})(\text{pdist}(\text{student}))(\text{met})] \\ &= [\text{no}](\bigcup \text{pdist}([\text{student}])(\bigcup (\text{pdist}([\text{student}]) \cap [\text{met}])) \\ &= [\text{no}]([\text{student}])(\bigcup \{P \in [\text{met}] : P \subseteq [\text{student}]\}) \\ &= \neg \exists x. [\text{student}(x) \wedge \exists P. x \in P \wedge P \in \text{meet} \wedge \forall y. y \in P \rightarrow \text{student}(y)] \end{aligned}$$

Given this, my LF for sentence (2b) is shown in (14). It is true iff its VP, the collective predicate *met in the park*, holds of at least one man-woman pair.

$$\begin{aligned} (14) \quad & [\text{dfit}(a)(\min(E(\text{man}) \text{ and } E(\text{woman}))) (\text{meet_in_the_park})] \\ &= \exists (\bigcup \text{mw-pair})(\bigcup (\text{mw-pair} \cap \text{meet_in_the_park})) \\ &= \exists x. \exists y. \text{man}(x) \wedge \text{woman}(y) \wedge \text{meet_in_the_park}(\{x, y\}) \end{aligned}$$

The semantics of *dfit* ensures that this sentence requires the man and the woman in question to have been part of the same meeting, as opposed to having each met separate people. If we had not used *dfit*, we would have predicted weaker truth conditions. The sentence would already be true if a man met some other men in the park, and a woman met some other women in the park in a separate meeting. So the assumption is crucial that determiners expect their arguments to

be of type $\langle et \rangle$, and that they adjust via $dfit$ when their arguments are of type $\langle et, t \rangle$.

Of course, noun-noun coordination does not require the VP to be collective. A sentence like (15a), with a distributive predicate in the VP, is represented as in (15b). Here, $pdist$ and $dfit$ make sure that the property of smiling is distributed over the two elements of any man-woman pair that makes the sentence true.

- (15) a. A man and woman smiled.
- b. $\llbracket dfit(a)(\min(E(\text{man}) \text{ and } E(\text{woman}))(pdist(\text{smile})) \rrbracket$
 $= \exists x. \exists y. \text{man}(x) \wedge \text{woman}(y) \wedge \{x, y\} \subseteq \text{smile}$

3 Comparison to Previous Work

Like any system that adopts a uniform meaning for *and*, this one avoids redundancy, which improves on Link (1984). Since the meaning I adopt is boolean, it generalizes to S, VP, and DP coordinations without problems (Gazdar, 1980). This improves on Heycock and Zamparelli (2005), one of the few journal-length treatments of the semantics of noun-noun coordination. Noun-noun coordination is discussed in Winter (1995) and Winter (1998) though not in Winter (2001). The present system is vastly different from the treatment of noun-noun coordination in Winter (1998). I now discuss Heycock and Zamparelli (2005) and Winter (1998) in more detail.

3.1 Heycock and Zamparelli (2005)

Heycock and Zamparelli (2005) adopt a non-boolean entry for *and* that is equivalent to the one in (16). Essentially, this entry combines two sets of sets by computing their cross-product, except that instead of putting any two elements together to form a pair, it forms their union. Heycock and Zamparelli (2005) call this operation *set product* in reminiscence of the notion of cross-product.

$$(16) \quad \llbracket \text{and}_{coll} \rrbracket = \lambda Q_{\langle \tau t, t \rangle} \lambda Q'_{\langle \tau t, t \rangle} \lambda P_{\tau t} \exists A_{\tau t} \exists B_{\tau t}. A \in Q \wedge B \in Q' \wedge P = A \cup B$$

Heycock and Zamparelli (2005) assume that nouns and VPs denote sets of singletons. For example, the noun *man* denotes the set of all singletons of men, $\lambda P. |P| = 1 \wedge P \subseteq \text{man}$. When the nouns *man* and *woman* are conjoined, the entry in (16) generates the following denotation:

$$(17) \quad \begin{aligned} & \llbracket \text{man and}_{coll} \text{ woman} \rrbracket \\ &= \lambda P_{et} \exists A_{et} \exists B_{et}. |A| = 1 \wedge A \subseteq \text{man} \wedge |B| = 1 \wedge B \subseteq \text{woman} \wedge P = A \cup B \\ &= \lambda P_{et} \exists x \exists y. \text{man}(x) \wedge \text{woman}(y) \wedge P = \{x, y\} \end{aligned}$$

This denotation is equivalent to the one my system generates, as seen in (9). In this respect, my system can be seen as a reconstruction of the one in Heycock and Zamparelli (2005) from first principles. But there is an important difference. I assume that all instances of *and* are boolean while Heycock and Zamparelli assume that all instances of *and* have the non-boolean denotation in (16). The latter assumption leads to problems when quantifiers are conjoined that are not upward entailing, as in the following cases:

- (18) a. No man and no woman smiled.
- b. Mary and nobody else smiled.

Assume first, as Heycock and Zamparelli do, that the simplex DPs are treated as generalized quantifiers, as shown in (19) for *no man* (the unusual types are due to the assumption that

nouns denote sets of singletons):

$$(19) \quad [\text{no man}] = \lambda Q_{(et,t)} \cdot \neg \exists X_{(et)} \cdot [\text{man}](X) \wedge Q(X)$$

Heycock and Zamparelli predict that the complex DP in (18a) holds of the union of any set **A** containing no man and any set **B** containing no woman. As **A** may contain women and **B** may contain men, the resulting truth conditions are too weak. For example, (18a) is true in a model that contains a smiling man called John, a smiling woman called Mary, and no other smilers. This is for the following reason. The entry for *no man* in (19) holds of the set containing nothing but the singleton of Mary, since that set contains no man; the corresponding entry for *no woman* holds of the set containing nothing but the singleton of John since that set contains no woman. According to entry (16), the DP in (18a) therefore holds of the union of these two sets, namely, the set containing nothing but the singletons of John and of Mary. But this set is precisely the denotation of *smiled* in this model. For analogous reasons, (18b) is predicted to be true in this model (assuming that *nobody else* in this context means *nobody other than Mary*).

Heycock and Zamparelli are aware of this problem and suggest that scope-splitting analyses of *nobody*, as proposed by Ladusaw (1992) and others for languages with negative concord, might help here. On these analyses, the lexical entry of *no* is separated into one part that contains only \neg and another part that contains everything else including $\exists x$, and the negation part is free to take scope in a higher position than the rest. But adopting such an approach would wrongly predict that (18b) means the same as *It's not the case that Mary and someone else smiled*. That sentence, unlike (18b), is true when Mary didn't smile but someone other than Mary smiled. And of course, standard English does not have negative concord, so adopting a split-scope analysis of *no* is not an available option for standard English.

3.2 Winter (1995, 1998)

In contrast to Winter (2001) discussed above, earlier work including Winter (1995) and Winter (1998, ch. 8) discusses noun-noun coordination, which is taken to require a departure from the boolean assumption. In that work, *and* always returns the denotations of its two conjuncts as an ordered pair. For example, *man and woman* is translated as the ordered pair in (20).

$$(20) \quad [\text{man and woman}] = \langle \lambda x. \text{man}(x), \lambda x. \text{woman}(x) \rangle$$

When such a pair combines with other items in the tree, it is first propagated upwards in a style reminiscent of alternative semantics (e.g., Rooth, 1985), in the sense that each of the two computations proceeds in parallel with the other. At any point in the derivation, this ordered pair can be collapsed back into a single denotation by application of \sqcap as defined in (4). When this operation happens immediately, it mimics the behavior of boolean *and*; the reason for introducing it is to give *and* the possibility to take arbitrarily wide scope. As Winter (1998) demonstrates, this leads to the right results in cases like (21), which is ambiguous between readings (21a) and (21b).

- (21) Every linguist and philosopher knows the Gödel Theorem.
- a. Everyone who is both a linguist and a philosopher knows the Gödel Theorem.
 - b. Every linguist knows the Gödel Theorem, and every philosopher knows the Gödel Theorem.

In Winter (1998)'s analysis of (21), if \sqcap is introduced immediately, this leads to the reading in (21a); if it is introduced after the conjuncts have combined with the determiner and optionally with the VP, the reading in (21b) is generated. On the present account, reading (21a) is obtained by intersection; reading (21b) is obtained by insertion of the type shifters E , min , and $dfit$ as demonstrated in the discussion of *man and woman* above.

However, the delayed introduction of intersection in Winter (1998) overgenerates. For example, the system does not prevent *No girl sang and danced* from meaning *No girl sang and no girl danced*. To see this, consider the following derivation:

- (22) a. $\llbracket \text{sang and danced} \rrbracket = \langle \lambda x. \text{sing}(x), \lambda x. \text{dance}(x) \rangle$
b. $\llbracket \text{no girl} \rrbracket = \lambda P. \neg \exists x[\text{girl}(x) \wedge P(x)]$
c. $(22b)((22a)) = \langle \neg \exists x[\text{girl}(x) \wedge \text{sing}(x)], \neg \exists x[\text{girl}(x) \wedge \text{dance}(x)] \rangle$
d. Application of \sqcap : $\neg \exists x[\text{girl}(x) \wedge \text{sing}(x)] \wedge \neg \exists x[\text{girl}(x) \wedge \text{dance}(x)]$

The problem here is similar to the one facing early accounts of VP coordination in Transformational Grammar via conjunction reduction. By allowing the subject to enter the computation twice and by giving *and* scope over it, such accounts overgenerate in many cases where the subject is a quantifier. The present system avoids this problem since *and* is interpreted as local, not delayed, intersection. A sentence like *No girl sang and danced* is interpreted simply by intersecting *sang* and *danced* locally.

To be sure, intersecting *sang* and *danced* locally is also a possible derivation in Winter (1998). The present account must be prevented from overgenerating by blocking the application of type shifters like E to verbs, like *sang* and *danced*. For this reason, I assume that the distribution of type shifters and other silent operators is not free but is constrained by syntax, just like the distribution of ordinary words. This assumption is discussed and defended at length in Winter (2001). Of course, one could adopt the system of Winter (1998) by constraining the application of \sqcap syntactically as well, for example by requiring pairs to be collapsed at certain nodes including VP. However, one might then as well adopt the present system and avoid the departure from the boolean hypothesis.

4 *And* vs. *Or*

Most authors who adopt the boolean analysis of coordination assume that it applies in equal ways to *and* and *or*. I will assume the same here. That is, I adopt the following entry for *or* based on Gazdar (1980), analogous to the boolean entry for *and* shown in (4):

$$(23) \quad \llbracket \text{or}_{\text{bool}} \rrbracket = \sqcup_{\langle \tau, \tau \tau \rangle} =_{\text{def}} \begin{cases} \vee_{\langle t, tt \rangle} & \text{if } \tau = t \\ \lambda X_\tau \lambda Y_\tau \lambda Z_{\sigma_1}. X(Z) \sqcup_{\langle \sigma_2, \sigma_2 \sigma_2 \rangle} Y(Z) & \text{if } \tau = \sigma_1 \sigma_2 \end{cases}$$

Bergmann (1982) challenges the boolean analysis based on examples that involve noun-noun coordination. The puzzle Bergmann raises is the following: Why are the sentences in (24a) equivalent while those in (24b) are not? The purpose of this section is to provide a solution for Bergmann's puzzle.

- (24) a. Every cat and dog is licensed. \Leftrightarrow Every cat or dog is licensed.
b. A cat and dog came running in. $\not\Leftrightarrow$ A cat or dog came running in.

For the sentences in (24a), the present system generates (among others) two equivalent LFs, shown in (25) and (26) along with their translations. For convenience, I treat the VPs *came*

running in and *be licensed* as unanalyzed predicates. They are distributive predicates, or atom predicates in the sense of Winter (2001), which means that they do not by themselves trigger determiner fitting. The application of *dfit* in (25a) is triggered by the type of the collective predicate *cat and dog*, which is treated in the same way as *man and woman* above. In a slight departure from Winter (1998), who assumes that applying *dfit* changes the pronunciation of *every* to “all”, I assume that the pronunciation of *every* is not affected when the conjoined DPs are singular.

- (25) a. $\text{dfit}(\text{every})(\min(\text{E(cat)} \text{ and } \text{E(dog)}))(\text{pdist}(\text{be_licensed}))$
 b. $\bigcup\{\{x, y\} | \text{cat}(x) \wedge \text{dog}(y)\} \subseteq \bigcup\{\{x, y\} | \text{cat}(x) \wedge \text{dog}(y) \wedge \{x, y\} \subseteq \text{be_licensed}\}$
- (26) a. $\text{every}(\text{cat or dog})(\text{be_licensed})$
 b. $\text{cat} \cup \text{dog} \subseteq \text{be_licensed}$

The translations in (25b) and (26b) are equivalent, as the reader may verify. As for the sentences in (24b), there is no way to generate equivalent LFs for them. For example, the LFs in (27a) and (28a) correspond to the most prominent (if not the only) readings of the two sentences in (24b), and they evaluate to the nonequivalent formulae in (27b) and (28b).

- (27) a. $\text{dfit}(\text{a})(\min(\text{E(cat)} \text{ and } \text{E(dog)}))(\text{pdist}(\text{come_running_in}))$
 b. $\exists x \exists y. \text{cat}(x) \wedge \text{dog}(y) \wedge \{x, y\} \subseteq \text{come_running_in}$
- (28) a. $\text{a}(\text{cat or dog})(\text{come_running_in})$
 b. $\exists x. (\text{cat}(x) \vee \text{dog}(x)) \wedge \text{come_running_in}(x)$

5 The Non-Ambiguity of *Or*

Unlike *and*, which is descriptively ambiguous between intersective and “non-boolean” uses, *or* has no such seeming ambiguity in any known language (Payne, 1985). As Winter (2001) emphasizes, this provides strong motivation against accounts that attribute collective uses of *and* to this word being ambiguous between a boolean and a non-boolean entry, since such accounts provide no explanation of the fact that *or* is not ambiguous in the same way. The type-shifting account of Winter (2001) provides a general answer to this question. Interestingly, this answer also extends to the present system. As discussed above, I have assumed that a surface string of the shape *N1 and N2* can correspond to the two LFs “*N1* \sqcap *N2*” and “ $\min(\text{E(N1)} \sqcap \text{E(N2)})$ ”. These two structures have completely different meanings. This explains why *and* sometimes looks like intersection and sometimes like collective formation. As for noun-noun disjunction, however, the situation is different. I assume that the same structures are generated: “*N1* \sqcup *N2*” and “ $\min(\text{E(N1)} \sqcup \text{E(N2)})$ ”. You might expect that this incorrectly predicts that *or* is ambiguous in an analogous way to *and*. But these two structures evaluate to almost the same thing, and because of determiner fitting, the remaining difference between them disappears in the course of the rest of the derivation. While “*N1* \sqcup *N2*” underlies the derivation in (26), “ $\min(\text{E(N1)} \sqcup \text{E(N2)})$ ” underlies the following derivation, which is equivalent to (26).

- (29) $\text{dfit}(\text{every})(\min(\text{E(cat)} \sqcup \text{min}(\text{E(dog)})))(\text{pdist}(\text{be_licensed}))$
 = $\text{dfit}(\text{every})(\min(\{P | P \cap \text{cat} \neq \emptyset \vee P \cap \text{dog} \neq \emptyset\})(\{P | P \neq \emptyset \wedge P \subseteq \text{be_licensed}\}))$
 = $\bigcup\{\{x\} | x \in (\text{cat} \cup \text{dog})\} \subseteq \bigcup(\{\{x\} | x \in (\text{cat} \cup \text{dog})\} \cap \{P | P \neq \emptyset \wedge P \subseteq \text{be_licensed}\})$
 = $(\text{cat} \cup \text{dog}) \subseteq ((\text{cat} \cup \text{dog}) \cap \text{be_licensed})$
 = $(\text{cat} \cup \text{dog}) \subseteq \text{be_licensed}$

6 Summary and Outlook

The boolean option is arguably the only unproblematic one outside the DP; I have shown that it is also preferable within the DP. The intersective behavior is expected. The collective behavior comes for free in the framework of Winter (2001), where we already have in place the right set of operators to generate collectivity effects. The system sketched here can be easily extended to handle plural nouns, which raise counting-related issues (e.g. *Ten men and women got married today*), and as my anonymous reviewer points out, also to “hydras” such as *every man and woman who met at the concert* (Link, 1984). As I will discuss elsewhere, accounting for related patterns across languages, discussed by Heycock and Zamparelli (2005) and references therein, leads to interesting questions about the semantics and scope of plural and agreement morphemes.

References

- Bergmann, M. (1982). Cross-categorial semantics for conjoined common nouns. *Linguistics and Philosophy*, 5:299–401.
- Gazdar, G. (1980). A cross-categorial semantics for coordination. *Linguistics and Philosophy*, 3:407–409.
- Heycock, C. and Zamparelli, R. (2005). Friends and colleagues: Plurality, coordination, and the structure of DP. *Natural Language Semantics*, 13(3):201–270.
- Ladusaw, W. (1992). Expressing negation. In *Proceedings of SALT*, volume 2, pages 237–259.
- Link, G. (1984). Hydras. On the logic of relative clause constructions with multiple heads. In Landman, F. and Veltman, F., editors, *Varieties of formal semantics*, GRASS 3. Foris, Dordrecht, Netherlands.
- Link, G. (1998). *Algebraic semantics in language and philosophy*. CSLI Publications, Stanford, CA.
- Payne, J. R. (1985). Complex phrases and complex sentences. In Shopen, T., editor, *Language typology and syntactic description, vol. II: Complex constructions*, pages 3–41. Cambridge University Press, Cambridge, UK, 1 edition.
- Rooth, M. (1985). *Association with focus*. PhD thesis, University of Massachusetts, Amherst, MA.
- Winter, Y. (1995). Syncategorematic conjunction and structured meanings. In Simons, M. and Galloway, T., editors, *Proceedings of SALT*, volume 5, pages 387–404.
- Winter, Y. (1998). *Flexible Boolean semantics: Coordination, plurality and scope in natural language*. PhD thesis, Utrecht University.
- Winter, Y. (2001). *Flexibility principles in Boolean semantics*. MIT Press, Cambridge, MA.

On the meaning of Intensifiers

Harris Constantinou*

University College London, London, U.K.
h.constantinou@ucl.ac.uk

Abstract

The intensifier, or emphatic reflexive, is a stressed anaphorically dependent element, exemplified by *himself* in *John himself built this house*. Depending on its distribution, the intensifier may take up to three radically different interpretations: adnominal, exclusive and inclusive. I claim that the common denominator of the three readings is the centrality effect imposed on the antecedent. The analysis departs from Eckardt (2001) and suggests that while the adnominal interpretation denotes the identity function ID operating on the domain of individuals D_e (i.e. $ID(x)$), the inclusive and exclusive respectively constitute clausal and manner adverbial variations of $ID(x)$, that is, $ID(e, x)$.

1 Introduction

In many languages the intensifier (INT for short) may be found adjoined to its antecedent or at some distance from it, adjoined to some clausal projection. This is illustrated for English in (1) – (3).

- 1) (It wasn't the director's secretary who went to the meeting.) The director *herself* went. (adnominal)
- 2) (Apart from Bill,) John has *himself* built a house, even though he wasn't happy about it. (inclusive)
- 3) (John did not build this house with Bill's help). John built it *himself*. (exclusive)

Interpreatively, the three distributional variants differ. When INT is adjoined to the antecedent, as in (1), it can be loosely paraphrased with 'in person'. When INT is found immediately after the auxiliary, as in (2), it has a meaning similar to additive focus particles (e.g. also). When INT is found post-verbally, as in (3), it may imply that the action denoted by the predicate is carried out 'without help'. In a different context, a post-verbal occurrence of INT may also take the additive reading. I adopt previous conventions (e.g. Siemund 2000, Gast 2006) in referring to the reading in (1) as the adnominal, the reading in (2) as the inclusive, and the reading in (3) as the exclusive.

Despite their different interpretation the three variants deserve to be considered related linguistic elements; they are morphologically the same (in many languages), they form a syntactic dependency with an (nominal) antecedent, they are adjuncts, and they are invariably stressed. This last feature is indicative of some kind of Information-Structural (IS) marking on INT that consistently induces alternatives. In all three cases, the alternatives generated contain an individual (e.g. the secretary in (1)) that is understood to be peripheral to the antecedent (e.g. the director in (1)). In opposite terms, the antecedent of INT is understood to be *central* against the alternative referent.

The notion of centrality has been characterized within the literature in terms of high status, prominence or importance of a referent (e.g. Siemund 2000, i.a., see Gast 2006 for criticism). In this paper, I adopt the following operational definition, based on König's (1991):

* I am grateful to Hans Van de Koot for his invaluable help in developing the ideas contained in this paper. This work is supported by the AHRC (UK) and the Leventis foundation.

- 4) An entity x is central against an entity y , if x ranks higher than y on some salient scale specified by the context.

The fact that centrality of x presupposes the existence of a value-scale proves particularly important for our purposes because it creates the expectation that in every case in which x is the antecedent of INT, we should be able to say with precision against which scale x qualifies as central against y .

The structure of the paper is as follows. In section 2 I suggest that each INT centralizes its antecedent in a different manner. In section 3 I propose a new semantics for each INT to account for their interpretive characteristics. I then conclude by briefly discussing how the Elsewhere Principle (Kiparsky 1973) regulates the meaning of INTs.

2 Interpretation

2.1 Adnominal Intensifier

The meaning contribution of the adnominal INT can be described as follows: a) it evokes a set of alternative referents $\{y, z, \dots\}$ to its antecedent x ; b) it structures this set into a central element x and peripheral elements $\{y, z, \dots\}$ (Eckardt 2001 i.a.). (a) becomes evident from the infelicity of (1) when inserted in an out of the blue context; alternative referents must be readily available in prior discourse. Most proposals, including the present one, assume that this is due to the IS marking of INT. In relation to (b), an alternative referent y may be understood to be peripheral to INT's antecedent x via world knowledge, as in (4).

- 4) *Context: Peter is the director's secretary.*

It wasn't Peter who went to the meeting. The director *herself* went.

In this context, Peter is understood to be peripheral to the director on the basis of general world-knowledge about companies (i.e. directors are in a hierarchically higher position than their employees). We can make sure that the scale of comparison is indeed company hierarchy by removing the contextual assumption that Peter is the director's secretary in (4). Such a scenario makes the use of the adnominal INT impossible. Importantly, we can also confirm that the adnominal INT requires a central referent to interact with by keeping everything constant but exchanging the positions of Peter and the director; the result is infelicity because Peter is not a central referent in discourse.[†]

Example (4) implies the exclusion of the alternative peripheral entity. However this is not always the case. The exclusion of the director's secretary in (4) is a result of contrastive focus on INT in that example. Below is a case of INT IS-marked as a contrastive topic. Notice that the centrality effect is still there but not the exclusion of alternatives.

- 5) A: Tell me about the director's secretary. Did he have a good time in Paris?

B: Well, I don't know about the director's secretary, but the director [*herself*]CT certainly did.

[†] The alternative referent y may also be understood as peripheral to the antecedent x , if y is identified via x . For instance if y is referred to as x 's *brother*. I do not discuss these cases here but see Constantinou (forthcoming) for details.

2.2 Exclusive Intensifier

Example (3) is an instance of the exclusive INT.[‡]

3) (John did not build this house with Bill's help). John built it *himself*.

The most obvious contribution of the exclusive INT in (3) is to negate an alternative version of the same event, in which the agent, John, receives Bill's help to build the house. Furthermore, the antecedent is understood to be somehow involved in the negated alternative description. Recall that this was not a requirement for the adnominal INT. Removing John from the picture in the alternative consistently results in an infelicitous exclusive INT.

6) (Mary did not build this house with Bill's help). #John built it *himself*.

More specifically, the role of the antecedent in the alternative must be that of the agent, and the one of the excluded referent that of the helper (a role that can be realized in a comitative PP). Reversing the two referents' roles, while keeping the antecedent constant, results in an infelicitous use of INT:

7) (Bill did not build this house with John's help). #John built it *himself*.

These data could be taken as supporting evidence for the centrality of the antecedent of the exclusive INT. The question though is *how* the antecedent is central. Considering that the antecedent must be involved in some way in the negated version of the event, the scale/criterion must be some sort of an event internal (or, loosely speaking, thematic) relation on the basis of which the antecedent is compared with, and ranks higher than, the alternative referent (see Ernst 2002 for details on event internal relations). We can hypothesize the following:

8) The exclusive INT centralizes its antecedent against other referents in an *event-internal* manner.

(8) presupposes a pretty specific state of affairs. In the first instance, compared referents must be holding the same event internal relation for them to be situated on the same scale. To elaborate on this, let's take a closer look at (3). The negated alternative description of (3) can presumably take the following linguistic realization: *John built this house with Bill*. Arguably, John and Bill, the compared referents, hold a similar role, that of a causer[§]. The two causers differ in one respect; the external referent is held accountable for the building of the house whereas the comitative referent is not (see Neeleman & Van de Koot 2012 for an explication of accountability). This indicates that the comparison is done on the basis of a scale of accountability for causing the building of the house. The external referent is situated higher than the comitative on such scale; thus the former is central against the latter. It is worth noting that the literature on intensification has never been concerned as to why a context like *John finished the race next to Bill* does not allow the subsequent use of the exclusive INT, but a context like *John finished the race with the help of Bill* does. This contrast follows naturally from (8) and, more specifically, from the requirement that alternative referents are compared on the basis of the same thematic relation. While in the latter both John and Bill are causers (of John finishing the race), in the former John is the agent while Bill is a locating point of reference.

[‡] Aside from the context in (3), the exclusive INT may occur in a context in which the antecedent is understood to benefit from the event at issue (e.g. Bill did not build this house for John, but John built it *himself*) or understood to be negatively affected by it (e.g. Bill did not ruin John's career, but John ruined it *himself*). See Constantinou (forthcoming) for details.

[§] I use 'causer' quite loosely in this paper. A referent *x* is a 'causer' if *x* is involved in a bringing about a result.

Additionally, (8) presupposes that a particular thematic relation can be held by a referent at different degrees. The accountability relation is one of them as is evident from the negated alternative of (3), in which both the external and comitative referents are causes of the event but the former is higher on the accountability scale than the latter. There are situations in which two referents hold the same thematic relation, as in (9), but the use of the exclusive INT is impossible.

- 9) A: John went to the cinema with Bill. (i.e. John is accompanied by Bill when going to the cinema)
B: # No, John went to the cinema himself.

(9) differs from (3) in terms of predicate type; (9) denotes an activity. In virtue of this, and given that both John and Bill participate in the event in the same manner, either of them may be classified as an agent. The infelicity of INT in (9B) lies in the fact that the two referents are found at the same point of the agent scale. John and Bill go to the cinema to the same extent. To put this more crudely, it is impossible for a referent to half-go somewhere. This implies that there cannot be central/peripheral referents in such context (with the intended reading in brackets).

Another important facet of the exclusive INT is the fact that all alternative propositions are understood as alternative descriptions of the same event and, as a result of this, they are all negated. This follows naturally from (8). A proposition with the exclusive INT may look as follows: *John built it himself*. The exclusive INT requires John to be central in an event internal manner, let's say, in terms of accountability. Suppose that the alternative, *John built the house with Bill*, denotes a different event. Then John would be event-internally central in this latter event but, crucially, *not* in the former simply because they are different events. Hence, the alternatives must denote the same event.

2.3 Inclusive Intensifier

Example (2) contains an inclusive INT and can be paraphrased with an additive focus particle (e.g. also) without a serious change in meaning.

- 2) (Apart from Bill,) John has *himself* built a house, even though he wasn't happy about it.

For the additive effect to be possible, it is a prerequisite that salient alternatives must be understood to be true. This is indeed how salient alternatives are interpreted in the inclusive case. In fact, previous authors (e.g. Gast 2006) suggest that the additive effect is delivered precisely for this reason. Thus, a theory of intensification ought to explain why all alternatives are understood to be true in order to derive the additive effect.

Aside from the additive effect, providing a principled explanation for the truth of all alternatives may also lead to an understanding of the subject orientation of the inclusive INT^{**}. In order to see why, consider again the (only type of) discourse that makes the inclusive INT possible.

- 10) A: Bill has raised three kids.
B: John has *himself* raised three kids, but he said that it was hard.

The key observation is that INT's antecedent is always interpreted as a switch topic. In (10), speaker A makes an utterance *about* Bill and speaker B responds by uttering something *about* some other

^{**} The adnominal and exclusive INTs may interact with any argument (for data from Dutch see Constantinou forthcoming).

topic, John. Given the tendency of topics to also be subjects in English, the subject-orientation of the inclusive INT comes as no surprise (see Gast 2006 for similar conclusions)^{††}.

A further characteristic of the inclusive INT, first observed by Siemund (2000), is the fact that the proposition p_1 that contains it is consistently interpreted as the premise/reason/explanation for some other salient proposition p_2 in discourse, corresponding to *he said that it was hard* in (10B). Once p_2 is removed from (10B), the inclusive INT in p_1 becomes impossible. Even more curiously, p_2 must be expressing something about INT's antecedent. For instance, the inclusive INT is impossible in (10B) if p_2 expresses something about the speaker (e.g. but I think it is hard to raise three kids), even though p_1 may still be interpreted as a premise. I take this to result from the existence of some specific relation between the antecedent and the event denoted in p_1 . This relation cannot be of the thematic/event internal type, as in the exclusive INT case, because the manner the event in p_1 (e.g. raising kids) takes place is not influenced by p_2 . On this basis, I advance the following hypothesis:

- 11) The inclusive INT centralizes its antecedent against other referents in an *event-external* manner.

In (10), the event external relation held between John and the event of raising three kids concerns the perceived difficulty of the event (i.e. how difficult John thinks it is to raise three kids). Of course, this relation is event external: the degree of perceived difficulty does not affect the event at issue. If (11) is correct, then, we expect the two alternative referents to be compared on the basis of the scale of the sort 'how hard a referent x thinks it is to raise three kids'; Since John is a good antecedent to the inclusive, it must be the case that John is understood to believe that it is harder to raise three kids compared to Bill; i.e. John ranks higher than Bill on the 'perceived difficulty in raising three kids' scale. The degradation of the following example, which explicitly expresses that John and Bill situate in the same point of this scale, supports (11):

- 12) A: Bill has raised three kids and he said that it was a hard thing to do.
B: ? John has *himself* raised three kids and he found it equally hard.

3 Deriving the meaning of Intensifiers

In this section, I propose a semantic analysis for the three INTs that is inspired by Eckardt (2001). She suggests that the core-meaning contribution of the adnominal intensifier is the identity function ID on the domain of objects D_e .

- 13) ID: $D_e \rightarrow D_e$

$$\text{ID}(x) = x \text{ for all } x \in D_e$$

According to this analysis, the adnominal intensifier is lexically specified with ID, which takes as its input value a nominal constituent x , the antecedent, and maps it onto the same output value. (14) exemplifies this operation for the DP *John himself*.

- 14) $\llbracket [\text{John}] \text{ himself} \rrbracket = \text{ID}(\llbracket \text{John} \rrbracket) = \llbracket \text{John} \rrbracket$

The assumption that the adnominal INT denotes ID(x) equals to saying that its core meaning contribution to the sentence amounts to nil. For INT to give rise to a meaningful interpretive effect, it

^{††} Of course, this analysis raises another question, that is, why would the inclusive INT prefer a topical antecedent to begin with. I am not aware of an answer to this, and I cannot offer one either. But see Constantinou (forthcoming) for directions for a solution on the basis of Kirfka's (2000) analysis of German post-posed additive particles, which are also topic-oriented.

must evoke a meaningful alternative. This accounts for the fact that INT is always IS marked (see Eckardt 2001 for a related proposal). I therefore make the further assumption that the IS marking of the adnominal INT induces the family of peripherality functions $\text{PER}_i(x)$. Each $\text{PER}_i(x)$ maps the nominal constituent x to an entity peripheral to x . (15) illustrates the semantic characteristics of this function. (16) exemplifies the operation of $\text{PER}(x)$.

$$15) \text{PER}_i: D_e \rightarrow D_e$$

$$\text{PER}_i(x) = x \text{ for all } x \in D_e$$

$$16) [[\text{John}] \text{PER}_i] = \text{PER}_i([\text{John}]) = [\text{an entity peripheral to John}]$$

The combination of (13) – (16) delivers the centrality effects on the adnominal INT's antecedent. It is crucial to note that for this to work, ID and PER_i must operate on the same referent x . This can be achieved by assuming that x is always given information, and thus present in all alternatives.

The intention is to use the semantics of the adnominal INT, and its alternative, as the basis for deriving the different types of centrality effect of the two adverbial cases. With minor changes, the analysis runs parallel to Ernst's (2002) analysis of predicational adverbs, such as the agent-oriented adverb *rudely*. Ernst suggests that *rudely* is a two-place predicate taking as arguments the agent and an event (of some type) and indicates that an event is such as to judge its agent as rude with respect to the event. *Rudely* may take two readings, a clausal and a manner, as exemplified below.

$$17) \text{Rudely, Sue left. (clausal)}$$

$$18) \text{Sue left rudely. (manner)}$$

The clausal reading can be paraphrased as ‘Sue was rude to leave’ (i.e. Sue is characterized as rude but the nature of her leaving need not be rude) whereas the manner reading reads as ‘the nature of Sue’s leaving was rude’ (e.g. she banged the door on her way out). Notice how the event is utilized in each case. In the clausal case, the event is merely used as the basis for characterizing the agent as rude, whereas in the manner case the properties of the event of leaving are characterized. In light of these differences, clausal *rudely* is an event external modifier whereas manner *rudely* is an event internal modifier. Ernst suggests that these two readings may be distinguished in a more principled fashion on the basis of event comparison classes. In (17), Sue is judged rude because of her leaving, as opposed to other things she could have done, which crucially need not be other leaving events. In (18) Sue is judged rude on the basis of something about her leaving, which distinguishes this from other leaving events, and only leaving events. On Ernst's theory, the different event comparison classes for (17) and (18) result from the selection of a different type of event (i.e. Fact-Event Object, FEO for short); the manner interpretation results from selecting a Specified Event, hence the comparison class is more specific, and the clausal interpretation results from selecting a clausal FEO (i.e. Event or Proposition or Fact). One of the attractive elements of Ernst's approach is that there is no need to assume two lexical entries for *rudely*. Since the duality of interpretation of *rudely* is particularly productive in natural language, Ernst assumes the existence of a ‘manner rule’ that takes clausal *rudely* and turns it into a manner adverb. (19) contains the lexical specification of *rudely*, (20) the semantic representation of clausal *rudely* and (21) that of manner *rudely* (the representations in (20)-(21) contain minor changes for reasons discussed in Constantinou forthcoming).

$$19) \text{Rudeness } (e, \text{Agt})$$

$$20) \text{Clausal ‘rudely’} = e [\text{REL warrants posing}] \text{rudeness in Agent as opposed to the norm for events.}$$

$$21) \text{Manner ‘rudely’} = e [\text{REL manifests}] \text{rudeness in Agent as opposed to the norm for Specified events.}$$

(20) and (21) differ in terms of the relation imposed between the two arguments. In the clausal case ‘warrant positing’ intends to capture the fact that the event as a whole forms the basis for or guarantees the attribution of rudeness to the agent. In the manner case ‘manifest’ captures the fact that manner adverbs describe some sort of external manifestation (e.g. Sue’s banging the door in (18)).

Following the above, I suggest that the inclusive and exclusive INT lexically differ from the adnominal in carrying an extra *e* variable in their entry. The alternative of the two adverbial INTs lexically differs from the adnominal’s in the same way. I represent these as follows:

- 22) ID (*e*, *x*)
- 23) PER (*e*, *x*)

The inclusive INT and its alternative constitute the clausal versions of (22) and (23), whereas the exclusive INT and its alternative constitute the manner versions of (22) and (23). These may be represented as follows:

- 24) Inclusive = *e* [REL warrants positing] ID in *x* as opposed to the norm for events.
- 25) Inclusive Alternative = *e* [REL warrants positing] PER in *x* as opposed to the norm for events.
- 26) Exclusive = *e* [REL manifests] ID in *x* as opposed to the norm for Specified events.
- 27) Exclusive alternative = *e* [REL manifests] PER in *x* as opposed to the norm for Specified events.

The combination of (24) and (25) delivers the interpretive effect of the inclusive, and the combination of (26) and (27) that of the exclusive. Starting from the exclusive, (27) expresses that there is a specified event that realizes a peripheral structure to the antecedent *x*. *x* denotes the same referent in ID (*e*, *x*) and PER (*e*, *x*). The events modified by ID (*e*, *x*) and PER (*e*, *x*) are spatiotemporally the same. Thus, if *x* is central in the event modified by PER (*e*, *x*), *x* is also central in the event modified by ID (*e*, *x*). Note that the centrality of *x* is motivated by the properties of the event itself, which implies that the centrality must be on the basis of an event related criterion (and not, let’s say, social hierarchy). It is also crucial to emphasize that the event’s FEO classification is one of a specified event. In the general case this implies event-internal modification (and a closed event comparison class, with the same predicate for all events). For the exclusive INT’s antecedent *x*, this implies that *x*’s centrality is such that it is evident *in* the event itself. That is, the event has certain properties that make *x* central. As elaborated in 2.2, centrality is indeed calculated on the basis of such event-internal properties, such as accountability of a referent for causing an event’s resultant state.

A similar rationale applies in the inclusive case. (25) expresses that there is an event *e* on the basis of which a peripheral structure to *x* is constructed. Again, *x* denotes the same referent in all alternatives. As opposed to the exclusive (and its alternative) case, the FEO classification of the event is clausal (and not specified event), something which implies event external modification (and an open event comparison class, with potentially different predicate for events). The immediate consequence of this is that in an event modified by (25) *x* is rendered central with respect to an event external criterion, and not some property of the event itself (as in the exclusive case). It is important to understand that the event is merely utilized as the basis, or as some sort of evidence, for the characterization of *x* as central. This means that the centrality is calculated on the basis of an event-related criterion, like in the exclusive case; however, the event participates in the calculation of centrality differently in each adverbial case. Going back to (25), its presence imposes centrality on *x* on the basis of the event it modifies. As discussed in 2.3 with reference to (10), the antecedent John may be judged central on the basis of the event-external criterion of him finding the event of raising three kids more difficult than Bill, who is the output of PER applied to *x*. Crucially, the ‘raising kids’ event that is utilized for characterizing John central against Bill is denoted by the alternative proposition (*p*₂) (i.e. (10A)), and not the event denoted by the proposition with the inclusive (*p*₁) (i.e. (10B)). The two events are of the same type though. Both of them talk about ‘raising three kids’. If

John is characterized as central on the basis of p_2 , then it must be the case that John is central on the basis of p_1 too. Thus, no issue arises with the inclusive INT's requirement to interact with a central antecedent. This perspective may also explain why the inclusive INT requires all the alternatives to be interpreted as true (and thus spatiotemporally different events). In (10B) the inclusive INT requires its antecedent John to be central against Bill with regard to a certain scale, that is, the 'perceived difficulty in raising three kids'. For John and Bill to be compared against this scale, it must be true that both of them have raised three kids. Now, suppose a scenario in which (10A) constitutes a false alternative description of (10B). This would mean that Bill has *not* raised three kids. It follows from this that Bill cannot fall on the scale of 'perceived difficulty in raising three kids'. Consequently, John cannot compare with Bill and deemed central as required. We may conclude that the use of the inclusive INT requires true alternatives because only in this way its antecedent may be central. As indicated in 2.3, the additive effect and the subject-orientation of this INT naturally follow from this.

4 Final remarks

As it stands, the overall account is unable to make the correct predictions. The adnominal INT's semantics is general enough to suggest that it should be able to associate with an antecedent that is qualified for any type of centrality, including the event internal/external types. What is needed, then, is a way to prevent the adnominal INT from associating with an adverbial interpretation. In order to achieve this, we could consider it to be the general, unmarked, case in competition with the two specific, marked, adverbial cases. Such a move is supported by the fact that the adnominal INT is universal, but not the adverbial variants. We could then invoke the Elsewhere Principle (Kiparsky 1973) to regulate the interpretation of the three INTs. If a language has one or both of the more highly specified adverbial cases, then the adnominal INT will not be used to centralize its antecedent in an event internal or external manner (depending on the INTs present in the language), because there are specific linguistic elements for this purpose. This would capture the inability of the adnominal INT to associate with an interpretation licensed by its adverbial counterparts in English. It would also capture the ability of the adnominal INT to associate with an adverbial interpretation in languages (e.g. Greek, Italian) in which not all three INTs are grammaticalized (for data see Constantinou forthcoming).

References

- Constantinou, H. (forthcoming). *Intensifiers: Meaning and Distribution*. PhD thesis:UCL, London.
 Eckardt, R. 2001. Reanalysing *selbst*. *Natural Language Semantics*, 9(4) , 371-412.
 Ernst, T. (2002). *The Syntax of Adjuncts*. Cambridge: Cambridge University Press.
 Gast, V. 2006. *The Grammar of Identity. Intensifiers and Reflexives in Germanic languages*. New York: Routledge.
 Kiparsky, Paul. 1973. 'Elsewhere' in Phonology. In S. Anderson and P. Kiparsky (eds.), *A Festshrift for Morris Halle*. New York: Holt, Rinehart and Winston.
 König, E. (1991). *The Meaning of Focus Particles: A Comparative Perspective*. London: Routledge.
 Krifka, M. 1999. Additive particles under stress. In: D. Strolovitch and A. Lawson (eds.). *Proceedings of SALT 8*. Ithaca NY: CLC Publications, 111–128.
 Neleman, A. & Van de Koot, H. (2012). The Linguistic Expression of Causation. In Everaert, M., Siloni, T., Marej, M. (Eds.). *The Theta System: Argument Structure at the Interface*. Oxford: Oxford University Press.
 Siemund, P. (2000). *Intensifiers in English and German. A comparison*. New York: Routledge.

Most: the View from Mass Quantification*

Carmen Dobrovie-Sorin

LLF-CNRS, Université Paris 7
carmen.sorin@linguist.univ-paris-diderot.fr

Abstract

It is currently assumed that *most* can quantify over mass domains [12][14]. The crosslinguistic data (English, Romanian, Hungarian and German) examined in the paper indicates that this is true of entity (type e)-restrictor *most* but not of set (type <e,t>)-restrictor *most*, which strongly supports the view that mass quantifiers denote relations between entities rather than relations between sets [26] [21] [14]. Nevertheless my proposal differs from previous proposals in assuming a more constrained view of the syntax-semantics mapping.

1 Introduction: The puzzle and its solution in a nutshell

Examples of the type in (1) are currently invoked in favor of the view that quantification over mass domains is allowed [12] [14], but examples like (2) go against this generalization. (2a-c) are from Matthewson ([22]:174); (2d) is my example:

- (1) a. Most milk from old goats is sour.
 b. Most water is liquid.
- (2) a. *Most milk in this fridge is sour.
 b. *I shoveled most snow that was in this yard.
 c. *Most mud that you traipsed in the house ended up on my rug.
 d. *Most furniture in this house is broken.

The problem was left open by Matthewson and to my knowledge it has only rarely been brought up in the subsequent literature (Szabolcsi [28] observes it, but leaves it aside). My solution relies on the distinction between entity-restrictor *most* and set-restrictor *most*:

- (3) a. Entity-restrictor *most* takes an entity-denoting (type e) restrictor.
 b. Set-restrictor *most* takes a set-denoting (type <e,t>) restrictor.

Given this distinction, the contrast between (2) and (4) can be captured by the generalization in (5):

- (4) a. Most of the milk in this fridge is sour.
 b. I shoveled most of the snow that was in this yard.
 c. Most of the mud that you traipsed in the house ended up on my rug.
 d. Most of the furniture in this house is broken.
- (5) Entity-restrictor *most* can apply to mass domains but set-restrictor *most* cannot do so.

Going back to (1a-b), they are problematic if we take the *most* that occurs there as having a set-denoting restrictor. The problem can be solved by assuming (6):

- (6) When occurring as complements of *most*, bare NPs *can* be kind-referring (type e) in English.
- (6) is weaker than Matthewson's [22] hypothesis, according to which bare NP complements of *most* are *necessarily* kind-referring. My weaker hypothesis correlates with the assumption that a set-restrictor *most* (see § 2.1 and § 2.3 below) exists (although it turns out to be illegitimate for mass domains), in addition to the entity-restrictor *most*. According to Matthewson herself, *most* always takes an e-type restrictor¹.

*This research is supported by the LABEX-PLU project. Thanks to D. Büring, P. Cabredo Hofherr, L. Champollion, I. Giurgea, D. Lassiter, F. Landman, B. H. Partee and A. Szabolcsi for comments on a previous draft of the ideas presented here. I am also indebted to my mathematician husband, Sylvain Sorin, who - after listening to me for 10 minutes - solved in 2 minutes the problem of 'measuring kinds' that constantly came up in my previous attempts of dealing with mass quantification. Sylvain just said : 'Well, you just take the whole as 1.'

¹See also [1] for the view that all Quantifiers always take DPs as complements.

2 MOST across languages

2.1 Romanian and Hungarian: MOST vs THE LARGEST PART OF

The examples in (7), (8) and (9) show that in English, Romanian and Hungarian the superlative forms of MANY, although quite different from each other in terms of morphosyntactic complexity, all allow the proportional interpretation characteristic of the quantificational use of *most* in English (this reading should be carefully distinguished from the relative/comparative reading, which in English is signalled by the presence of *the*, e.g., *Who read the most books?* In Romanian and Hungarian the definite article also appears in the proportional use):

- (7) Most students in my class left early.
- (8) Cei mai multi elevi din clasa mea au plecat devreme².
 the more many students in my class have left early.
 'Most students in my class left early.'
- (9) A legtöbb fiú már hazament³⁴.
 the most boy already is-gone-home.
 'Most boys are gone home already.'

Turning now to the mass domain, the constraint in (5) correctly predicts the unacceptability shown in (10), which is parallel to the unacceptability of (2a-d) in English:

- (10) a. *Cel mai mult lapte din frigiderul ăsta e acru.
 the more much milk in fridge-the this is sour.
- b. *Cea mai multă mobilă din această casă e veche.
 the_{fem} more much_{fem} furniture in this_{fem} house is old.

Although she does not provide explicit examples, Szabolcsi [28] [27] makes it clear that the proportional reading of a *legtöbb* is disallowed with NP_{mass}⁵.

The intended meanings of (10a-b) can be rendered by using partitive configurations of the type *the largest part of DP* (which are the closest counterparts of the English *most of DP*)⁶:

- (11) Cea mai mare parte din untul din acest frigider e acru.
 the more big part of butter-the in this fridge is sour.
 'The largest part of/most of the butter in this fridge is sour'

Examples of this type (their Hungarian counterparts are also acceptable) obey the constraint in (5) - on a par with the English examples in (4) - because the complement of *the largest part* is a DP, which is an e-type expression.

In sum, the following generalizations hold in English, Romanian and Hungarian:

²Romanian superlatives are built like in the other Romance languages: the (strong form of the) definite article + *mai* 'more' + Adj, e.g., *cel_{msg} mai bun unt* 'the_{msg} more good butter' 'the best butter', *cele_{fempl} mai bune eleve* 'the_{fempl} more good_{fempl} students_{fempl}' 'the best women students'.

³The fact that *sok* 'much', *több* 'more' and a *legtöbb* 'the most' are not marked as plural correlates with the fact that their NP complements, just like the complements of cardinals, are not plural marked in Hungarian, e.g., *három fiú* 'three boy' meaning 'three boys'.

⁴Anna Szabolcsi (p.c., November 10, 2012) observes that the example in (9) - which she kindly provided - 'is kind of okay, if I am saying this more in virtue of what I deduce from the fact that it's midnight now, than in virtue of having checked on the individual boys.' This comment is in line with Szabolcsi [28] [27], according to whom a *leggtöbb* is acceptable with count NPs, but only 'in generic contexts'. Whatever the fine-grained description of the data might be, the crucial point for our present purposes is that a *legtöbb* is allowed with NP_{count}.

⁵In both Romanian and Hungarian, relative superlative readings of *THE MOST NP_{mass}* are allowed, e.g., Romanian *Cine a băut cel mai mult vin?* and Hungarian *Ki ivott a legtöbb bort?* 'Who drank the most wine?'.

⁶In Romanian (as well as in Hungarian) the distinction between entity restrictors and set restrictors perfectly correlates with the distinction between partitive and non-partitive constituents. In line with Matthewson [22] I assume that partitive prepositions or Case markers are morphosyntactically motivated but do not bear on the semantics, i.e., they are expletive (*contra* Ladusaw [17], followed by others, who assumes that partitive prepositions apply to e-type expressions and yield <e,t> type). Anticipating on the discussion in § 2.3, it should be pointed out that the correlation between e-type denotation and morphosyntactic partitive marking breaks down in English, because (i) bare NPs can be either set-denoting (type <e,t>) or kind-denoting (type e) and (ii) regardless of their denotational type, bare NPs are not marked with a partitive preposition.

- (12) a. MUCH/MANY_{superl} NP_{count} VP allows the proportional reading.
 b. MUCH_{superl} NP_{mass} VP disallows the proportional reading.

2.2 The German puzzle

The German example in (13)⁷ seems to go against (12b):

- (13) Maria hat den meiste Kaffee weggewischt.
 Maria has the most coffee wiped-away
 (aber es ist noch ein bisschen in den Ecken geblieben).
 Mary wiped away most of the coffee (but there is still a bit in some corners)'

This problem can be solved by adopting Roehrs' ([25]:107) hypothesis that in German the definite article raises to D° from a low position labelled (Art(icle), which means that *der meiste NP* 'the most NP' is base-generated as */meiste [ArtP der NP]/* '[most [ArtP the NP]]'. Assuming that *der* is interpreted in the low position, example (13) turns out to obey the generalization in (5), since – at the relevant level of representation – *meiste* does not take NP (type <e,t>), but rather ArtP (type e) as a complement.

2.3 MOST in generic contexts

Romanian examples of the type in (14) contrast with their English counterparts in (15):

- (14) a. *Cea mai multă apă e lichidă.
 the more much water is liquid
 b. *Cel mai mult lapte de capre bătrâne e acru.
 the more much milk of goats old is sour.
- (15) a. Most water is liquid.
 b. Most milk from old goats is sour.

This contrast is parallel to the contrast in (16)-(17), which shows that bare NPs in argument positions can be kind-referring in English but not in Romanian [8], [9], [11], [7]:

- (16) a. *Apă e lichidă.
 Water is liquid
 b. *Lapte de capre bătrâne e acru.
 Milk of goats old is sour
- (17) a. Water is liquid.
 b. Milk from old goats is sour.

The two sets of contrasts can be correlated by assuming that kind-referring bare NPs are allowed in the complement position of MOST only if kind-referring bare NPs are allowed in argument positions. Kind-referring bare NPs [2] can be analyzed as DPs headed by a null Det with the semantics of Chierchia's [3] Down operator:

- (18) *[QP[Qcea mai multă] [NPapă]] e lichidă. (= (14a))
 the more much water is liquid
- (19) [QP[QMost] [DP[DetØ] [NPwater]]] is liquid. (= (15a))

⁷The judgments indicated in this section, due to Patricia Cabredo-Hoffherr and Barbara Hemforth, confirm the generalization reported in Szabolcsi [28] where the judgments are attributed to Lucas Champollion.

The contrast between (14) and (15) can now be explained on the basis of the generalization in (5). The English examples are acceptable because the restriction/complement position of MOST is filled with kind-referring (type e) mass bare NPs, i.e., DPs headed by a covert Down operator (see (19)). The Romanian examples are ruled out because in this language bare NPs cannot be kind-referring (as indicated by the unacceptability of (16)); the only option is shown in (18), where [apă] is a genuine NP, a configuration that is ruled out by the ban against set-denoting restrictors⁸.

Going back to the English examples in (2), repeated under (20), they are unacceptable because stage-level modifiers prevent bare NPs from referring to kinds (recall Carlson's observations regarding *parts of this machine*); *milk in this fridge* can only be analyzed as a set-denoting bare NP, which is ruled out by (5):

- (20) a. *[_{QP}[_QMost] [_{NP}milk in this fridge]] is sour.
- b. *I shoveled most snow that was in this yard.
- c. *Most mud that you traipsed in the house ended up on my rug.
- d. *Most furniture in this house is broken.

Note that it is not modification *per se* that blocks kind-reference. Those modifiers that are compatible with kind-reference are allowed, in particular (certain) i-level predicates, e.g., *black* or *from old goats* in *black cats* or *milk from old goats*:

- (21) a. Most black cats are intelligent.
- b. Milk from old goats is sour.

Let us finally consider the examples in (22) from Matthewson [22], who attributes them to V. Dayal (p.c.):

- (22) a. Most men who came to the party left early.
- b. Most people at yesterday's rally were Democrats.
- c. Most voters surveyed indicated that...
- d. [context: Last night I threw a party and a bunch of linguists and philosophers got drunk]
 Most linguists who got drunk merely passed out, but most philosophers who got plastered revealed interesting things about their colleagues.

The contrast between (20) and (22) cannot be explained by Matthewson's [22] assumption that the NP complement of *most* is always kind-referring⁹. I take this contrast to support (12a-b), showing that a set-denoting (type <e,t>) restrictor is allowed with count quantifiers but not with mass quantifiers (see § 3 for the explanation). The contrast in (23) provides further evidence in favor of the same generalization:

- (23) a. I am sure most men will arrive late.
- b. *I know most wine will be delivered late.

In this pair of examples the kind-reference of the restrictor is blocked by the s-level nature of the main predicate. We are again left with the only other possibility, a set-denoting restrictor, which is allowed with NP_{count} but disallowed with NP_{mass}.

In sum, the contrast between MOST_{count} and MOST_{mass} can be observed in English only in non-generic contexts, more precisely in those contexts in which the kind-reference of the restrictor is blocked. In Romanian and Hungarian, on the other hand, the contrast can be observed even in generic contexts, because in these languages the NP-complements of MOST are always set-denoting. Which is allowed for count quantification, but not for mass quantification.

3 Towards an explanation

The fact that *most* is the superlative of *much/many* has lead Hackl [13] to propose that proportional *most* is not a quantificational Determiner (*contra* the Generalized Quantifier Theory (GQT) analysis) but rather an adjective (both syntactically and semantically). This analysis incorrectly predicts that (THE)

⁸The counterparts of (14) and (16) are also unacceptable in Hungarian.

⁹But see Crnic [4], who relies on heavy type-shifting combined with stipulated constraints.

MOST NP_{mass} can take a proportional reading¹⁰. The systematic contrast between the acceptability of the relative reading and the unacceptability of the proportional reading of (THE) MOST NP_{mass} can be explained if we assume (24a,b):

- (24) a. On its relative reading, THE MOST is a superlative adjective.
b. On its proportional reading, (THE) MOST is a quantificational Det.

3.1 Set Quantification

According to the set-theoretical analysis of *most* [23] [24] examples of the type in (25) are true iff (26) is satisfied; \cap notates the general lattice-theoretic operation 'meet' (intersection is meet applied to sets):

- (25) Most students in my class have left early.

$$(26) |\{x: \text{student}(x)\} \cap \{\text{left-early}(x)\}| > |\{x: \text{student}(x)\} \cap \{\text{not-left-early}(x)\}|$$

In words, (26) says that the set of students in my class for which the property denoted by the VP (*leave early*) is true has a greater cardinality than the set for which the VP-property is false.

Let us now turn to the unacceptable example in (27), which illustrates the ban on applying the set-restrictor MOST to mass domains (capital X notates a variable that ranges over non atomic entities):

- (27) *Most milk in the fridge is sour.

The condition in (28) - written on the model of (26) *modulo* cardinality being replaced by a measure function notated μ - is illegitimate (as indicated by #) because mass NPs denote non-atomic join semi-lattices [20] [18] and meet, which is required in (28), is not defined¹¹¹² on such a poor algebraic structure.

$$(28) \# \mu (\{X: \text{butter}(X)\} \cap \{\text{sour}(X)\}) > \mu (\{X: \text{butter}(X)\} \cap \{\text{not-sour}(X)\})$$

The unacceptability of (27) may thus be attributed to the fact that a condition of the type in (28) can be computed on count domains but not on mass domains. It is worth pointing out that the contrast between count and mass quantification is not related to the impossibility of counting. Cardinality is a particular measure function, and for the purposes of proportional quantification an unspecified measure function is sufficient (more on this below).

3.2 Entity-restrictor MOST

Turning now to the entity-restrictor MOST, I will assume that it denotes a relation between two entities [26] [21] [14], which are respectively supplied by the restrictor and by the maximal sum obtained by applying the \sum operator to the predicate in the nuclear scope (\sum is a nominalizing operator¹³ that applies to a lattice structure and picks up the maximal element in that lattice)¹⁴. Given this analysis, (29) is true iff (30a) or (30b) is satisfied (which one of these is more appropriate is irrelevant here):

- (29) Most of the milk in the fridge is sour.

$$(30) \text{ a. } \mu([[\text{the milk in the fridge}]] \cap \sum X. \text{sour}(X)) > 1/2 \mu [[\text{the milk in the fridge}]]$$

¹⁰This prediction is explicitly mentioned in Hackl [13] (footnote 9): 'Although I believe that the analysis of *most* given below can be straightforwardly extended to *much-est*, I will ignore proportional mass quantifiers here.' Hackl's [13] proposal is confronted with two further problems: (i) proportional *most* lacks *the* in English, whereas the relative superlative *most* requires *the*, e.g., *Who read the most books?* [15] [28] [6]; (ii) it cannot explain the fact that *le plus de NP* in French cannot take a proportional reading [6].

¹¹This explanation is inspired by Szabolcsi & Zwarts [29], according to whom weak island violations are explained by the impossibility of performing the operations required by the weak island trigger on the domain of the extracted element. The general point is that a syntactically well-formed sentence is semantically unacceptable if computing it requires performing an operation on a structure for which the operation is not defined.

¹²The ban on quantifying over elements ordered by the part-of relation, which is currently invoked for quantification over situations (e.g., Kratzer [16]:169), can also be invoked [10] [7], but it merely restates the observation.

¹³Chierchia's 1998 Down operator strongly resembles Higginbotham's \sum operator

¹⁴According to Higginbotham [14], the analysis in (30) explains the Homogeneity Constraint on mass quantification (e.g., **Most of this milk is heavy*) observed by Lønning [21]: the conditions in (30a-b) rely on the Sigma operator, which can apply to homogeneous predicates (*liquid*, *red*, etc.), because such predicates denote join semi-lattices, but not to non-homogeneous predicates (*heavy*, *cost 200 euros*, *intelligent*, etc.), because such predicates denote sets that are not ordered by the part-of relation, which do not have a supremum.

- b. $\mu([[\text{the milk in the fridge}]] \cap \sum X. \text{sour}(X)) > \mu([[\text{the milk in the fridge}]] - [[\text{the milk in the fridge}]] \cap \sum X. \text{sour}(X))$

The computations required in (30a-b) are legitimate because in these formulae meet applies to two mass entities (type e): $[[\text{the milk in the fridge}]]$, i.e., the overall amount of milk in the fridge on the one hand and $\sum X. \text{sour}(X)$, i.e., the maximal amount of sour stuff in the world.

As to the measure function μ on which (30a-b) rely, a particular choice of a measure unit (for the same dimension) does not affect truth conditions. We can even disregard canonical measure units (meters, liters, etc.) and use ratios in order to measure the parts of an object with respect to the object itself (this is possible because size is a ratio scale [19]): the measure of the whole (in this case $[[\text{the milk in this fridge}]]$) is 1 and the measure of any part of the whole is a ratio r comprised between 0 and 1. In order to check whether (27) is true we need to calculate $r = \frac{\text{vol}([[the-milk-in-this-fridge]] \cap \sum X. \text{sour}(X))}{\text{vol}([[the-milk-in-this-fridge]])}$; (27) is true iff $r > \frac{1}{2}$.

This type of computation is particularly useful for indeterminate/infinite entities such as kinds, which cannot be measured by using regular measure units. But since all we need is the measure of parts of kinds with respect to the kind itself, we can use ratios by taking the measure of the kind itself to be 1. Thus, (31) is true iff (32) is satisfied:

- (31) Most milk from old goats is sour.

$$(32) r = \frac{\text{vol}([[milk-from-old-goats]] \cap \sum X. \text{sour}(X))}{\text{vol}([[milk-from-old-goats]])} \text{ is bigger than } \frac{1}{2}$$

4 Constraining the syntax-semantics mapping

My proposal confirms the view that mass quantifiers denote relations between entities rather than relations between sets. It is however worth stressing that the details of the analysis differ from previous implementations regarding the syntax-semantics mapping. Following Link [20], I have assumed that mass NPs denote non-atomic join semi-lattices, a type of denotation that is incompatible with set quantification (because meet is not defined on join semi-lattices). I have shown that examples of the relevant type, i.e., examples built with a mass NP in the complement position of MOST are unacceptable.

Such ungrammaticalities are expected neither on Lønning's [21] nor on Higginbotham's [14] assumptions regarding mass NPs. According to Lønning, mass NPs are peculiar in that they denote entities (type e) rather than sets of entities (type $\langle e, t \rangle$), as count NPs do. This assumption predicts the crosslinguistic counterparts of MOST NP_{mass} to be acceptable, which is disconfirmed by the Romanian and Hungarian data examined in section 2. Inside English itself, examples of the type in (2), e.g., **Most milk in this fridge is sour*, are also incorrectly predicted to be acceptable: since according to Lønning [_{NP}milk in this fridge] denotes an entity (the overall amount of milk in this fridge), *most* should be able to denote the relation between this entity and the overall amount of sour stuff in the world.

According to Higginbotham [14], the sister of *most* is a mass NP (type $\langle e, t \rangle$) that gets nominalized due to a default application of the Sigma operator:

- (33) Most water is liquid.

- (34) a. $[[NP_{\text{water}}]]$: the property of being water, i.e., the set of quantities of water (type $\langle e, t \rangle$)
b. $\sum [[NP_{\text{water}}]]$: the maximal sum of water (type e)

Although technically different from Lønning's assumptions regarding the syntax-semantics of mass NPs, Higginbotham's proposal is itself confronted with the problem of the crosslinguistic data in §2 and with the unacceptability of (2) in English: the default application of the Sigma operator in the restriction of MOST incorrectly predicts that mass quantification should always be acceptable.

In order to account for the observed unacceptabilities, type-shifting needs to be severely constrained:

- (35) No type-shifting (no default application of \sum) in the restriction of quantifiers.

Given (35), the examples in (2) are unacceptable because mass NPs denote join semi-lattices and meet is not defined on join semi-lattices. Examples like (31) or (33) are on the other hand acceptable because they are built with kind-referring (type e) bare NPs and meet *can* apply to entities.

In sum, my analysis follows Lønning [21] and Higginbotham [14] in assuming that mass quantifiers denote relations between entities, but crucially differs from these authors regarding the syntax-semantics analysis of mass NPs and mass DPs. My proposal has the advantage of relying on the 'null hypothesis' regarding mass NPs: they are $\langle e, t \rangle$ type expressions that can be shifted to e-type denotation by applying a (overt or covert) Det¹⁵.

What may seem unconstrained is the assumption of two types of MOST. I believe that the evidence is too compelling to avoid making this assumption. Besides the benefits invoked in the present paper, this hypothesis may shed new light on the analysis of other quantifiers, e.g., *all* and adverbs of quantification, including the GEN operator¹⁶.

References

- [1] A. Cardinaletti and G. Giusti. The syntax of quantified phrases and quantitative clitics. In Martin Everaert and Henk van Riemsdijk, editors, *The Blackwell Companion to Syntax*, volume V. Blackwell Publishing, 2006.
- [2] GregN. Carlson. A unified analysis of the english bare plural. *Linguistics and Philosophy*, 1(3):413–457, 1977.
- [3] Gennaro Chierchia. Reference to kinds across language. *Natural Language Semantics*, 6(4):339–405, 1998.
- [4] Luka Crnic. On the (non-) cumulativity of cumulative quantifiers. In *Proceedings of Sinn und Bedeutung*, volume 14, pages 117–133, 2009.
- [5] Carmen Dobrovie-Sorin. On the genericity of mass (in)definites. In CUNY, editor, *LSRL abstract*, New-York, 2013.
- [6] Carmen Dobrovie-Sorin. Strong indefinites are not quantificational: proportional many vs most. *Revue roumaine de linguistique*, 2013.
- [7] Carmen Dobrovie-Sorin and Claire Beyssade. *Redefining indefinites*, volume 85. Springer, 2012.
- [8] Carmen Dobrovie-Sorin and Brenda Laca. Generic bare nps. *Unpublished manuscript*, 1996.
- [9] Carmen Dobrovie-Sorin and Brenda Laca. La généricté entre la référence à l'espace et la quantification générique. In P. Sauzet, editor, *Langues et Grammaire*, volume 3, chapter Syntaxe, pages 163–177. Presses Universitaires de Vincennes, Saint-Denis, 1999.
- [10] Carmen Dobrovie-Sorin, Alda Mari, et al. Generic plural indefinites: sums or groups? In *PROCEEDINGS-NELS*, volume 37, page 205, 2008.
- [11] Donka F Farkas and Henriëtte De Swart. Article choice in plural generics. *Lingua*, 117(9):1657–1676, 2007.
- [12] Brendan S Gillon. Towards a common semantics for english count and mass nouns. *Linguistics and philosophy*, 15(6):597–639, 1992.
- [13] Martin Hackl. On the grammar and processing of proportional quantifiers: most versus more than half. *Natural Language Semantics*, 17(1):63–98, 2009.
- [14] Jim Higginbotham. Mass and count quantifiers. *Linguistics and Philosophy*, 17(5):447–480, 1994.
- [15] Sveta Krasikova. Definiteness in superlatives. In *Logic, language and meaning*, pages 411–420. Springer, 2012.
- [16] Angelika Kratzer. Stage-level and individual-level predicates. *The generic book*, 125:175, 1995.
- [17] William Ladusaw. Semantic constraints on the english partitive construction. In *Proceedings of WCCFL*, volume 1, pages 231–242, 1982.
- [18] Fred Landman. *Structures for semantics*. Springer, 1991.
- [19] Daniel Lassiter. Measurement and modality: The scalar basis of modal semantics. *New York: New York University dissertation*, 2011.
- [20] Godehard Link. The logical analysis of plurals and mass terms: A lattice-theoretical approach. *Formal semantics: The essential readings*, pages 127–146, 2002.
- [21] Jan Tore Lønning. Mass terms and quantification. *Linguistics and philosophy*, 10(1):1–52, 1987.
- [22] Lisa Matthewson. Quantification and the nature of crosslinguistic variation. *Natural Language Semantics*, 9(2):145–189, 2001.

¹⁵I have left aside mass indefinite DPs, e.g., *much water*, *some water*, *one liter of water*, which are necessarily weak (in the sense of Milsark 1977) and are best analyzed as generalized existential quantifiers over amounts ([7]: 61–63).

¹⁶According to Dobrovie-Sorin ([7], [5]), Q-adverbs are entity-quantifiers when they take bare plurals or definite plurals in their restriction and set-quantifiers when they are built with singular indefinites or with inherently plural indefinites [10] [7].

- [23] Andrzej Mostowski. On a generalization of quantifiers. *Studies in Logic and the Foundations of Mathematics*, 93:311–335, 1979.
- [24] Nicholas Rescher. Plurality quantification. *The Journal of Symbolic Logic*, 27:373–374, 1964.
- [25] Dorian Roehrs. *Demonstratives and definite articles as nominal auxiliaries*, volume 140. John Benjamins, 2009.
- [26] Peter Roeper. Semantics for mass terms with quantifiers. *Noûs*, 17(2):251–265, 1983.
- [27] Anna Szabolcsi. Compositionality without word boundaries:(the) more and (the) most, 2012.
- [28] Anna Szabolcsi. Wordless compositionality: case studies in quantification, 02 2012.
- [29] Anna Szabolcsi and Frans Zwarts. Weak islands and an algebraic semantics for scope taking. *Natural language semantics*, 1(3):235–284, 1993.

Dowty's aspect hypothesis segmented

Tim Fernando

Trinity College Dublin, Ireland
Tim.Fernando@tcd.ie

Abstract

A form of the aspect calculus hypothesized in Dowty 1979 to explain “the different aspectual properties of the various kinds of verbs” is reinterpreted by chaining segments incrementally, as in Landman 2008. Segmenting an interval brings out, it is argued, a notion of event that can be represented at bounded but variable granularities by strings of sets of temporal propositions, tracking change. Simple accounts of telicity, durativity and composition interpreting Moens and Steedman 1988 are given in terms of such strings.

1 Introduction

The hypothesis from [Dow79] that “the different aspectual properties of the various kinds of verbs can be explained by postulating a single homogeneous class of predicates – *stative predicates* – plus three or four sentential operators and connectives” (page 71) has, over the years, profoundly influenced work on lexical aspect, in spirit, if not always in letter. Propositions are interpreted relative to interval world pairs, and a proposition φ said to be *homogeneous* when it holds at an interval world pair $\langle I, w \rangle$ precisely if it holds at $\langle \{t\}, w \rangle$ for every $t \in I$. Sentential operators are then applied for “a reductionist analysis of the aspectual classes of verbs” (page 71). Under a simplified reformulation from [Rot04] (page 35) that departs from the letter, if not the spirit, of the hypothesis, the operators DO, BECOME and CAUSE yield activities (1a), achievements (1b), and accomplishments (1c), with CAUSE reworked in (1c) using a culmination function Cul and a summation operation \sqcup_S producing singular entities.

- (1) a. activities $\lambda e. (\text{DO}(\varphi))(e)$
b. achievements $\lambda e. (\text{BECOME}(\varphi))(e)$
c. accomplishments $\lambda e. \exists e' [(\text{DO}(\varphi))(e') \wedge e = e' \sqcup_S \text{Cul}(e)]$

As an approximation of the accounts in [Dow79, Rot04], (1) is very rough, but serves to bring out events e , conspicuously absent in [Dow79]. It is clear from, say, the *eventuality structures* in [Lan08] that interpreting propositions as sets of interval world pairs (satisfying the propositions) does not preclude events. But we can test the reductionism in [Dow79] by asking: are events not *already* implicit in the interval world interpretation of propositions? The claim of the present paper is that they are, and that they can be made explicit along the “segmental” and “incremental axes” in [Lan08].

The segmental-incremental divide is used in [LR12] to refine the notion of homogeneity above (for stative φ) to an incremental notion for eventive predicates “sensitive to the arrow of time” (page 97). Turning to telicity, note the arrow of time differentiates the second half of a run to the post-office from its preceding half. Inasmuch as a run to the post-office is telic, and its second half (but not its first) is also a run to the post-office, it is problematic to equate telic predicates with quantized predicates ([Kri98]). Let us represent the arrow of time by a linear order \prec on temporal instants, lifted to intervals I and I' universally for “whole precedence”

$$I \prec I' \iff (\forall t \in I)(\forall t' \in I') t \prec t'.$$

A sequence $I_1 \dots I_n$ of intervals is a *segmentation* of an interval I if $\bigcup_{i=1}^n I_i = I$ and $I_i \prec I_{i+1}$ for $1 \leq i < n$ ([Fer13a]). That is, a segmentation of I is a finite partition of I ordered according to \prec . The point in partitioning an interval is to track changes within the interval in propositions, including statives that make achievements and accomplishments telic (by culminating).

The importance of choosing a suitable segmentation is illustrated by the rule (2), where we write $I \mathbf{m} I'$ to mean the sequence II' is a segmentation of $I \cup I'$ ¹

$$I \mathbf{m} I' \iff I \prec I' \text{ and } I \cup I' \text{ is an interval.}$$

$$(2) \quad I \mathbf{m} I' \text{ and } \langle I, w \rangle \not\models \varphi \text{ and } \langle I', w \rangle \models \varphi \implies \varphi \text{ becomes } w\text{-true between } I \text{ and } I'$$

One problem with (2) is that the interval I mentioned there may have a segmentation $I_1 I_2$ such that $\langle I_1, w \rangle \not\models \varphi$, but $\langle I_2, w \rangle \models \varphi$, in which case, under (2), φ becomes w -true between I_1 and I_2 (i.e., within I), as well as between I and I' . Such circumstances can be avoided through a negation \neg , called *predicate* negation in [Ham71] and *strong* negation in [AF94], with $\neg\varphi$ w -satisfied by intervals that have *no* subinterval w -satisfying φ

$$\langle I, w \rangle \models \neg\varphi \iff (\forall J \sqsubseteq I) \langle J, w \rangle \not\models \varphi.$$

(where the *subinterval* relation \sqsubseteq is \subseteq restricted to intervals). We then strengthen $\langle I, w \rangle \not\models \varphi$ in (2) to $\langle I, w \rangle \models \neg\varphi$ in (3).

$$(3) \quad I \mathbf{m} I' \text{ and } \langle I, w \rangle \models \neg\varphi \text{ and } \langle I', w \rangle \models \varphi \implies \varphi \text{ becomes } w\text{-true between } I \text{ and } I'$$

Note that $\langle I, w \rangle \models \neg\varphi$ collapses to $\langle I, w \rangle \not\models \varphi$ if the interval I is (φ, w) -homogeneous in that I has *no* subintervals J and J' such that $\langle J, w \rangle \models \varphi$ and $\langle J', w \rangle \not\models \varphi$ — i.e., if

$$(\exists J \sqsubseteq I) \langle J, w \rangle \models \varphi \iff (\forall J \sqsubseteq I) \langle J, w \rangle \models \varphi.$$

Under the assumptions in (3), $I \cup I'$ cannot be (φ, w) -homogeneous, though I' could (along with I), which would mean φ is w -satisfied by every subinterval of I' . Lest change get buried in a non-homogeneous interval (such as $I \cup I'$ in (3)), we will chain together homogeneous intervals (such as I and I') for a segmentation (such as II') bringing out change.

Change is detected through propositions φ , which it will be useful to assume have certain properties commonly associated with states since [BP72] — the subinterval property

$$\langle I, w \rangle \models \varphi \text{ and } J \sqsubseteq I \implies \langle J, w \rangle \models \varphi$$

and the additive property

$$\langle I, w \rangle \models \varphi \text{ and } \langle I', w \rangle \models \varphi \implies \langle I \cup I', w \rangle \models \varphi \quad \text{whenever } I \cup I' \text{ is an interval}$$

illustrated by the entailments (4) and (5), respectively.

$$(4) \quad \text{Ed slept from 1pm to 4pm} \implies \text{Ed slept from 2pm to 3pm}$$

$$(5) \quad \text{Ed slept from 1pm to 4pm and from 2pm to 5pm} \implies \text{Ed slept from 1pm to 5pm}$$

Combining these properties, let us say φ is *w-segmented* if for all intervals I and I' whose union $I \cup I'$ is an interval,

$$\langle I, w \rangle \models \varphi \text{ and } \langle I', w \rangle \models \varphi \iff \langle I \cup I', w \rangle \models \varphi.$$

This is somewhat weaker than the notion of homogeneity for statives φ in [Dow79] (mentioned in the opening paragraph above), under which φ is *w-pointwise* in that for every interval I ,

$$\langle I, w \rangle \models \varphi \iff (\forall t \in I) \langle \{t\}, w \rangle \models \varphi.$$

It turns out that if φ is *w-segmented*,

¹In [Ham71], \mathbf{m} is *abutment on the left*; in [KR93] *abutment*; and in [AF94] *meet*.

- (R1) a segmentation of I consists of (φ, w) -homogeneous intervals iff it w -tracks φ in I
and if moreover, φ is w -pointwise,
(R2) an interval I has a segmentation into (φ, w) -homogeneous subintervals iff φ does not
change its w -truth value infinitely often in I

where the right hand sides of (R1) and (R2) are made precise in the next section, section 2.
(R1) reinforces the view that segmentations into (φ, w) -homogeneous intervals are the way to represent φ 's changes in w , whereas (R2) describes the price to be paid for such representations. But why should we be interested in these representations? Because from representations of change, we can expect to extract representations of events that are parts of that change. To carry this out, we will work with not just one w -segmented proposition φ but a set X of such. We shall come under pressure to assume X is finite, and discover that X provides a useful notion of granularity that is bounded but refinable. Aspectual notions such as telicity and durativity will focus our attention in section 3 on the propositions in X , reducing, for every world w ,

- (i) an interval I to the set

$$X_w(I) = \{\varphi \in X \mid \langle I, w \rangle \models \varphi\}$$

of propositions $\varphi \in X$ w -satisfied at I , and

- (ii) a segmentation $I_1 \cdots I_n$ to its (X, w) -diagram, the string

$$X_w(I_1 \cdots I_n) = X_w(I_1) \cdots X_w(I_n)$$

over the alphabet 2^X of subsets of X .

For a fixed granularity X , we represent events and reconstruct the connectives **DO** and **BECOME** directly in terms of strings in $(2^X)^+$. Events can be conceived as truthmakers [Dav67] (page 91), relative to a notion \models_w of satisfaction between a segmentation $I_1 \cdots I_n$ and a string $\alpha_1 \cdots \alpha_m \in (2^X)^+$ requiring that $m = n$ and $\langle I_i, w \rangle$ satisfy each proposition in α_i

$$I_1 \cdots I_n \models_w \alpha_1 \cdots \alpha_m \iff n = m \text{ and } (\forall i \in [1, n])(\forall \varphi \in \alpha_i) \langle I_i, w \rangle \models \varphi$$

(where $[i, j]$ is the set of integers $\geq i$ and $\leq j$). The world parameter w is largely inactive throughout (sitting idly along for the ride). In fact, working with strings, we can dispense with worlds at the outset, constructing as many as we wish from a notion of branching time that enriches the incremental axis of [Lan08]. But to link up with [Dow79] and the tradition from which it sprang, we have no choice but to keep the worlds around.

Section 2 provides model-theoretic justification for a string-based approach to a form of the aspect calculus of [Dow79] close to [MS88]. This approach is taken up in section 3 and can be understood, up to a point, without section 2's justification.² Minimizing the technical fuss, we can summarize section 2 as follows. Relative to a set Φ of propositions and its family

$$Fin(\Phi) = \{X \subseteq \Phi \mid X \text{ is finite}\}$$

of finite subsets, an interval world pair $\langle I, w \rangle$ meeting assumptions largely familiar from [Dow79] is represented by a $Fin(\Phi)$ -indexed family $\{s_X\}_{X \in Fin(\Phi)}$ of strings $s_X \in (2^X)^*$ that picture $\langle I, w \rangle$ up to granularity X . Suitable substrings of the strings s_X represent (up to granularity X) events that happen within the interval world pair pictured. Section 3 bases its account of telicity and durativity on a simple choice of X . That account can be refined by expanding X .

²Indeed, a variant of section 3 is described in [Fer13a] without the benefit of Propositions 1 and 2 of section 2. Instead, [Fer13a] locates segmented propositions in an adjunction with (so-called) whole propositions, applying that adjunction to the imperfective/perfective divide (analyzed as viewpoint aspect).

2 From segmentations to strings

Attending promptly to (R1), let us track a w -segmented proposition φ in I through segmentations of I satisfying the following definition. A segmentation $I_1 \cdots I_n$ of I is (φ, w) -fine if for every subinterval I' of I , φ holds at $\langle I', w \rangle$ precisely if I' is covered by components I_i that w -satisfy φ

$$\langle I', w \rangle \models \varphi \iff I' \subseteq \bigcup \{I_i \mid i \in [1, n] \text{ and } \langle I_i, w \rangle \models \varphi\}.$$

Proposition 1. Let φ be a w -segmented proposition, and $\mathbb{I} = I_1 \cdots I_n$ be a segmentation of I .

- (a) \mathbb{I} is (φ, w) -fine iff $(\forall i \in [1, n]) I_i$ is (φ, w) -homogeneous.
- (b) If \mathbb{I} is (φ, w) -fine, then so is any segmentation that is a finer partition of I .
- (c) If \mathbb{I} is (φ, w) -fine, then so is any segmentation into $n - 1$ subintervals obtained from \mathbb{I} by merging two components I_i and I_{i+1} such that

$$(\langle I_i, w \rangle \models \varphi \text{ and } \langle I_{i+1}, w \rangle \models \varphi) \quad \text{or} \quad (\langle I_i, w \rangle \not\models \varphi \text{ and } \langle I_{i+1}, w \rangle \not\models \varphi).$$

An easy corollary of Proposition 1(c) is that if I has a (φ, w) -fine segmentation, it has a (φ, w) -fine segmentation $J_1 \cdots J_k$ such that

$$\langle J_i, w \rangle \models \varphi \iff \langle J_{i+1}, w \rangle \not\models \varphi \quad \text{for every } i \in [1, k - 1]$$

and this is the coarsest (and shortest) of all (φ, w) -fine segmentations of I . But if we want a (φ, w) -fine segmentation of I that is also (φ', w) -fine for a different w -segmented proposition φ' , the coarsest (φ, w) -fine segmentation may not do (making Proposition 1(b) relevant).

Collecting w -segmented propositions of interest into a set X , let us call a segmentation (X, w) -fine if it is (φ, w) -fine for every $\varphi \in X$. When does an interval have an (X, w) -fine segmentation? Clearly, a necessary condition is that no φ in X alternate between w -true and w -false infinitely often in I . More precisely, let us call a sequence $I_1 \cdots I_n$ of subintervals of I a (φ, w, n) -alternation in I if for all $i \in [1, n - 1]$, $I_i \prec I_{i+1}$ and

$$\langle I_i, w \rangle \models \varphi \iff \langle I_{i+1}, w \rangle \not\models \varphi.$$

An interval I is (X, w) -stable if for every $\varphi \in X$, there is an integer $n > 0$ such that no (φ, w, n) -alternation in I exists. The necessary condition that I be (X, w) -stable is also sufficient for finite X , provided we assume that each $\varphi \in X$ is not only w -segmented but (as [Dow79] does for statives) w -pointwise, making (R2) from the introduction precise in

Proposition 2. Fix an interval I , world w and finite set X of w -pointwise propositions.

- (a) I has an (X, w) -fine segmentation iff I is (X, w) -stable.
- (b) If I is (X, w) -stable, then there is a coarsest (and shortest) (X, w) -fine segmentation of I .

To describe the coarsest (X, w) -fine segmentation of an (X, w) -stable interval, it is helpful to define a segmentation $I_1 \cdots I_n$ to be (X, w) -compressed if for all $i \in [1, n - 1]$,

$$X_w(I_i) \neq X_w(I_{i+1})$$

where (as mentioned in the introduction) $X_w(I_i)$ is the set of propositions in X that $\langle I_i, w \rangle$ satisfy. Given an (X, w) -fine segmentation $I_1 \cdots I_n$ of I , we can merge contiguous blocks $I_i I_{i+1} \cdots I_{i+j}$ of components that w -satisfy the same propositions in X

$$X_w(I_i) = X_w(I_{i+1}) = \cdots = X_w(I_{i+j})$$

to form an (X, w) -compressed segmentation refined by $I_1 \cdots I_n$, which is, furthermore, the coarsest (and shortest) (X, w) -fine segmentation of I .

Focusing on strings, including the (X, w) -diagram $X_w(I_1) \cdots X_w(I_n)$ of a segmentation $I_1 \cdots I_n$, let us call a string $\alpha_1 \cdots \alpha_k$ *stutterless* if for all $i \in [1, k-1]$, $\alpha_i \neq \alpha_{i+1}$. Clearly, a segmentation is (X, w) -compressed iff its (X, w) -diagram is stutterless. Moreover, we can express the formation above of an (X, w) -compressed segmentation in terms of the *block compression* $bc(s)$ of a string s , compressing all contiguous blocks α^{j+1} of a symbol α to α

$$bc(s) = \begin{cases} bc(\alpha s') & \text{if } s = \alpha \alpha s' \\ \alpha \ bc(\beta s') & \text{if } s = \alpha \beta s' \text{ with } \alpha \neq \beta \\ s & \text{otherwise} \end{cases}$$

so that $bc(s)$ is stutterless. Next, we fix an arbitrary (possibly infinite) set Φ of w -segmented propositions, and picture a Φ -stable pair $\langle I, w \rangle$ through a function $\Delta_{I,w}$ with domain the set $Fin(\Phi)$ of finite subsets of Φ such that for every finite $X \subseteq \Phi$, $\Delta_{I,w}(X)$ is the (X, w) -diagram of the coarsest (X, w) -fine segmentation of I . The functions $\Delta_{I,w}$ belong to the *inverse limit* $\mathfrak{IL}(\Phi)$ of the $Fin(\Phi)$ -indexed family $\{bc_X\}_{X \in Fin(\Phi)}$ of functions $bc_X = \rho_X; bc : (2^\Phi)^* \rightarrow (2^X)^*$ mapping $s \in (2^\Phi)^*$ to the block compression $bc(\rho_X(s))$ of the X -restriction $\rho_X(s)$ intersecting each component of s with X

$$\rho_X(\alpha_1 \cdots \alpha_k) = (\alpha_1 \cap X) \cdots (\alpha_k \cap X).$$

An example is (6), where boxes are drawn instead of curly braces $\{\}$ for sets as symbols.

$$(6) \quad bc_{\{E\}}([E | E, R | E]) = bc(\rho_{\{E\}}([E | E, R | E])) = bc([E | E | E]) = [E]$$

For the record, $\mathfrak{IL}(\Phi)$ is the set of functions f with domain $Fin(\Phi)$ such that

$$f(X) = bc_X(f(Y)) \text{ whenever } X \subseteq Y \in Fin(\Phi).$$

Φ -stable pairs $\langle I, w \rangle$, all of which are represented in $\mathfrak{IL}(\Phi)$, vary wildly, giving $\mathfrak{IL}(\Phi)$ an intensional dimension. *Incremental inclusion* \subseteq_i from [Lan08] on intervals I and I'

$$I \subseteq_i I' \iff I \sqsubseteq I' \text{ and not } (\exists t' \in I') \{t'\} \prec I$$

can be rendered into strings s and s' as the prefix relation

$$s \text{ prefix } s' \iff (\exists s'') ss'' = s'$$

construing s within a string $s\alpha$ as the past of α . We lift *prefix* to $\mathfrak{IL}(\Phi)$, building into an irreflexive relation \prec_Φ on $\mathfrak{IL}(\Phi)$ the prefix requirement at every granularity X

$$f \prec_\Phi f' \iff f \neq f' \text{ and } (\forall X \in Fin(\Phi)) f(X) \text{ prefix } f'(X).$$

As noted in [Fer13a], \prec_Φ is tree-like ([Dow79]); i.e. transitive and left linear: for all $f \in \mathfrak{IL}(\Phi)$,

$$f_1 \prec_\Phi f_2 \text{ or } f_2 \prec_\Phi f_1 \text{ or } f_1 = f_2 \quad \text{whenever } f_1, f_2 \prec_\Phi f.$$

No element of $\mathfrak{IL}(\Phi)$ is \prec_Φ -maximal, which is to say all intervals I that form Φ -stable pairs (I, w) can be extended further, whether or not they are conceived within w to stretch through an entire timeline. That said, some small part of $\mathfrak{IL}(\Phi)$ may suffice for a specific purpose. In particular, a fixed $f \in \mathfrak{IL}(\Phi)$ may do, inducing at every granularity $X \in Fin(\Phi)$ an interpretation of $\varphi \in X$ from $f(X) = \alpha_1 \cdots \alpha_n$, with temporal instants $t \in [1, n]$ and interval satisfaction

$$[i, j] \models \varphi \iff (\forall t \in [i, j]) \varphi \in \alpha_t$$

whenever $1 \leq i \leq j \leq n$. Worlds are absent in this interpretation, and, borrowing a phrase from [Kah11], what you see is all there is (WYSIATI).³

3 Composing transitions, telicity and durativity

We now step from propositions (interpreted over intervals) that represent states to strings $\alpha_1 \cdots \alpha_n$ of length $n > 1$ (interpreted over segmentations) that represent events with precondition α_1 and postcondition α_n . We build on [MS88] for (7) and (8) below.⁴

	non-durative	durative
(7)	telic	achievement $\boxed{\neg\varphi} \boxed{\varphi}$
	atelic	semelfactive $\boxed{\psi}$
		accomplishment $\boxed{\neg\varphi} \boxed{\neg\varphi, \psi} \boxed{\neg\varphi, \psi}^+ \boxed{\varphi}$
		activity $\boxed{\psi} \boxed{\psi}^+$

- (8)
 - a. $\alpha_1 \cdots \alpha_n$ is *durative* if its length n is ≥ 3
 - b. $\alpha_1 \cdots \alpha_n$ is *telic* if there is some φ in α_n such that for all $i \in [1, n - 1]$, the negation $\neg\varphi$ of φ appears in α_i
 - c. $\text{iterate}_{;}(\boxed{\psi}) = \boxed{\boxed{\psi} \boxed{\psi}}^+$
 - d. $\boxed{\boxed{\psi} \boxed{\psi}}^+ ; \boxed{\neg\varphi} \boxed{\varphi} \approx \boxed{\neg\varphi} \boxed{\neg\varphi, \psi} \boxed{\neg\varphi, \psi}^+ \boxed{\varphi}$

(8a) and (8b) give simple definitions of durative and telic strings, exemplified in (7). (8c) reduces an activity in $\boxed{\psi} \boxed{\psi}^+$ to the iteration of the semelfactive $\boxed{\psi}$, while (8d) decomposes an accomplishment into an activity and achievement, not unlike (1c) in the opening paragraph above (with \sqcup_S as ;). (8c) and (8d) depend on operations, ; and $\text{iterate}_{;}$, defined in (9). In (9a), β and α range over subsets of some fixed set Φ of propositions, while s and s' range over strings in $(2^\Phi)^*$, including the empty string ϵ mentioned in (9b). L and L' in (9b) and (9c) are languages, understood as subsets of $(2^\Phi)^*$.

- (9)
 - a. $s\beta ; \alpha s' = s(\beta \cup \alpha)s'$
 - b. $L; L' = \{s; s' \mid s \in L - \{\epsilon\} \text{ and } s' \in L' - \{\epsilon\}\}$ (regular if L and L' are)
 - c. $\text{iterate}_{;}(L) = \text{the } \subseteq\text{-least set } Z \text{ such that } L; L \subseteq Z \text{ and } Z; L \subseteq Z$

Observe that ; is essentially string concatenation, except that the last symbol β of the first string is merged with the first symbol α of the second string. The rationale behind this overlap is a notion of inertia that is expressed, for instance, in page 49 of [Com76]

³More about (and around) Propositions 1 and 2 in [Fer13b].

⁴(7) here is comparable to Figure 1, page 17 of [MS88], and (8) to Figure 2, page 18, with non-durative \sim atomic, durative \sim extended, telic \sim +conseq, atelic \sim -conseq, $\varphi \sim$ consequent state, $\psi \sim$ progressive state, and achievement/accomplishment/semelfactive/activity \sim culmination/culminated process/point/process.

unless something happens to change [a] state, then the state will continue

together with the assumption that in $s; s'$, nothing happens between the end of s and the start of s' . That same notion of inertia accounts for the discrepancy behind the wavy \approx in (8d)

$$\boxed{\boxed{\psi} \boxed{\psi}}^+ ; \boxed{\neg\varphi} \boxed{\varphi} = \boxed{\boxed{\psi}}^+ \boxed{\neg\varphi, \psi} \boxed{\varphi} \neq \boxed{\neg\varphi} \boxed{\neg\varphi, \psi} \boxed{\neg\varphi, \psi}^+ \boxed{\varphi}$$

with $\neg\varphi$ spreading (within the accomplishment) back from the achievement $\boxed{\neg\varphi} \boxed{\varphi}$. But then why doesn't ψ in the activity $\boxed{\boxed{\psi} \boxed{\psi}}^n$ also spread? Quoting more fully from [Com76], page 49

With a state, unless something happens to change that state, then the state will continue ... With a dynamic situation, on the other hand, the situation will only continue if it is continually subject to a new input of energy.

The reason $\neg\varphi$ flows in (8d) while ψ does not is that $\neg\varphi$ represents a state in $\boxed{\neg\varphi} \boxed{\varphi}$, whereas in an activity $\boxed{\boxed{\psi} \boxed{\psi}}^n$, the proposition ψ represents a dynamic situation that persists only if forced (moving loosely from a scalar, energy, to a vector, force). A precise mechanism for regulating inertial flow in the absence and presence of force is described in [Fer08], under which $\neg\varphi$ is inertial, whereas ψ is not.

But just what is ψ ? For an answer, it is instructive to cast $\boxed{\neg\varphi} \boxed{\varphi}$ as $\boxed{\hat{\varphi}}$ with $\hat{\varphi}$ set to $\varphi \wedge \langle \text{mi} \rangle \neg\varphi$, where $\langle \text{mi} \rangle$ is the diamond modal operator given by the inverse mi of meet m

$$\langle I, w \rangle \models \langle \text{mi} \rangle \chi \iff (\exists I' \text{ m } I) \langle I', w \rangle \models \chi$$

so that $\hat{\varphi}$ describes a dynamic situation

$$\langle I, w \rangle \models \varphi \wedge \langle \text{mi} \rangle \neg\varphi \iff (\exists I' \text{ m } I) I' I \models_w \boxed{\neg\varphi} \boxed{\varphi}.$$

A segmentation cannot w -satisfy $\boxed{\hat{\varphi}}^n$ for $n \geq 2$ any more than it can w -satisfy a string in $\boxed{\neg\varphi} \boxed{\neg\varphi, \varphi}^+ \boxed{\varphi}$, which under (9c), is $\text{iterate}_+(\boxed{\neg\varphi} \boxed{\varphi})$. But we can modify $\hat{\varphi}$ so that strings in the set $\boxed{\boxed{\psi} \boxed{\psi}}^+ = \text{iterate}_+(\boxed{\boxed{\psi}})$ are w -satisfiable. One way is through a linearly ordered set D of degrees to which a function $\deg_{\varphi, w}^D$ maps temporal instants. Propositions $d < \varphi\text{-deg}$ and $\varphi\text{-deg} \leq d$ are then interpreted over intervals by universal quantification (as with whole precedence \prec)

$$\begin{aligned} \langle I, w \rangle \models d < \varphi\text{-deg} &\iff (\forall t \in I) d < \deg_{\varphi, w}^D(t) \\ \langle I, w \rangle \models \varphi\text{-deg} \leq d &\iff (\forall t \in I) \deg_{\varphi, w}^D(t) \leq d \end{aligned}$$

(conflating, for simplicity, semantic entities $d, <$ and \leq to the left of \models with their syntactic representations to the right of \models). It is natural to assume a contextually given threshold \hat{d} reducing φ to $\hat{d} < \varphi\text{-deg}$. Existentially quantifying degrees, we put

$$\varphi_{\uparrow}^D = (\exists d \in D) (d < \varphi\text{-deg} \wedge \langle \text{mi} \rangle \varphi\text{-deg} \leq d)$$

and stretch instantaneous change $\boxed{\neg\varphi} \boxed{\varphi}$ to a more graduated sequence $\boxed{\varphi_{\uparrow}^D}^n$ associating an increasing sequence $d_1 < d_2 < \dots < d_n$ of degrees with a segmentation $I_0 I_1 \dots I_n$

$$\begin{aligned} I_0 I_1 \dots I_n \models_w \boxed{\varphi_{\uparrow}^D}^n &\iff (\exists d_1 \in D) \dots (\exists d_n \in D) (\forall i \in [1, n]) \\ &\quad (\forall t \in I_{i-1}) (\forall t' \in I_i) \deg_{\varphi, w}^D(t) \leq d_i < \deg_{\varphi, w}^D(t'). \end{aligned}$$

But is φ_{\uparrow}^D *w*-segmented? Not necessarily. Let us say that $\deg_{\varphi,w}^D$ is *increasing within* an interval I if for all $t, t' \in I$, if $t \prec t'$ then $\deg_{\varphi,w}^D(t) < \deg_{\varphi,w}^D(t')$. Given $\langle I, w \rangle \models \varphi_{\uparrow}^D$, one can show

$$\begin{aligned}\deg_{\varphi,w}^D \text{ is increasing within } I &\iff (\forall J \sqsubseteq I) \langle J, w \rangle \models \varphi_{\uparrow}^D \\ &\iff \langle I, w \rangle \models [\Box] \varphi_{\uparrow}^D\end{aligned}$$

where $[\Box]$ is the necessity operator induced by the inverse of the subinterval relation. The proposition $[\Box] \varphi_{\uparrow}^D$ can be assumed to be *w*-segmented, whereas φ_{\uparrow}^D cannot (unless interpreted against say, a string supporting WYSIATI, as described in the previous section). It is not entirely obvious that as a representation of a dynamic situation, the proposition ψ in an activity $\boxed{\psi} \boxed{\psi}''$ should be required to be *w*-segmented. What is clear from the lively literature on aspectual composition is that there is much more to say about what ψ to box (not to mention how to do so compositionally from syntactic input) under the scheme (7) – (9) or some revision thereof. The strings in (7) are simple examples of how to flesh out (8) and (9) based on a minimal set X of propositions that we can, under the inverse limit $\mathfrak{IL}(\Phi)$ in section 2, expand indefinitely, refining granularity (to express, for instance, forces and grammatical aspect).⁵

References

- [AF94] James F. Allen and George Ferguson. Actions and events in interval temporal logic. *J. Logic and Computation*, 4(5):531–579, 1994.
- [BP72] Michael Bennett and Barbara Partee. Toward the logic of tense and aspect in English. Technical report, System Development Corporation, Santa Monica, California, 1972. Reprinted in Partee 2008.
- [Com76] Bernard Comrie. *Aspect*. Cambridge University Press, 1976.
- [Dav67] Donald Davidson. The logical form of action sentences. In Nicholas Rescher, editor, *The Logic of Decision and Action*, pages 81–95. University of Pittsburgh Press, 1967.
- [Dow79] David R. Dowty. *Word Meaning and Montague Grammar*. Reidel, Dordrecht, 1979.
- [Fer08] Tim Fernando. Branching from inertia worlds. *J. Semantics*, 25(3):321–344, 2008.
- [Fer13a] Tim Fernando. Segmenting temporal intervals for tense and aspect. In *13th Meeting on the Mathematics of Language (MoL 13)*, pages 30–40. ACL Archive, 2013.
- [Fer13b] Tim Fernando. Partitions representing change homogeneously. Manuscript (5 pages), 2013.
- [Ham71] Charles Leonard Hamblin. Instants and intervals. *Studium Generale*, 24:127–134, 1971.
- [Kah11] Daniel Kahneman. *Thinking, Fast and Slow*. Farrar, Straus and Giroux, 2011.
- [KR93] Hans Kamp and Uwe Reyle. *From Discourse to Logic*. Kluwer, 1993.
- [Kri98] Manfred Krifka. The origins of telicity. In Susan Rothstein, editor, *Events and Grammar*, pages 197–235. Kluwer, 1998.
- [Lan08] Fred Landman. 1066: On the differences between the tense-perspective-aspect systems of English and Dutch. In Susan Rothstein, editor, *Theoretical and Crosslinguistic Approaches to the Semantics of Aspect*, pages 107–166. John Benjamins, 2008.
- [LR12] Fred Landman and Susan Rothstein. The felicity of aspectual for-phrases, part 2: Incremental homogeneity. *Language and Linguistics Compass*, 6(2):97–112, 2012.
- [MS88] Marc Moens and Mark Steedman. Temporal ontology and temporal reference. *Computational Linguistics*, 14(2):15–28, 1988.
- [Rot04] Susan Rothstein. *Structuring Events*. Blackwell, Oxford, 2004.

⁵Not a single durative string in (7) is stutterless. Such a string cannot make it into $\mathfrak{IL}(\Phi)$ on its own, but can do so as part of a stutterless string, for example, $\boxed{\psi, R} \boxed{\psi}$ with a Reichenbachian reference time R .

A two dimensional analysis of the future: modal adverbs and speaker's bias

Anastasia Giannakidou^{1*}and Alda Mari²

¹ University of Chicago, Chicago USA
giannaki@uchicago.edu

² Institut Jean Nicod (CNRS/ENS/EHESS, Paris, France
alda.mari@ens.fr

Abstract

Whether future morphemes in languages are temporal or modal operators is a central question in the semantics of the future. Most analyses (with the exception of [14]; see its rebuttal in [14]), agree that future morphemes do convey modality, and do not always make reference to future times. We study here the temporal (future time) reading of the future (*John will arrive at 5pm*), which we call ‘predictive’. The modality of the predictive reading is often assumed to be purely metaphysical (e.g. [13]). Based on novel Greek and Italian data with modal adverbs, we argue that prediction involves both a metaphysical and an epistemic dimension, and we offer an analysis that relies on knowledge of the speaker at the utterance time t_0 . This knowledge restricts the set of the futures (metaphysical branches) only to reasonable ones (in the sense of [18]). In these branches the prejacent comes out true. We also propose that when predicting the speaker has a degree of *bias* that the actual world-to-come at t_0 will be reasonable. This bias, to our knowledge not observed before, is reflected in the use of adverbs, and although by default it expresses high confidence, it can be quite variable, from very strong to relatively weak— as is evidenced by the wide range of modal adverbs that can co-occur with the future morpheme.

Contents

1	Introduction	1
2	The predictive reading: basic data with modal adverbs	3
3	Future: metaphysical alternatives and speaker's bias	5
4	Conclusion: what is common in epistemic and predictive future	7

1 Introduction

The question of whether the category ‘future’ in natural languages is a tense or modality has received a lot of attention in linguistic semantics, and both answers have been explored (for modal accounts, see e.g. [4]; [13], [9], [16], [19], [11]; [2]; [15]; for a defense of the temporal analysis see [14]). In previous work ([19], [9], [11]) we showed that Greek and Italian future morphemes have non-temporal, epistemic modal uses with past and present, such as:

*We would like to thank the participants at the seminar ‘How do we know what happens?’ Institut Jean Nicod on subjective meaning and non-veridicality, and colloquia at the Université Libre de Brussels and Groningen where this material was presented. We are thankful to Paul Egré, Jack Hoeksema, Salvador Mascarenhas, John Nerbonne, Bryan Renne, Seth Yalcin, and especially Janneke Huitink, Mikhail Kissine and Dan Lassiter for their written comments on earlier drafts of this material. Finally, our thanks to the Amsterdam Colloquium reviewers for their helpful suggestions.

- (1) a. I Ariadne tha ine arrosti.
 the Ariadne FUT be.3sg sick
 - b. Giovanni sarà malato.
 Giovanni FUT-be sick
 'Ariadne/Giovanni must be sick.'
- (2) a. I Ariadne tha itan arrosti.
 the Ariadne FUT be.past.3sg sick
 - b. Giovanni sarà stato malato.
 Giovanni FUT-be been. PERF. sick
 'Ariadne/Giovanni must have been sick.'

As we see, Italian and Greek employ future morphemes - which we will call FUT in this paper: a bound one (Italian), and a particle that precedes the tensed verb (Greek). FUT is followed by a nonpast or a past tense form in the sentences above, and produces an epistemic, non-predictive reading – notice that in the translation we use "must". *Will* would have been inappropriate in this use. This link of the category 'future' to epistemic modality appears to be explored quite productively in languages (see also [2] for Dutch, and [4] for similar ideas about Dutch *zullen* and *will*). Given this purely epistemic usage and the lack of future reference, it becomes appealing to treat the future morphemes in Greek and Italian like epistemic modals, which we did in our earlier work (cited earlier). In the present paper, we want to address the canonical, predictive reading of the future. We argue that the future in this reading conveys an assessment, based on knowledge and reasonability, that the speaker makes about an event that hasn't happened yet. In the end, we discuss briefly how the purely epistemic and the predictive reading are related, and what is the semantics they share (see for extended discussion [10]). A typical predictive use is illustrated in (3):

- (3) a. Tha vrexí avrio. (Greek)
 FUT rain tomorrow
- b. Domani pioverà. (Italian)
 Tomorrow rain.FUT.3sg
 'It will rain tomorrow'

Epistemically, what will happen in the future is unknown at the present time, therefore the future is non-veridical [7] and allows for *p* and *not p* (i.e. to rain or not rain tomorrow). However, speakers have knowledge at the speech time that determines what would count as reasonable course of events, and, we argue, they use that knowledge when they make predictions to create a bias towards *p*. A bias towards *p* for the future does not imply commitment of the speaker to *p*, as is the case in veridical unmodalized past and present assertions. Our proposal in a nutshell is that in reasonable worlds (i.e. those worlds in the metaphysical modal base in which nothing unexpected happens, see [18]), the prejacent is true at a future time. So, although the future is objectively and epistemically non veridical (or 'unsettled' [13]), speakers project their knowledge to carve out the metaphysical alternatives, into reasonable and unreasonable ones, i.e. *p* worlds and *non p* worlds. After this clean-up, FUT *p* conveys the semblance of veridicality because it universally quantifies over the reasonable worlds. This, however, breaks down once we add modal adverbs.

Our analysis is in line with [4] and [13], who also treat the future as a universal quantifier over a restricted set of alternatives (most normal ones, or boulethic, etc). *But this is not the whole story for us:* we present new data where FUT co-occurs with modal adverbs of variable force and texture (meaning *possibly*, *probably*, *definitely*, *really*, or containing numerals). What

is the function of these adverbs? We will argue that they reveal a supplementary component of the future: a presupposition that *the speaker is assessing how likely it is that the actual world to come will be a reasonable p world*. We suggest that these two components, truth conditional and presuppositional, are also at play in the epistemic reading, thus adding a presuppositional layer in epistemic modality within Kratzer's semantics, while keeping the same ontological structure.

In section 2 we present the adverbs data, and in section 3 our analysis for prediction. Section 4 concludes the paper.

2 The predictive reading: basic data with modal adverbs

Reference to a future event (the ‘predictive reading’) is done with the particular combination of FUT plus a lower non-past (which in Greek must additionally be perfective; [8], [9]).

- (4) a. O Janis tha ftasi avrio.
The John FUT arrive.perfective.nonpast.3sg tomorrow.
b. Gianni arriverà domani.
Gianni FUT-arrive tomorrow.
‘Gianni will/#must arrive tomorrow’.

As we see, the paraphrase with *must* is pretty odd in this case. Because of the adverb, the speaker has specific temporal information about where to place the eventuality. She might, for instance, be knowledgeable, at the time of utterance, about Gianni’s schedule for the day, or she might know that Gianni every day takes the train from Gare du Nord at 4:30 which will place him at the Jean Nicod at 5 pm. So, the future reading is forward shifting of the eventuality time - not of the evaluation time of FUT - which remains PRES [11] - due to the adverb scoping above the TP but lower than the FUT:

- (5) FUT > alle 5 del pomeriggio > arrive

The forward shifted, predictive reading of FUT is non-veridical - at the time of knowledge/assessment, the future event hasn’t happened yet, so FUT *p* does not entail *p* [7]. FUT *p*, however, does seem to express, as we said earlier, a positive bias of the speaker that the event will happen. Crucially, this bias is not full commitment of the speaker to *p* in the sense that an unmodalized past or present assertion is (which are veridical). FUT *p* does not commit the speaker to *p*, does not add *p* to the common ground. So FUT *p* is weaker than PAST *p*, and the modal adverbs are important in bringing about this weakness as we see next.

FUT (including *will*) appears to be compatible, in additional to the temporal adverbs, also with modal adverbs – and these can be of various strength, including some really weak ones.

- (6) a. Arriverà forse/certamente/probabilmente alle 4.
b. O Janis tha erthi isos/sigoura/malon/isos stis 4.
‘John will maybe/definitely/probably/maybe arrive at 4.’

The appearance of those modal adverbs – which are also observed in the non-predictive, purely epistemic reading of the future [11], makes us pause and think immediately of ‘modal concord’, as was indeed suggested by [9]:

- (7) I Ariadne malon/profanos/sigoura tha ine jatros.
the Ariadne probably/obviously/certainly FUT be.3sg doctor
‘Ariadne must probably/obviously/certainly be a doctor.’

However, in current theories of modal concord (e.g. [12]), there is matching force of the modal and the adverb. [16], [17] after [1], notes that in the non-predictive use, in Italian, the future is compatible with adverbs of very weak force like *forse*. With a weak adverb such as *isos*, *forse* 'maybe', the force of the future— which is that of a universal modal— seems incompatible.

- (8) Giacomo sarà forse/probabilmente/sicuramente un dottore.
 Giacomo be.FUT.3sg maybe/probably/certainly a doctor
 'Giacomo must maybe/probably/certainly be a doctor.'

The same observation holds for the predictive reading, as we just saw, in both Greek and Italian. In the predictive reading, FUT is compatible with adverbs of very weak force. A prediction with *isos*, *force* 'maybe' is hardly a strong one, and if the future morpheme is a universal it seems inaccurate to talk about 'concord'. *Will*, crucially, shows the same pattern, and can combine, just like the Greek and Italian futures, with something as weak as *maybe*, as we see. Things get even more 'loose' when we consider low-probability adverbials (like 10% *chance*, 30% *chance*):

- (9) a. Ine 10%, 30%, 90% sigouro oti o Janis tha erthi stis 5.
 is 10% 30% 90% certain that the John FUT come at 5
 b. Gianni arriver, al 30%, 30%, 90%, alle 5.
 Gianni arrive.FUT.3sg, at 30%, 30%, 90%, at 5pm
 'There is 10%/30%/90% probability that John will come at 5.'

So, when we predict, we seem to be in a paradoxical situation: on the one hand, we feel secure enough to make the prediction, but on the other, we are aware that our prediction is not fullproof. The adverb is there to signal that the prediction is not commitment to *p*, but rather bias toward *p*. We will argue that it is used as a probability measure, as a way for the speaker to indicate the degree of confidence she has in her prediction. This is an important point to make, and to our knowledge it has not been noticed before. The indication of confidence is highly sensitive to what the speaker knows or takes as a reliable information, and is suggestive of an epistemic layer in the prediction.

What will happen in the future is objectively unsettled at the speech time [13] and therefore objectively non-veridical: the future event hasn't happened yet. But knowledge at the present time creates bias, i.e. it allows the speaker to project alternatives in which *p* is true, as we propose next (we will call these metaphysical alternatives 'reasonable futures' using a notion of normality borrowed from [18]). Reasonable futures are metaphysical alternatives, not yet realized, and nothing guarantees that they will be. So, clearly, we are not proposing an account in which the future can be, or is, known; FUT *p* does not commit the speaker to *p*, as is the case of know *p* and other veridical operators. We are proposing an account where the speaker (a) has enough knowledge to propose *p* as being true in a reasonable future, (b) has high confidence the the actual world will be a *p* world (what we called bias towards *p*), while (c) remaining uncertain as to whether the actual world will in fact be a *p* world.

The flexibility observed with adverbs allows us to say that the predictions (how likely is that the actual world to come belongs to the set of alternatives in which *p* is true) fall generally into two categories: predictions of high confidence (based on adequate knowledge that allows such high confidence), and predictions based on low confidence (based on less adequate knowledge). One can think of such low confidence predictions as *guesses*. In our semantics next, we will try to capture these ideas in a precise way, by using the notion of reasonable future [18]. Once we define reasonable futures, we will claim that the future morpheme universally quantifies over these, and that the adverbs measure how confident the speaker is that the actual world to come is in the set of reasonable futures.

3 Future: metaphysical alternatives and speaker's bias

Let us introduce the notion of *reasonable future* (from [18]). Consider an example:

- (10) Gianni arriverà a Roma con la macchina alle 4pm.
John will arrive in Rome by car at 4pm

When uttering this sentence, the speaker considers only reasonable futures and states that the event of arrival at 4pm will take place in these futures. What is a reasonable future?

Following [18], a reasonable future determined at time t is such that the rules that hold at t are maintained. As FUT sets the time of the perspective at present, the time at which reasonable futures are considered is the time of the utterance. Reasonable futures are such that also habits and behaviors do not change there. Of course habits can change. However, those futures in which they change are discarded (for a comparison between the notion of ‘reasonability’ and inertia ([5], see [18])). Our notion of reasonability borrowed from [18] is reduced to the total absence of interferences that would disrupt the course of the law/habit/behavior that holds at the time at which reasonable futures are considered. The future, we propose, is a universal quantifier over reasonable alternatives (as these are metaphysical alternatives). *However, the actual world to come is not known at the utterance time to belong to the set of reasonable future at a later time t' . We claim that the future presupposes confidence of the speaker that the actual world to come belongs to the set of reasonable ones, i.e. presupposes confidence that the actual world to come is a p world.*

We use the framework designed in [18]. We use a $W \times T$ forward-branching structure [20] [3]. A three-place relation \simeq on $T \times W \times W$ is defined such that (i) for all $t \in T$, \simeq_t is an equivalence relation; (ii) for any $w, w' \in W$ and $t, t' \in T$, if $w' \simeq_{t'} w$ and t precedes t' , then $w' \simeq_t w$ (we use the symbols \prec and \succ for temporal precedence and succession, respectively). In words, w and w' are historical alternatives at least up to t' and thus differ only, if at all, in what is future to t' .

For any given time, a world belongs to an equivalence class comprising worlds with identical pasts but different futures. Let w_0 be the actual world.

For any time $t \in T$, we define the *common ground* $cg(t)$ as the set of worlds that are identical to the actual world w_0 at least up to and including t .

$$(11) \quad cg(t) := \{w \mid w \simeq_t w_0\}$$

In the case depicted in Figure 1, the common ground at t is the set given in (12).

$$(12) \quad cg(t) = \{w_1, w_2, w_0, w_3, w_4\}$$

So defined, the common ground includes any world branching from the actual world at a time equal to or after t , including those worlds that are highly different in their causal laws from the actual world as well as those worlds in which current causal and social laws malfunction.

[18], p.17 defines ‘reasonable futures’ as in (13). For any $t \in T$,

$$(13) \quad ReasFut(t) := \{w_i \in cg(t) \mid w_i \text{ is such that the set of rules fixed at } t \text{ continue to hold in } w_i\}$$

Let us assume that w_3 in Figure 1 is a world in which causal and social laws malfunction. The set of reasonable futures defined at time t does not include w_3 . In the case depicted in Figure 2, the set of reasonable futures fixed at t is given in (14).

$$(14) \quad ReasFut(t) = \{w_1, w_2, w_0, w_4\}$$

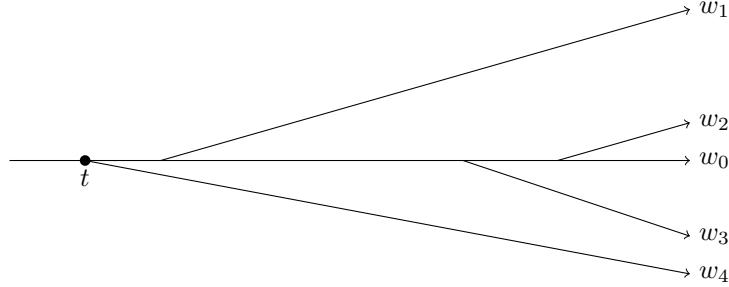


Figure 1: $cg(t)$ [18], p.17, fig. 3

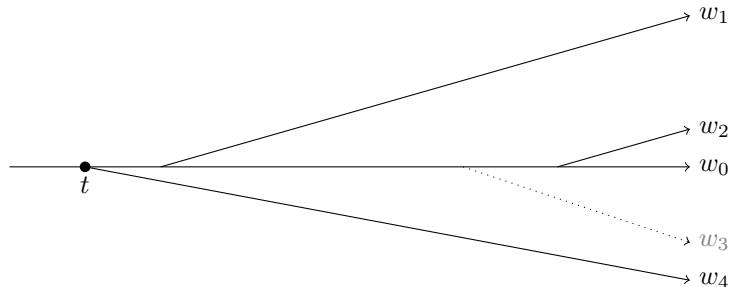


Figure 2: Reasonable Future Worlds (ReasFut) [18], p. 18, fig 4

As we have explained, however, the set of reasonable worlds determined at a certain time t is different according to different epistemic agents. Hence, amending (13), we define (15), where epistemic agents are added as parameters.

- (15) Given the epistemic agent i , $\text{ReasFut}(t) := \{w_i \in cg(t) \mid w_i \text{ is such that the set of rules fixed at } t \text{ and considered by } i \text{ continue to hold in } w_i\}$

Reasonable futures are not a projection of speakers' preferences and beliefs; *they are metaphysical alternatives* that do not exist yet at the time of the utterance. Let us recall with [18], p. 18, that, at t_u , one can state what the reasonable futures of time t_u are; however, one cannot state whether the actual-world-to-come belongs to the set of reasonable possibilities, since the actual-world-to come does not exist yet at t_u . Yet the speaker has positive bias, i.e. confidence that the actual world to come will be a reasonable one. In the bare case (*John will come at 4pm*), the assumption that the actual world to come will coincide with a reasonable future at 4pm is ranked very highly, which means, associated with a high degree of confidence. The adverbs are the measures of this confidence.

How can this dimension of bias/confidence be captured? One thing seems clear: the adverbs do not directly compose with p or $FUT\ p^1$. Given the variability in force we observed in section 2, especially given that we have even existential adverbs such as *forse*, and in English *maybe*

¹The adverbs directly compose with FUT at least in the purely epistemic reading in Greek (see [9]), they do not in Italian, even for the epistemic reading, see [19].

(*Maybe John will come*), it is impossible to combine FUT (a universal) with the adverbs in a concord like manner. Even the most flexible theories of modal concord (such as [12]) would require at least matching force. So, the adverb is forced to contribute at the non-at issue level, something consistent with the fact that modal adverbs are speaker oriented [6] and seem to ‘scope’ high in the sentence (Ernst therefore argued that they are positive polarity items).

We will claim that the modal adverbs contribute a presupposition that there is a measure of bias in the assessment of the epistemic agent i (in the default unembedded case, the speaker) which measures how committed i is that the actual world will be within the set of the reasonable worlds. For instance, if the adverb is *probably*, and the measure interval is between 1 and 100, then *probably* FUT(p) says that, according to the speaker, the probability of the actual world to be a reasonable future (where p is true) is between 85 and 100 percent. We are now ready to provide an analysis for the predictive use of the future.

- (16) FUT(NON-PAST(p)) = 1 iff $\forall w' \in ReasFut(t_u) : \exists t' \in [t_u, \infty] \wedge p(w't')$

Presupposition: There is a probability measure $\mu_{likelihood}$ that measures the likelihood, according to the speaker, that the actual-world-to-come will be in *ReasFut*(t_u). The default value of $\mu_{likelihood}$ is *probably*, above 80%. (positive bias towards p)

Truth condition: p is true only in the the reasonable futures, given the evidence: $\forall w' \in ReasFut(t_u) : p(w') = 1$. All reasonable futures are p futures.

Hence the future, as a universal quantifier over the worlds in *ReasFut*, does express commitment of the speaker in that set. However, the speaker is still uncommitted about whether the actual world to come is in that set. This accounts for the non veridical nature of the predictive future, and allows us to unify it, as we suggest next, with the purely epistemic future.

4 Conclusion: what is common in epistemic and predictive future

We proposed here an analysis of the predictive reading of FUT using metaphysical, not epistemic alternatives. We said hat the speaker carves out a set of metaphysical alternatives in which p comes out true in the future. We explained that these are the metaphysical alternatives in which ‘nothing goes wrong’, given present knowledge. Modal adverbs, at the same time, measure the speaker bias towards p , i.e. they measure how likely the speaker thinks is that the actual world to come will be a reasonable world. This bias, especially with weaker adverbs, appears to weaken the strength of the future. In [10], we offer a parallel analysis of the epistemic reading:

- (17) Epistemic FUT (= future as MUST exs 1,2)

Epistemic-FUT p asserts *necessarily* p , relative to an epistemic modal base, an ordering source and an epistemic agent i .

Presupposition: there is a probability measure function $\mu_{likelihood}$ determined by i that measures the likelihood, according to i that the actual world is within the set of the best worlds. The default value of $\mu_{likelihood}$ is *probably*, above 80%.

Truth condition: p is true only in the best worlds, given the evidence: $\forall w' \in Best_{g(w)}(\cap f(w)) : p(w') = 1$.

Here, the modal domain is the set of best words (rather than reasonable future worlds). We can therefore generalize that future morphemes, in Greek and Italian at least, are modals of necessity in both uses. As such, they expresses commitment within the modal domain (*ReasFut* and best worlds), but it remains unclear, at the time of the utterance, whether the actual world

is, or is not, in the domain. This renders FUT in both uses non veridical, and explains why assertions with the future pattern with modal assertions (also non veridical) in not positively asserting *p* and in licensing polarity and free choice items ([7], e.g. *Ariadne will talk to anybody, Ariadne may talk to anybody*, etc).

References

- [1] Pier Marco Bertinetto. Alcune ipotesi sul nostro futuro (con osservazioni su "potere" e "dovere"). *Rivista di Grammatica Generativa*, 4(1-2):77–138, 1979.
- [2] Hans Broekhuis and Henk J Verkuyl. Binary tense and modality. *Natural Language and Linguistic Theory*, 2013.
- [3] Cleo Condoravdi. Temporal interpretation of modals. *The construction of meaning*, pages 59–88, 2002.
- [4] Bridget Copley. *The semantics of the future*. PhD MIT, 2002.
- [5] David R Dowty. *Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ*, volume 7. Springer, 1979.
- [6] Thomas Ernst. Speaker-oriented adverbs. *Natural Language and Linguistic Theory*, 27(3):497–544, 2009.
- [7] Anastasia Giannakidou. Affective dependencies. *Linguistics and Philosophy*, 22(4):367–421, 1999.
- [8] Anastasia Giannakidou. The dependency of the subjunctive revisited: Temporal semantics and polarity. *Lingua*, 119(12):1883–1908, 2009.
- [9] Anastasia Giannakidou. The greek future as an epistemic modal. In *Proc. ICGL 10*, 2012.
- [10] Anastasia Giannakidou and Alda Mari. Epistemic weakening with future and must: non veridicality, evidentiality, and partial knowledge. *Ms. University of Chicago and Institut Jean Nicod*, 2013. <http://ling.auf.net/lingbuzz/001836>.
- [11] Anastasia Giannakidou and Alda Mari. The future of greek and italian: an epistemic analysis. In Emmanuel Chemla, Vincent Homer, and Grégoire Winterstein, editors, *Proceedings of Sinn und Bedeutung 17*, pages 255–270, Paris, 2013.
- [12] Janneke Huitink. Modal concord: a case study of dutch. *Journal of semantics*, 29(3):403–437, 2012.
- [13] Stefan Kaufmann. Conditional truth and future reference. *Journal of Semantics*, 22(3):231–280, 2005.
- [14] Mikhail Kissine. From predictions to promises: how to derive deontic commitment. *Pragmatics & cognition*, 16(3):471–491, 2008.
- [15] Peter Klecha. Diagnosing modality in predictive expressions. *Journal of Semantics*, forthcoming.
- [16] Alda Mari. Disambiguating the italian future. *Proceedings of Generative Lexicon*, pages 209–216, 2009.
- [17] Alda Mari. Future, judges and normalcy conditions. *Ms. Institut Jean Nicod*, 2009. <http://hal.archives-ouvertes.fr/docs/00/35/44/62/PDF>.
- [18] Alda Mari. Each other, asymmetry and reasonable futures. *Journal of Semantics*, 2013.
- [19] Alda Mari. On the modal meaning of italian future tense. *Lingvisticae Investigationes*, 2014.
- [20] Richmond H Thomason. Combinations of tense and modality. In *Handbook of philosophical logic*, pages 135–165. Springer, 1984.

No ordered arguments needed for nouns*

Scott Grimm and Louise McNally

Universitat Pompeu Fabra
`{scott.grimm,louise.mcnamara}@upf.edu`

Abstract

Syntacticians have widely assumed since [11] that there is a fundamental difference between so-called argument structure nominals (AS-nominals, also called Complex Event Nominals), e.g. *destruction*, and non-AS-nominals, e.g. *book* ([1, 5], i.a.). Grimshaw provided a list of properties characterizing AS-nominals, most notably that they have obligatory arguments (e.g. *the destruction *(of Carthage) by the Romans*). She and others have associated having argument structure with having event structure, but it has never been clear what having or lacking such structures amounts to semantically. In this paper we present extensive corpus evidence that AS-nominals do not in fact exist as a distinct class. This result, we argue, removes an important challenge to [9]'s hypothesis that eventuality-denoting nouns systematically lack an ordered-argument semantics.

1 Introduction

Syntacticians have widely assumed since [11] that there is a fundamental difference between so-called argument structure nominals (hereafter, AS-nominals, also called Complex Event Nominals), e.g. nominalizations like *destruction*, and non-AS-nominals like *book* or *trip* ([1, 5], i.a.; see also [3, 12]). Grimshaw provided a list of properties characterizing AS-nominals, most notably that they have obligatory arguments (e.g. *the destruction *(of Carthage) by the Romans*). She and others have associated having argument structure with having event structure; this, in turn, has been crucial to motivating a fundamental grammatical distinction between *Complex Event Nominals*, on the one hand, and *Simple Event Nominals* and *Result Nominals*, on the other. However, it has never been clear in this literature what having or lacking AS or event structure amounts to semantically. In this paper we address precisely this issue. We begin by presenting extensive corpus evidence that AS-nominals do not in fact exist as a distinct class. This result, we argue, removes an important challenge to [9]'s hypothesis that eventuality-denoting nouns systematically lack an ordered-argument semantics.

[9], in a context outside the nominalization literature, distinguished two general sorts of analyses for any predicate describing an eventuality: (i) as an ordered n -tuple of individuals, where the participants must be fed to the predicate by function application in order to form a proposition, following classical Montague semantics (e.g. (1)); vs. (ii) as a neo-Davidsonian 1-place property of eventualities (e.g. (2-a)) with the participants added effectively as adjuncts by conjunction, mediated by thematic role predicates (e.g. (2-b); see also [6, 18]).

- (1) $\lambda y \lambda x \lambda e [\mathbf{destruction}(e, x, y)]$
- (2) a. $\lambda e [\mathbf{destruction}(e)]$ b. $\lambda e [\mathbf{destruction}(e) \wedge \mathbf{Ag}(e, \mathbf{romans}) \wedge \mathbf{Th}(e, \mathbf{carthage})]$

*We thank the anonymous reviewers for the Colloquium for comments. This research has been supported by Spanish Ministry of Education and Science grant FFI2010-15006 and an ICREA Foundation Academia award.

Dowty hypothesized that verbs have the former sort of semantics, while nouns have the latter, elaborating little on why this should be. However, he acknowledged (1989:90-91) that the existence of some cases of nominals with apparently obligatory PPs, AS-nominals in different terminology, posed a challenge for this hypothesis.

Though the intuitions supporting the existence of AS-nominals are initially convincing, counterexamples to the claims concerning their diagnostics are well known (see [5, 17, 14] references cited therein). We therefore decided to test the strength of these diagnostics on a broad empirical level. Our results show that (i) none of the putative properties in fact discriminate between nominalizations with PPs expressing participants and those without, and (ii) the presence of such PPs with a nominalization is instead predictable based on the discourse context. We argue that the data can be explained with a single neo-Davidsonian semantics for nominals plus a general pragmatic principle that requires all token eventualities (and certain subeventualities) to be anchored to at least one token discourse referent on first mention, in the spirit of the Argument-Per-Subevent condition of [21].

2 An Empirical Study of the Diagnostics of AS-Nominals

[11] argued nominalizations are ambiguous between different readings, in turn influencing the presence of arguments. These different readings are exemplified in (3), reproduced from [2]. *Examination* has a “complex event interpretation” (involving participants) in (3-a), while *examination* in (3-b) is argued to refer to an event without an articulated event structure, and therefore no arguments. In (3-c), *examination* refers to a physical object, again not possessing event structure, thereby lacking arguments.

- (3) a. The examination of the patients took a long time. (AS-nominal)
b. The examination took a long time. (Simple Event Nominal)
c. The examination was on the table. (Result Nominal)

[11] argued that a cluster of properties distinguished AS-nominals from non-AS-nominals, whether Simple Event Nominals or Result Nominals. AS-nominals 1) have obligatory arguments, 2) denote eventualities; 3) take agent-oriented modifiers such as *deliberate* or *intentional* which are banned for non-AS-nominals; 4) take Saxon genitives and *by*-phrases as arguments, not adjuncts; 5) allow implicit argument control; 6) take aspectual modifiers such as *in an hour* or *for an hour*; 7) allow modifiers such as *frequent* or *constant* in the singular, while non-AS-nominals require the plural; and 8) must be singular, while non-AS-nominals may be plural. We conducted an extensive corpus analysis in order to examine whether the ambiguity in nominalizations hypothesized by Grimshaw and others was observable. In particular, we examined properties 3), 6), 7), and 8), which are easily observable with corpus techniques.

Using the information on derivational morphology available in the CELEX database [4], we selected 150 nouns derived by *-(at)ion* or *-ment*, including all *-ment* or *-ation* nouns in the database as well as a random selection of frequent nouns derived by *-ion* (e.g. *destruction*). For each noun, we extracted up to 1000 random occurrences, or as many as present if less than 1000, from the COCA corpus (<http://corpus/byu.edu/coca/>).

We then processed the data in several steps, using a dependency parser and then extracting the information thereby obtained. First, we developed a Python script to standardize the search results from COCA, extracting the sentence where the search term occurred and discarding any surrounding fragments in the search result. These sentences were then passed to the Stanford Dependency Parser [7], which returned both a parse of the sentence and a list of

the dependencies. Another Python script used the information coded in the dependencies to automatically extract information relevant to the properties put forth by Grimshaw, including whether arguments were present (*of*-phrase, *by*-phrase, etc.), number (singular/plural), the presence of adjectival modifiers and any further PP modifiers. At this stage, instances which appeared as a part of a compound, as in *examination criteria*, were excluded as they clearly were not of interest for the hypotheses under investigation. Finally, an R script was developed which constructed a large data frame in which all the information for each nominalization was collected, tallied and submitted to the appropriate statistical tests where relevant.

Overall, the procedure was highly accurate for the questions we address. Although [7] estimate that the Stanford Dependency Parser has a per-dependency error rate of 80.3% correct, as we were only examining a subset of the dependencies, which were mostly very local, the number of errors that were relevant for our study were far fewer. An error analysis on 400 sentences resulted in an effective error rate of 97% correct. We also examined across 5 nouns how reliably the *of*-phrase extracted by this procedure corresponded to a true argumental *of*-phrase. This varied greatly according the meaning of the noun and potential for polysemy. Some nouns had nearly exclusively argumental *of*-phrases (*abandonment* 99%), while others had a minority (*adornment* 25%). As hand-correcting all of the occurrences is not currently feasible, we simply checked the data points which were relevant to the properties investigated to ensure reliability, e.g. the *of*-phrases for each occurrence where *deliberate* appeared as an adjectival modifier. We now discuss the results with respect to each of the properties.

Number: The overwhelming majority of deverbal nouns had nearly all of their occurrences in the singular, whether an *of*-phrase was present or not. While [11] claims that all AS-nominals only occur in the singular, 20% of AS-nominals in our data set occurred at least once in the plural. More generally, number does not distinguish between AS- and non-AS-nominals, since whether the nominal had an *of*-phrase or not had no effect on its occurrence in the plural. Nouns with high rates of singular (plural) occurrences when PPs were present also had high rates of singular (plural) occurrences when no PPs were present: e.g., 98% of the occurrences of *allegation*, both with and without an *of*-phrase, were singular, while for *observation*, 46% of the occurrences, with and without an *of*-phrase, were singular. For each nominal, we ran Fischer's Exact Test comparing the number of occurrences in singular and in plural of the AS-nominal with the number of occurrences in singular and in plural of the non-AS-nominal. For nearly all the nominalizations considered, there was no significant difference between nominals with and without an *of*-phrase. The only exceptions were *application*, *assessment*, and *examination*—all of which have substantive uses in a non-eventive result reading.

Adjectival Modifiers: [11] and subsequent authors claimed that (i) *constant* and *frequent* should occur with singular forms of AS-nominals, but with plural forms of nominalizations without an *of*-phrase and that (ii) *deliberate* and *intentional* only occur with AS-nominals as they crucially refer to an agent. Table 1 displays the distribution of these adjectival modifiers observed in the data. Neither claim is substantiated: *constant* and *frequent* appear happily with singular forms of nominalizations without an *of*-phrase and *deliberate* and *intentional* occur equally in nominalizations with and without an *of*-phrase. The behavior of these modifiers then, rather than distinguishing between AS-nominals and nominals without *of*-phrases, indicates that the nominalization behaves similarly regardless of the presence of an *of*-phrase.

Aspectual Modifiers: Aspectual modifiers of the type *for an hour/in an hour* were extremely rare. Out of 3363 data points which had *for-* or *in-*phrases, only 3 were aspectual modifiers, all of which were *for*-phrases and occurred with non-AS-nominals, contrary to the claim of [11]. An example is given in (4):

- (4) Rosenblatt and co-authors (1987) also reported that at least 40 mg/L ClOsub2 **gas**

Constant	Sing	Plur	Frequent	Sing	Plur
AS	16	0	AS	15	0
Non-AS	46	0	Non-AS	10	0
Deliberate	Sing	Plur	Intentional	Sing	Plur
AS	8	0	AS	6	0
Non-AS	21	0	Non-AS	8	0

Table 1: Distribution of Adjectival Modifiers *constant*, *frequent*, *deliberate* and *intentional*

treatment for one hour at 60 percent RH effectively sterilized Whatman 3-mm filter paper strips ... (COCA)

Obligatoriness of Arguments We then further tested some of the strongest evidence that has been used to argue for AS-nominals: (i) the *of*-phrase is obligatory when a *by*-phrase is present ([13]), and (ii) certain nominalizations, such as *sending* or *handing* appear to be banned without an *of*-phrase ([20]), e.g. *the sending* *(*of the letters*). We found that, in the right context, neither generalization holds, as the attested examples in (5) and (6) show. Interestingly, this behavior is not unique to deverbal or even deadjectival nominals (e.g. *awareness*): Even those derived from nouns, such as *friendship*, show similar behavior (see (7)).

- (5) How does a country recover from 40 years of **destruction by an unchallenged tyrant?** (Newsweek)
- (6) While the originator—or his or her computerized agent—purposefully sent the information item into Jurisdiction B, **the sending** occurred instantaneously in response to the address supplied by the client, without any advance arrangement by the content originator. (http://www.kentlaw.edu/perritt/publications/41%20VILL._L._REV._1.htm)
- (7)
 - a. ??Bob’s friendship vs. ^{ok}Bob’s friendship with Sue
 - b. Sue has known Bob for years. Bob’s friendship means the world to her.

To conclude this study, the properties proposed by Grimshaw and others do not accord with the data that we examined. While the nominalization literature asserts a fundamental ambiguity between AS-nominals and eventive nominalizations without arguments (Simple Event Nominals), the examples in (6)–(7) indicate a different possibility: the optionaliy of PPs with a nominalization may be simply due to the overall discourse context, rather than to an inherent ambiguity between AS-nominals and non-AS-nominals.

3 The Discourse Behavior of AS-Nominals

We conducted a second study using the Brown Corpus to investigate the general relation between the presence of PPs and discourse context. The investigation centered on two related questions. First, if the nominalization is eventive, and no argument is present, is the argument recoverable? Second, what is the information status of the discourse referent specified by the *of*-phrase or that of the recovered argument? If the lack of an *of*-phrase corresponds with an interpretation where no internal argument is specifiable, this is strong evidence for an ambiguity account. If, however, the lack of an *of*-phrase corresponds with internal arguments that are recoverable from the discourse context, that supports a view upon which nominalizations simply do not need to specify the internal argument if its referent is sufficiently salient.

We used the Brown corpus for this study since the entire discourse context, e.g. an entire news article, is available to examine. We automatically searched for instances of a large subset of the *-ation* and *-ion* nominals from our first corpus study, extracting their occurrences along with the entire context. This resulted in 867 occurrences of 37 different noun types.

We ensured that we were comparing like interpretations of the nominalizations, namely eventive and full noun phrase uses of nominalizations which potentially would have an internal argument, by hand-checking each occurrence. We excluded all instances which referred to physical objects, i.e. “result nominal” interpretations (e.g. *All of the jackets carry a fairly technical and detailed explanation of this new recording program.*). We also excluded instances which were not full noun phrases, but occurred in light-verb constructions (*It would allow presentation to the public of a unified approach.*) or occurred bare with prepositions (*under occupation*). We also set aside nominals which selected for a different preposition than *of*, e.g. *invitation for*. Finally, we excluded nominals occurring as part of a compound (*a German recitation*) or with possessive pronouns which would specify the internal argument. 259 occurrences remained.

For the cases in which no *of*-phrase was present, to determine whether an internal argument was recoverable, we considered whether the occurrence was paraphrasable with an *of*-phrase. For instance, in the example in (8), *interpretation* occurs without an *of*-phrase, but it is clear that the interpretation is of something, namely ‘the information’ which is already present as the subject of the sentence, rather than a “simple event” of interpreting.

- (8) This information was accepted with **the frequent interpretation** that those persons who did not show arm-levitation must be preventing it.

In general, when no *of*-phrase was present, the internal argument was nearly always recoverable. Only in 4 instances did there seem to be a true “Simple Event” usage, and these were essentially limited to the noun *presentation*, e.g. *The presentation was made before several hundred persons at the annual meeting . . .*

We annotated the information status of the *of*-complement, if present, or if not present but recoverable, that of the recovered discourse referent. We followed the coding scheme in [8] which has three primary information structure categories: *given* if the discourse referent was previously mentioned, *accessible* if the discourse referent is generally known or inferable in some manner from the discourse, and *new* if the discourse referent is not previously mentioned or otherwise inferable. *Given* has two subcategories: *active*, when the discourse referent is present within the current or immediately preceding sentence, and *inactive*, if mentioned earlier than the preceding sentence. (*Accessible* also has subcategories which were coded for but are not relevant in what follows.)

If the *of*-complement was not present but the internal argument was recoverable, in 82% of the cases it was *given* and in 18% of the cases it was *accessible*. Most all occurrences classified as *given* were *given-active* (91%). Thus, when an eventive nominalization occurs without a PP, the content of the missing PP is nearly always recoverable and mostly very salient in the discourse. In contrast, for those nominalizations where the *of*-phrase was present, 47% of the referents in the *of*-phrase were classified as *new*. Of the remaining cases, 34% were *given* and 19% *accessible*. Those that were given were almost even split between *given-active* (51%) and *given-inactive* (49%). Thus, when an eventive nominalization occurs with a PP, the discourse referent is nearly half of the time new, and not frequently mentioned in the immediately preceding discourse (17% of the time). In summary, the presence or absence of a PP does not generally correspond to AS-nominal or Simple Event readings, respectively, although there may certainly be uses of certain nominalizations corresponding to a Simple Event Nominal, such as *presentation*. Instead, the driving force behind the occurrence of PPs appears to be the discourse salience of its referent.

4 Analysis

Based on the data, we see no argument for a distinct category of AS-nominals whose arguments are obligatorily expressed. Rather, the data indicate that nominalizations are simply relation-*entailing* nouns, as represented in (2b); the participant variables are free and can be contextually valued (see [22] for such a proposal). To this we add the condition on reference to a particular token eventuality in (9), which is a semantic counterpart of the Argument-Per-Subevent condition of [21], (p. 779) in (10):

- (9) **Event Instantiation Condition:** The introduction of a token discourse referent for an eventuality e requires that e be anchored to a discourse referent corresponding to at least one of its participants.
- (10) **The Argument-Per-Subevent Condition:** There must be at least one argument XP in the syntax per subevent in the event structure.

The crucial restriction to the *introduction* of a *token* eventuality is not a simple stipulation but rather is grounded in the deeper hypothesis that token eventualities can only be properly identified and apprehended if we identify at least one of their participants, whereas our identification and apprehension of *types* of eventualities involves other cognitive mechanisms for which the identification of participants is not essential (see [15] for related discussion).

We see various advantages to an account based on (9). First, it predicts the possibility of variability in the presence vs. absence of participant PPs. If we appealed to the Argument-Per-Subevent condition, we would predict no such variation because the subevent structure associated with a word should be the same across all of its uses. Second, (9) squares with the observation that AS-nominals tend to denote (unique) token eventualities, while non-AS-nominals are not so restricted. The Argument-Per-Subevent condition also fails to make this prediction because there is no reason to think that the event structure for kind- vs. token-level eventualities should differ.

Third, we predict a clear correlation between the anaphoric use of nominals and the omissibility of arguments. If a discourse referent for an eventuality is already introduced, anaphoric reference to it will not require re-anchoring it via some participant. In contrast, if the nominal is being used to introduce a token eventuality referent for the first time, as we would argue is the way in which the classic examples from the literature are most naturally interpreted, the participant anchor for the eventuality must be specified, either via a PP or in some cases via a Saxon genitive.

We can exemplify the process of specification via PPs with *of*-phrases. Following [19] i.a., we take (genitive) *of*-phrases to introduce an underspecified relation, e.g. $\lambda x[R(x, \text{carthage})]$, x of the entity or event sort. The value for this relation will be picked up from the relational noun. As is well known, *of*-phrases with nominalizations are ambiguous, for example between agent and theme interpretations. Conjoining a representation for nominalizations of the sort given in (2) with that for the *of*-phrase results in specifying either the agent or theme, depending on which role relation R picks up on. More “contentful” PPs will work analogously, but contribute additional lexical entailments. For example, *by*-phrases specify that the participant is the initiator of the event.

Fourth, an analysis grounded in (9) immediately carries over to deadjectival and denominational forms, since nothing in it hinges on verb-specific properties. Analyses grounded in event structure do not obviously predict this similarity. Fifth, it allows us to eliminate the otherwise unmotivated simple vs. complex event nominal distinction from the grammar. Finally, though space limitations preclude extensive discussion here, it avoids the difficulties faced by alter-

native accounts on which the noun's arguments are claimed to be present but simply elided or expressed by null pronominals or similar elements. The non-expression of the material in question simply does not appear to be sensitive to the structural restrictions that are typical of ellipsis in either nominals or VPs (the latter observation, by the way, further calls into question the use of verbal structure in the analysis of these nominals). An analysis in terms of null pronominal arguments is complicated by the need to deal with material corresponding to the missing preposition(s) and by the difficulties of integrating such an analysis into a theory of null pronominals in English more generally. On our account, there is no need to fit the facts into either the theory of ellipsis or null pronominals.

Rescuing Dowty's claim that nouns lack an ordered-argument semantics leaves us with the question of why nouns and verbs would differ in this regard, as he suggested. In fact, it would not surprise us if, contrary to what Dowty argues, [6] was right in defending a neo-Davidsonian analysis for verbs as well. Dowty based his claim concerning the differences between verbs and nouns on the fact that seemingly truth-conditionally equivalent verbs (e.g. those in (11)) differ in the obligatoriness of their arguments.

- (11) a. John ate (the meal). b. John devoured *(the meal).

However, these asymmetries might be explainable without appealing to an ordered-argument semantics.¹ The first crucial observation is that nonpronominal verb phrases, unlike noun phrases, are arguably never anaphoric, and they are almost always used for token reference. Neither characteristic applies to nominals. The participant realization properties of nouns and verbs will thus be systematically different, but we need not conclude that this is accompanied by the sort of distinction enshrined in the AS-nominal literature.

As for the contrast in (11), there are at least two possible explanations for it that do not depend on ordered arguments. One could appeal to the well-known aspectual and distributional differences between *eat* and *devour* to argue that *eat* is an activity predicate with no interesting subevent structure, while *devour* is necessarily a change of state predicate with activity and result state subevents (see e.g. the references cited in [16]). Alternatively, we could follow [10], who shows that *eat* and *devour* differ in that the latter occurs with a much wider variety of arguments; their presence may therefore be necessary to identify exactly what sort of eventuality is being described. Either way, the strong tendency for the object argument in (11-b) but not (11-a) could follow directly.

It remains to replicate this study with -*ing* nominals to see whether the observations reported here extend to them as well. For the time being, however, we tentatively conclude that no ordered arguments are needed for nouns.

References

- [1] Artemis Alexiadou. *Functional structure in nominals*. John Benjamins, Amsterdam, 2001.
- [2] Artemis Alexiadou and Jane Grimshaw. Verbs, nouns and affixation. In F. Schäfer, editor, *SinSpeC (1): Working Papers of the SFB 732, University of Stuttgart*, pages 1–16, 2008.
- [3] Mona Anderson. Prenominal genitive NPs. *The Linguistic Review*, 3:1–24, 1983.
- [4] R. Harald Baayen, Richard Piepenbrock, and H. van Rijn. The CELEX lexical data base on CD-ROM. 1993.

¹We are optimistic that the account we sketch here for verbs can also address a concern raised by an anonymous reviewer as to whether (10) by itself is enough to predict all of the facts.

- [5] Hagit Borer. Exo-skeletal vs. endo-skeletal explanations. In John Moore and Maria Polinsky, editors, *The Nature of Explanation in Linguistic Theory*, pages 31–67. CSLI Publications, Stanford, CA, 2003.
- [6] Gregory Carlson. On the role of thematic roles in linguistic theory. *Linguistics*, 22:259–279, 1984.
- [7] Marie-Catherine de Marneffe, Bill MacCartney, and Christopher D. Manning. Generating typed dependency parses from phrase structure trees. In *LREC*, 2006.
- [8] Stefanie Dipper, Michael Götze, and Stavros Skopeteas, editors. *Information Structure in Cross-Linguistic Corpora: Annotation Guidelines for Phonology, Morphology, Syntax, Semantics and Information Structure*. Universitätsverlag Potsdam, 2007.
- [9] David Dowty. On the semantic content of the notion ‘thematic role’. In Gennaro Chierchia, Barbara Partee, and Raymond Turner, editors, *Properties, Types, and Meaning, II*, pages 69–129. Kluwer, Dordrecht, 1989.
- [10] Lelia Glass. If looks could kill: Honing a pragmatic account of English implicit objects. Talk presented at the 37th Penn Linguistics Colloquium, 2013.
- [11] Jane Grimshaw. *Argument Structure*. MIT Press, 1990.
- [12] James Higginbotham. Logical form, binding, and nominals. *Linguistic Inquiry*, 14:679–708, 1983.
- [13] N. Hornstein. S and X’ convention. *Linguistic Analysis*, 3:137–176, 1977.
- [14] Andrew McIntyre. Argument linking and complex predicates in English nominalizations. Talk presented at JeNom5, UPF, 2013.
- [15] Louise McNally, Gemma Boleda, and Marco Baroni. Conceptual vs. referential affordance in concept composition. Talk presented at the Workshop on Concept Composition and Experimental Semantics/Pragmatics, U. Utrecht, 2013.
- [16] Åshild Næss. The grammar of eating and drinking verbs. *Language and Linguistic Compass*, 4/5:413–423, 2011.
- [17] Frederick Newmeyer. Current challenges to the lexicalist hypothesis. In William D. Lewis, Simin Karimi, Heidi Harley, and Scott O. Farrer, editors, *Time and Again: Theoretical Perspectives on Formal Linguistics*, pages 91–117. John Benjamins, Amsterdam, 2009.
- [18] Terence Parsons. *Events in the Semantics of English: A Study in Subatomic Semantics*. Current Studies in Linguistics Series 19. MIT Press, Cambridge, MA, 1990.
- [19] Barbara Partee. Uniformity vs. versatility: the genitive, a case study. In Johan van Benthem and Alice ter Meulen, editors, *The Handbook of Logic and Language*, pages 464–470. Elsevier, Amsterdam, 1983/1997.
- [20] Malka Rappaport. On the nature of derived nominals. In B. Levin, M. Rappaport, and A. Zaenen, editors, *Papers in Lexical-Functional Grammar*, pages 113–142. Indiana University Linguistics Club, 1983.
- [21] Malka Rappaport Hovav and Beth Levin. An event structure account of English resultatives. *Language*, 77:766–797, 2001.
- [22] Alessandro Zucchi. *The Language of Propositions and Events*. Kluwer, Dordrecht, 1993.

The Intervention effect*

Focus alternatives or indefinite alternatives?

Experimental evidence

Andreas Haida and Sophie Repp

Humboldt University, Berlin, Germany
{andreas.haida|sophie.repp}@hu-berlin.de

Abstract

In many languages, quantificational and focusing elements may not intervene between a *wh*-phrase and the interrogative complementizer by which it is licensed. Among the semantic approaches to explaining this phenomenon, there are two that make different predictions as to the intervener status of focus particles. According to the first approach, all focus particles are interveners. According to the second approach, *only* and *even* are interveners while *also* is not. We present evidence from two speeded-acceptability judgement experiments and a self-paced reading experiment in German which tested the acceptability and online processing of intervention sentences with *only* and *also* with regard to the predictions of the two theories. The results of the three experiments converge in showing that *also* is not an intervener in German. We argue that this can be taken as evidence for theories where focus is not the key property in the emergence of the intervention effect, at least for German.

1 Introduction: the intervention effect

In many languages, we find paradigms such as (1) and (2), see [2].

- | | | | |
|-----|---|-----|--|
| (1) | a. *Minsu-man nuku-lül po-ass-ni?
Minsu-only who-ACC see-PAST-Q
<i>exp:</i> ‘Who did only Minsu see?’ | (2) | a. *Wer hat niemand wo gesehen?
who has no one where seen
<i>exp:</i> ‘Where did no one see whom?’ |
| | b. Minsu-nun nuku-lül po-ass-ni?
Minsu-TOP who-ACC see-PAST-Q
‘Who did Minsu see?’ | | b. Wer hat Luise wo gesehen?
who has Luise where seen
‘Where did Luise see whom?’ |

The Korean data in (1) show that the focus particle *-man* ‘only’ may not precede the *wh*-word *nuku-lül* ‘who-ACC’. The German data in (2) show the same phenomenon with the negative quantifier *niemand* ‘nobody’ relative to the *wh*-word *wo* ‘where’.¹ Observations like these have led to the generalization in (3), cf. [7, 2].

- (3) *[Q_i [. . . [intervener [. . . wh-phrase_i . . .] . . .] . . .]]

Certain elements may not intervene between a *wh*-phrase and the interrogative complementizer by which it is licensed.

*This research was supported by the German Research Council (DFG) as part of the SFB 632 *Information Structure*. Many thanks to M. Krifka, R. Šimk, B. Surányi, and M. Zimmermann for helpful discussion.

¹Since German is a *wh*-ex-situ language, the phenomenon under consideration can only be evoked with multiple *wh*-questions. When discussing the different theories, we only consider structures with a single *wh*-phrase. However, the results readily translate to multiple *wh*-questions.

The emergence of deviance in (3) is called the intervention effect.² The semantic approaches to explaining the intervention effect share a line of thinking going back to [6]: *wh*-words induce semantic alternatives, a proper question denotation can be derived only if these alternatives are visible to the question operator, and an intervener renders the alternatives invisible to the question operator. The approaches differ with regard to the source of the alternatives: (i) *Wh*-words induce alternatives by being focused, see [2]. (ii) *Wh*-words induce alternatives by being indefinites, see [3, 5, 8]. Correspondingly, the list of interveners proposed by Beck can be divided into two classes: (i) focusing elements such as *only*, *even*, and *also*, (ii) the sentence negation *not*, nominal quantifiers such as (*almost*) *every*, *most* and *few*, and adverbial quantifiers such as *always*, *often* and *never*. Furthermore, the approaches differ with regard to the way the deviance is explained, as will be shown for [2] and [5] in section 2.

2 The theories

BECK'S FOCUS APPROACH in [2] is based on Rooth's focus semantics [10].³ Every structure has two semantic values, viz. an ordinary semantic value ('o-value' for short) and a focus semantic value ('f-value'). These values are given by the (partial) interpretation functions $\llbracket \cdot \rrbracket^o$ and $\llbracket \cdot \rrbracket^f$, respectively. The way Beck derives the truth conditions of the LF-structure of the sentence *Only John left* in (4) is completely standard. The o-value of a syntactic atom is its usual denotation, see (5a) and (6a). The f-value of the F-marked proper name $John_F$ is a set of individual alternatives, see (5b). The f-value of a non-F-marked atom is the singleton set of its o-value, see (6b). The semantic values of the phrase α are derived by (pointwise) function application, see (7).

$$(4) \quad [\text{only } C [\sim C [\alpha \text{ } John_F \text{ } left]]]$$

$$(5) \quad \begin{array}{ll} \text{a. } \llbracket John_F \rrbracket^o = J & \text{b. } \llbracket John_F \rrbracket^f = \{J, M, \dots\} \end{array}$$

$$(6) \quad \begin{array}{ll} \text{a. } \llbracket left \rrbracket^o = \lambda x \lambda w. x \text{ left in } w & \text{b. } \llbracket left \rrbracket^f = \{\lambda x \lambda w. x \text{ left in } w\} \end{array}$$

$$(7) \quad \begin{array}{ll} \text{a. } \llbracket \alpha \rrbracket^o = \lambda w. J \text{ left in } w & \text{b. } \llbracket \alpha \rrbracket^f = \{ \lambda w. J \text{ left in } w, \lambda w. M \text{ left in } w, \dots \} \end{array}$$

Next, the \sim operator passes on the o-value of its complement, see (8a), it evaluates the focus anaphor C with a subset of the f-value of its complement (by means of a definedness condition), see (8a,b), and it resets the f-value to the singleton set of the o-value, see (8b).

$$(8) \quad \begin{array}{ll} \text{a. } \llbracket [\sim C \alpha] \rrbracket_g^o = \lambda w. J \text{ left in } w, \text{ if } g(C) \subseteq \llbracket \alpha \rrbracket^f, \text{ undefined otherwise} & \\ \text{b. } \llbracket [\sim C \alpha] \rrbracket_g^f = \{ \lambda w. J \text{ left in } w \}, \text{ if } \dots \text{ (same as above)} & \end{array}$$

The reset is the root cause of the intervention effect in Beck's account: the \sim operator renders the focus alternatives invisible to operators outside its c-command domain. Outside this domain, they can only be accessed via the focus anaphor C. Of course, *only* comes with C so that it can express truth conditions relative to the focus alternatives induced by $John_F$, see (9).

$$(9) \quad \llbracket (4) \rrbracket_g^o = \lambda w. \forall p \in g(C) : p(w) = 1 \rightarrow p = [\lambda w'. J \text{ left in } w'], \text{ if } g(C) \subseteq \llbracket [John_F \text{ left}] \rrbracket^f$$

The answer set of the *wh*-question *Who left?*, i.e. the o-value of (10), is derived as follows.⁴

$$(10) \quad [Q [\alpha \text{ } who_F \text{ } left]]$$

²Intervention effects also arise in other environments than *wh*-questions, see [11].

³Beck uses syntactic variables for deriving focus semantic values but for convenience we specify the usual alternative sets. Alternative sets can be readily derived from Beck's interpretation function $\llbracket \cdot \rrbracket^{g,h}$. The alternative set of a structure $\alpha = \{ \llbracket \alpha \rrbracket^{g,h} \mid h' \text{ is a total f-variable assignment function in } \mathcal{D}_e \}$, where g is an o-variable assignment function.

⁴In (10), we ignore the coindexation of Q and the *wh*-phrase. This simplification leaves Beck's analysis intact.

Beck assumes that *wh*-words are inherently focused (i.e. F-marked). Thus who_F induces a set of individual alternatives, see (11a), and the f-value of α is a proposition set which is the Hamblin denotation of the question under consideration, see (11b).

$$(11) \quad \text{a. } \llbracket \text{who}_F \rrbracket^f = \{J, M, \dots\} \quad \text{b. } \llbracket \alpha \rrbracket^f = \{\lambda w. J \text{ left in } w, \lambda w. M \text{ left in } w, \dots\}$$

The Q operator promotes the f-value of its complement to the o-value of the overall structure:

$$(12) \quad \llbracket [Q \alpha] \rrbracket^o = \llbracket \alpha \rrbracket^f = \{\lambda w. J \text{ left in } w, \lambda w. M \text{ left in } w, \dots\}$$

On these assumptions, the deviant string **Only John saw who?* (Korean with English lexemes) cannot be assigned a *wh*-question meaning. Consider the LF of this string in (13).⁵

$$(13) \quad [Q [\gamma \text{ only } C [\beta \sim C [\alpha \text{ John}_F \text{ saw who}_F]]]]$$

The f-value of α still captures the focus alternatives induced by who_F , see (14). But since the \sim operator resets the f-value of β to the singleton set of the o-value of α , see (15), the focus alternatives induced by who_F are invisible to Q.

$$(14) \quad \llbracket [\alpha \text{ John}_F \text{ saw who}_F] \rrbracket^f = \{\lambda w. J \text{ saw M in } w, \lambda w. M \text{ saw J in } w, \dots\}$$

$$(15) \quad \llbracket [\beta \sim C [\alpha \text{ John}_F \text{ saw who}_F]] \rrbracket^f = \{\llbracket \alpha \rrbracket^o\}, \text{ if } g(C) \subseteq \llbracket \alpha \rrbracket^f, \text{ undefined otherwise}$$

This means that even if the o-value of (13) was defined (it is not, see below), it would not be the answer set of a *wh*-question. Thus, in Beck's account the distribution of the \sim operator explains the distribution of the intervention effect.

So far, (13) seems to have an o-value (albeit not the o-value of a *wh*-question). To rule out any meaning assignment, Beck stipulates that *wh*-words do not have an o-value:

$$(16) \quad \llbracket \text{who} \rrbracket^o \text{ is undefined}$$

(16) does not do harm to 'normal' *wh*-questions but it crashes the meaning assignment to structures like (13): Because of (16), the o-value of α is undefined and hence the f-value of β is too, see (15). But then the f-value of γ is undefined, too, and so is the o-value of the overall structure. This explains the deviance of the string **Only John saw who?*.

HAIDA'S INDEFINITES APPROACH in [5] is a dynamic semantic analysis. Dynamic semantic frameworks serve to explain anaphoric binding such as the relation indicated in (17).⁶

$$(17) \quad \text{Someone}_i \text{ left. He}_i \text{ closed the door.}$$

Following [9], Haida assumes that *[someone₁ left]* denotes (18).

$$(18) \quad [\lambda k \lambda k' \lambda w. k \text{ and } k' \text{ differ at most in the content of } r_1, \\ \text{and in } k' \text{ the individual in } r_1 \text{ left in } w]$$

(18) is the intension of a relation between input contexts k and output contexts k' . Contexts can be thought of as inventories of the content of an array of registers $\langle r_1, r_2, \dots \rangle$. An indefinite deposits an individual in a register (thereby pushing out any old content), i.e., it has context change potential. The condition in the first line of (18) states that the context change potential of *someone₁* is limited to r_1 : *someone₁* deposits individuals in r_1 and in no other register (and the

⁵Recall that in Beck's analysis f-values are derived by means of (indexed) variables. This means that the question operator can selectively bind F-marked *wh*-phrases (with which Q is coindexed) and ignore F-marked non-*wh*-phrases. The simplified structure in (13) is not rich enough to enable selective binding. Still, since Beck assumes that the \sim operator is an unselective binder our discussion still conveys the gist of her explanation.

⁶There are four reasons for a dynamic semantic analysis of *wh*-questions: 1. Questions words are anaphoric binders, see (i). 2. In many languages, unfocused *wh*-words are indefinites, see [5]. 3. Anaphoric binding is prone to intervention, see (21) below. 4. In the partition theory of questions, anaphoric binding potential can be seen as the defining semantic property of question words [4, 5].

(i) Who_i won the women's high jump? What height did she_i jump?

same for any other index). The condition in the second line states that any deposited individual is an individual that left in the evaluation world w . Any output context k' of $[\text{someone}_1 \text{ left}]$ can be passed on as input context for the interpretation of the subsequent sentence $[\text{he}_1 \text{ closed the door}]$. The pronoun he_1 looks up the content of register r_1 , and the sentence adds to the input-output relation in (18) the condition that in k' the individual in r_1 closed the door in w . This explains the possibility of anaphoric binding.

The input-output relation in (18) can be translated into a set of propositional alternatives. The function in (19) takes a sentence structure α and translates its context change potential with respect to the register r_i into an alternative set. The alternative set of $[\text{someone}_1 \text{ left}]$ for r_1 is derived in (20). It is the same set as the one specified by $[\text{who}_F \text{ left}]$ in Beck's approach.

$$(19) \quad \mathcal{A}lt_i(\alpha) = \{\llbracket \alpha \rrbracket(k)(k') \mid k \text{ and } k' \text{ differ at most in } r_i\}$$

$$\begin{aligned} (20) \quad \mathcal{A}lt_1([\text{someone}_1 \text{ left}]) &= \{\llbracket [\text{someone}_1 \text{ left}] \rrbracket(k)(k') \mid k \text{ and } k' \text{ differ at most in } r_1\} \\ &= \{\lambda w. \text{in } k' \text{ the individual in } r_1 \text{ left in } w \mid k' \text{ is a context}\} \\ &= \{\lambda w. x \text{ left in } w \mid x \text{ is an individual}\} \end{aligned}$$

Anaphoric binding is prone to intervention:⁷

- (21) a. John didn't consider buying a car._i. *It_i was too expensive.
- b. Most students considered buying a car._i. *It_i was very cheap.
- c. John often considered buying a car._i. *It_i was very cheap.

Interveners for anaphoric binding do not pass on the context change specified by indefinites in their scope, and this follows from their meaning. For example, *not* expresses that the context change specified by its complement is incompatible with the facts in the evaluation world. Therefore, the alternative set of $[\text{not} [\text{someone}_1 \text{ left}]]$ for r_1 is the singleton set of the proposition that no one left, see (22). This is the root cause of the intervention effect in the indefinites approach.

$$\begin{aligned} (22) \quad \mathcal{A}lt_1([\text{not} [\text{someone}_1 \text{ left}]]) &= \\ &= \{\lambda w. \neg \exists k'. \llbracket [\text{someone}_1 \text{ left}] \rrbracket(k)(k')(w) = 1 \mid k \text{ is a context}\} \\ &= \{\lambda w. \neg \exists k'. \text{in } k' \text{ the individual in } r_1 \text{ left in } w\} \\ &= \{\lambda w. \neg \exists x. x \text{ left in } w\} \end{aligned}$$

Haida assumes that *wh*-words are indefinites. Thus the question *Who left?* has the denotation in (23). Note that the question operator in (23) defines a partition from the alternative set specified by its complement.

$$\begin{aligned} (23) \quad \llbracket [\text{Q}_1 [\alpha \text{ who}_1 \text{ left}]] \rrbracket &= [\lambda w \lambda w'. \forall p \in \mathcal{A}lt_1(\alpha) : p(w) = 1 \leftrightarrow p(w') = 1] = \\ &= [\lambda w \lambda w'. \forall p \in \{\lambda w''. x \text{ left in } w'' \mid x \text{ is an individual}\} : p(w) = 1 \leftrightarrow p(w')] \end{aligned}$$

The string **Not John saw who?* (Hindi with English lexemes) has the denotation in (24), i.e., it is not assigned a *wh*-question meaning but the *yes/no*-question meaning 'Did John see anyone?'. According to [5] it is deviant because *wh*-questions come with an existence presupposition. So the string presupposes a complete answer to the question it expresses, see (25) (recall that *not* is a presupposition hole). That is, it defines a trivial partition of the logical space. This explains the deviance.

$$\begin{aligned} (24) \quad \llbracket [\text{Q}_1 [\alpha \text{ not} [\text{John saw who}_1]]] \rrbracket &= [\lambda w \lambda w'. \forall p \in \mathcal{A}lt_1(\alpha) : p(w) = 1 \leftrightarrow p(w') = 1] \\ &= [\lambda w \lambda w'. \exists x (\text{J saw } x \text{ in } w) \leftrightarrow \exists x (\text{J saw } x \text{ in } w')] \end{aligned}$$

- (25) *Not John saw who?
 derived meaning: 'Did John see anyone?' presupposition: 'John saw someone.'

⁷We discuss focus particles in the next section.

3 The intervener status of focus particles

The focus approach of [2] readily predicts that focus particles such as *only*, *even*, and *also* are interveners. For interveners like *not*, *most*, and *often*, [2] assumes that they always come with a \sim operator. This makes the prediction that focus association across these elements is impossible. In German, this prediction is wrong, see [5]. Therefore, we hold that the focus approach cannot be maintained for these elements in German.

The indefinites approach of [5] readily predicts that expressions like **not**, **most**, and **often**, which are interveners for anaphoric binding, are interveners in *wh*-questions. However, the approach does not predict that focus particles are interveners in *wh*-questions since they are not interveners for anaphoric binding:

- (26) a. {Only | Even} John considered buying a car_i. It_i was very {expensive | cheap}.
- b. Auch Anna erwog, einen Wagen_i zu kaufen. Er_i war sehr billig.
‘It is also the case that Anna_F considered buying a car_i. It_i was very cheap.’

However, [5] gives independent reasons for the observations that question words cannot occur in the scope of *only* and *even*. We illustrate for *only* (for *even*, see [5]). Recall that Haida assumes that *wh*-questions have an existence presupposition. This means that the question in (27) has the content given underneath, which is either trivial or contradictory or it forces a presupposition violation.⁸

- (27) *Only Ann saw who?
‘For which *x*: only Ann_i saw *x* where *x* is a person that she_i saw?’

Thus (27) does not partition the logical space in a non-trivial way. Hence, it is deviant. The meaning of *also* does not clash with the existence presupposition.

4 The experiments

EXPERIMENT 1 tested the predictions of [2] and of [5] in a speeded-acceptability task with word-by-word presentation. We tested if the focus particles (FPs) *nur* and *auch* have differential effects on the acceptability of multiple *wh*-questions in German where the FP asymmetrically c-commands the *in-situ wh*-word, i.e. where there is an intervention constellation (+INT), or where the *in-situ wh*-word asymmetrically c-commands the FP, i.e. where there is no intervention constellation (−INT). An example set of the test sentences is given below.

Context: Nachher auf dem Hinterhof wird der Deutschlehrer {auch | nur} die Austauschschüler streng rügen. ‘Later in the backyard, the._{NOM} German teacher will sternly scold {also | only} the exchange students.’

- (1) −INT *auch* Welcher Lehrer wird wo auch die Austauschschüler streng rügen?
which._{NOM} teacher will where also the exchange students sternly scold
- (2) −INT *nur* Welcher Lehrer wird wo nur die Austauschschüler streng rügen?
- (3) +INT *auch* Welcher Lehrer wird auch die Austauschschüler wo streng rügen?
- (4) +INT *nur* Welcher Lehrer wird nur die Austauschschüler wo streng rügen?

The two theories make the following predictions for these test sentences. Both [2] and [5] predict that there should not be a difference in acceptability between the −INT conditions, i.e. between

⁸To see this more clearly, consider the bound reading of the sentence in (i). If there is no *y* ≠ Ann, (i) is a tautology. Otherwise, (i) is a contradiction.

(i) #Only Ann_i saw someone she_i saw. ‘For all *y* ≠ Ann, *y* did not see someone *y* saw.’

(1) and (2). For the +INT conditions (3) and (4), the theories make differential predictions. If [5] is correct about *auch* not being an intervener and *nur* being an intervener, condition (3) with *auch* should be more acceptable than condition (4) with *nur*. Overall, [5] predicts a drop in acceptability for the +INT conditions vs. the -INT conditions because of the expected drop in acceptance for condition (4). [2], in contrast, predicts only one difference: -INT should be better than +INT.

Method. **Participants.** 40 native speakers of German from the Berlin-Brandenburg region took part in the experiment after giving informed consent. They received 7 Euros in return.

Design and materials. The design was a 2×2 within-subject design with the factors FP (*auch* vs. *nur*) and INT (intervention configuration or not), yielding the four conditions given above. There were $4 \times 10 = 40$ experimental items per participant and condition, which were distributed over lists in a Latin square design, including 88 unrelated fillers. Each experimental item included a context sentence and the critical *wh*-question, which asked for information that was given in the context sentence. All subject *wh*-phrases in the experiment had masculine gender so that the first *wh*-phrase was unambiguously marked for nominative. **Procedure.** Participants were tested individually by using the software DMDX.⁹ The items were randomly presented in the centre of the screen, word-by-word with 400 milliseconds per word plus 100 ms between words, and 200 ms after the last word of each item. Participants judged the well-formedness of the question - whether it sounded *natürlich* 'natural' or not - within a maximal interval of 3000 ms by pressing one of two buttons. 1000 milliseconds after the response, the next trial began. Prior to the experimental session, participants were told that the term *natürlich* meant that they were to judge the question using their own intuitions rather than prescriptive grammar rules, i.e. that they were to judge for grammaticality. For illustration, they were given some unrelated examples for grammatical and ungrammatical sentences with and without syntactic and semantic violations. They received practice trials in order to ensure that they had understood the task.

Results. A statistical analysis of the acceptance rates was performed by fitting a generalized linear mixed model [1] with a logistic link function, for planned comparisons. We coded contrasts for the factor INT (+1 for conditions (1) and (2), -1 for (3) and (4)), for the FP in the non-intervention constellation (+1 for *auch*, -1 for *nur*), and for the FP in the intervention constellation (+1 for *auch*, -1 for *nur*). Participants and items were random factors. Figure 1 illustrates the mean acceptance proportions for conditions (1) through (4).

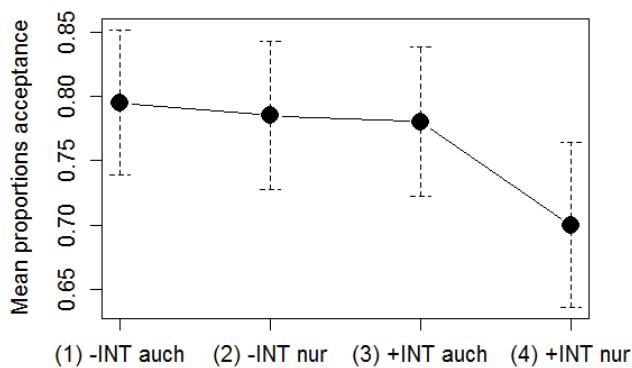


Figure 1: Mean proportions of acceptance with 95% CI in experiment 1

The comparison of the intervention constellation vs. the non-intervention constellation yielded a marginally significant effect (*estimate* = 0.168, *se* = 0.098, *z* = 1.718). -INT was accepted more often than +INT. The comparison of conditions (1) and (2) yielded no difference between these conditions (*estimate* = 0.043, *se* = 0.142, *z* = 0.300). The comparison of conditions (3) and (4) yielded a significant effect (*estimate* = 0.271, *se* = 0.135, *z* = 2.009). Condition (3) with the FP *auch* was accepted more often than condition (4) with the FP *nur*.

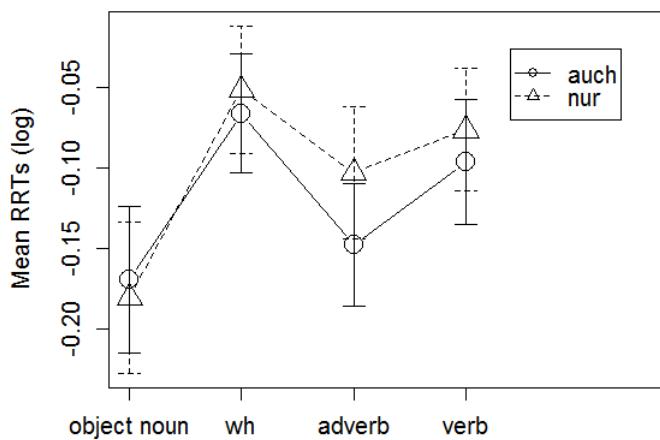
⁹<http://www.u.arizona.edu/~jforster/dmdx.htm>

Discussion. These results fit the predictions of the approach of [5], suggesting that *auch* and *nur* differ in their intervener status, i.e. *auch* does not function as an intervener between the question operator and an *in-situ wh*-word whereas *nur* does.

EXPERIMENT 2 was designed to investigate the online processing of *auch* vs. *nur* as potential interveners. It was a self-paced reading study where participants read the same experimental items as in experiment 1 but only in conditions (3) and (4), i.e. in the +INT configuration. A direct comparison with (1) and (2) is not meaningful due to the different positions of the *in-situ wh*-phrase. For the conditions tested in this experiment, [5] predicts an increase in reading time for questions with *nur* compared to questions with *auch* when, or shortly after, the reader reads the *in-situ wh*-word because this is where the problematic semantic integration is expected to be carried out. [2] predicts no difference between questions with *auch* vs. questions with *nur*.

Method. 40 native speakers of German from the Berlin-Brandenburg region took part in the experiment after giving informed consent. They received 7 Euros in return. **Design and materials.** The design was a one-factorial within subjects-design with the two-level factor FP (*auch* vs. *nur*). There were $2 \times 10 = 20$ experimental items, which were distributed over two lists, including 112 fillers from other experiments. The items were the same as in experiment 1, conditions (3) and (4). **Procedure.** Participants were tested individually by using the software *Presentation*. They read the items at their own pace. The items were randomly presented word-by-word non-cumulatively in a stationary window in the centre of the screen. Participants pressed the space bar to proceed from word to word. After each item, they answered the question by choosing one of three answers as the correct one, which ensured that they read the discourses for their meaning.

Results. The statistical analysis (linear mixed effect model with contrasts +1 for *auch* and -1 for *nur*), was conducted over residual reading times (RRTs), which were calculated per participant on the basis of test & filler items.



The mean RRTs for the clause-final region are given in figure 2. There was a weak effect on the adverb after the *in-situ wh*-word (*estimate* = -0.022, *se* = 0.012, *t* = -1.828, *p* (*MCMC-estimated*) = 0.068) but not in any other position. RRTs were higher in questions with *nur* than in questions with *auch*. There were no effects on the accuracy or speed in the answer task. The mean answer accuracy was 88%.

Figure 2: Mean RRTs for the clause-final region with 95% CI in experiment 2

Discussion. These results again support the approach of [5]. The observed effect in the reading times arises at a moment during processing when this is expected if *nur* is an intervener whereas *auch* is not – right after the *in-situ wh*-word, when we would expect interpretation difficulties to arise.

EXPERIMENT 3 was a speeded acceptability study which used the same method as exper-

iment 1 but tested object questions rather than subject questions as in experiment 1. There was an additional control condition with a non-focused DP (*der Deutschlehrer* ‘the German teacher’) as potential intervener. For reasons of space we only sketch the most important results here. There were reduced acceptance proportions for questions with *nur* in the +INT condition compared to the –INT condition. There were no effects of the factor INT for conditions with a non-focussed DP and for conditions with *auch*. Overall, the results of experiment 3 corroborate the findings of experiments 1 and 2, again supporting [5].

5 Conclusion

The results of the three experiments converge in confirming the predictions of the indefinites approach of [5] and in disconfirming the predictions of the focus approach of [2] for German. This suggests that, at least in German, there are two types of interveners: those that block the question operator from accessing the context alternatives induced by the *wh*-phrase, i.e. interveners in the literal sense. In addition, there are operators that must not have a *wh*-phrase in their scope as this leads to a semantic clash with the meaning of the operator. The focus particle *nur* is among them but *auch* is not. This raises the question of why German seems to be different from Korean and Malayalam with respect to the intervention status of *also*, see [2]. The most obvious difference between these languages are the *wh-ex-situ* vs. *wh-in-situ* question formation strategies. However, it is not quite clear how this would have a bearing on the compositional semantics of *wh*-questions. This requires further empirical work in other languages.

References

- [1] Douglas Bates and Deepayan Sarkar. lme4: Linear mixed-effects models using S4 classes (R package version .9975-11), 2007.
- [2] Sigrid Beck. Intervention effects follow from focus interpretation. *Natural Language Semantics*, 14:1–56, 2006.
- [3] Alastair Butler. Semantic control and intervention. In Catherine Pilière, editor, *Proceedings of the ESSLLI-2000 Student Session*, pages 45–54, 2000.
- [4] Jeroen Groenendijk and Martin Stokhof. A note on interrogatives and adverbs of quantification. In Chris Barker and David Dowty, editors, *Proceedings from the Second Conference on Semantics and Linguistic Theory*, pages 99–124, Columbus, Ohio, 1992. The Ohio State University.
- [5] Andreas Haida. *The Indefiniteness and Focusing of Wh-Words*. PhD thesis, Humboldt-Universität, Berlin, 2007.
- [6] Martin Honcoop. Towards a dynamic semantics account of weak islands. In Teresa Galloway and Justin Spence, editors, *Proceedings from Semantics and Linguistic Theory VI*, pages 93–110, Ithaca, 1996. Cornell University.
- [7] Shin-Sook Kim. Focus matters: Two types of intervention effects. Paper presented at WCCFL XXI, 2002.
- [8] Clemens Mayr. Intervention effects and additivity. *Journal of Semantics*, 2013.
- [9] Reinhard Muskens. Combining montague semantics and discourse representation. *Linguistics and Philosophy*, 19:143–186, 1996.
- [10] Mats Rooth. A theory of focus interpretation. *Natural Language Semantics*, 1:75–116, 1992.
- [11] Anna Szabolcsi. How unitary are intervention effects? Handout of a talk at the Brussels Conference on Generative Linguistics, June 8–10, 2006.

Monadic Quantifiers Recognized by Deterministic Pushdown Automata

Makoto Kanazawa

National Institute of Informatics, Tokyo, Japan
kanazawa@nii.ac.jp

Abstract

I characterize the class of type $\langle 1 \rangle$ quantifiers (or, equivalently, type $\langle 1, 1 \rangle$ quantifiers satisfying Conservativity and Extension) that are recognized by deterministic pushdown automata in terms of the associated semilinear sets of vectors in \mathbb{N}^2 . These semilinear sets are finite unions of linear sets with at most two generators each, which are taken from a common three-element set of the form $\{(k, 0), (0, l), (m, n)\}$. This answers a question that was left open by Mostowski (1998). A consequence of my characterization is that the type $\langle 1 \rangle$ quantifiers recognized by deterministic pushdown automata are already recognized by deterministic *one-counter* machines with zero tests, i.e., deterministic pushdown automata whose stack alphabet contains just one symbol (besides the bottom-of-stack symbol).

1 Introduction

A *type $\langle 1 \rangle$ quantifier* is a class of finite first-order structures of the form (U, P) , with $P \subseteq U$, which is closed under isomorphism. (We allow $U = \emptyset$.) Some examples of type $\langle 1 \rangle$ quantifiers are¹

$$\begin{aligned}\exists &= \{(U, P) \mid P \neq \emptyset\}, \\ \forall &= \{(U, P) \mid U = P\}, \\ \mathbf{D}_n &= \{(U, P) \mid n \text{ divides } |P|\} \quad (n = 1, 2, \dots), \\ Q^R &= \{(U, P) \mid |U - P| < |P|\}.\end{aligned}$$

Linguistically, the interest of type $\langle 1 \rangle$ quantifiers mostly owes to their correspondence with a subclass of the *type $\langle 1, 1 \rangle$ quantifiers*, which are isomorphism-closed classes of first-order structures of the form (U, P_1, P_2) (with $P_1, P_2 \subseteq U$). The correspondence is through the operation of *relativization*:

$$Q^{\text{rel}} = \{(U, P_1, P_2) \mid (P_1, P_1 \cap P_2) \in Q\}.$$

Relativizations of type $\langle 1 \rangle$ quantifiers are precisely those type $\langle 1, 1 \rangle$ quantifiers satisfying *Conservativity* and *Extension* (see Peters and Westerståhl, 2006):

$$\begin{aligned}\text{Conservativity:} \quad (U, P_1, P_2) \in Q &\iff (U, P_1, P_1 \cap P_2) \in Q. \\ \text{Extension:} \quad (U, P_1, P_2) \in Q &\iff (P_1 \cup P_2, P_1, P_2) \in Q.\end{aligned}$$

Many natural language determiners apparently express type $\langle 1, 1 \rangle$ quantifiers satisfying Conservativity and Extension:

$$\begin{aligned}\text{some} &= \{(U, P_1, P_2) \mid P_1 \cap P_2 \neq \emptyset\}, \\ \text{every} &= \{(U, P_1, P_2) \mid P_1 \subseteq P_2\}, \\ \text{an-even-number-of} &= \{(U, P_1, P_2) \mid |P_1 \cap P_2| \text{ is even}\}, \\ \text{more-than-half-of} &= \{(U, P_1, P_2) \mid |P_1 - P_2| < |P_1 \cap P_2|\}.\end{aligned}$$

¹Here, $|P|$ denotes the cardinality of the set P . Elsewhere, we sometimes write $|w|$, where w is a string, to denote the length of w . Context should make it clear which is intended.

These are relativizations of $\exists, \forall, \mathbf{D}_2, Q^R$, respectively.

Because of the correspondence via relativization, classifications of type $\langle 1 \rangle$ quantifiers in terms of their complexity translate into classifications of type $\langle 1, 1 \rangle$ quantifiers satisfying Conservativity and Extension, and vice versa. In this paper, I mostly speak of type $\langle 1 \rangle$ quantifiers, but all the results equally pertain to type $\langle 1, 1 \rangle$ quantifiers satisfying Conservativity and Extension.

Because of isomorphism closure, each type $\langle 1 \rangle$ quantifier Q has two alternative presentations: (i) a set V_Q of vectors in \mathbb{N}^2 , and (ii) a commutative (i.e., permutation-closed) set W_Q of strings over a two-letter alphabet, say, $\{a, b\}$:

$$V_Q = \{ (|U - P|, |P|) \mid (U, P) \in Q \}, \quad W_Q = \{ w \in \{a, b\}^* \mid \#(w) \in V_Q \}.$$

Here, $\#(w)$ is the Parikh vector associated with w , defined by

$$\#(w) = (\#_a(w), \#_b(w)),$$

where $\#_c(w)$ denotes the number of occurrences of symbol c in w . Conversely, any subset of \mathbb{N}^2 (and any commutative subset of $\{a, b\}^*$) determines a type $\langle 1 \rangle$ quantifier (and via relativization, a type $\langle 1, 1 \rangle$ quantifier satisfying Conservativity and Extension). These correspondences are bijections.

Van Benthem (1986) studied the relationship among the three presentations of quantifiers, proving, among other things, that W_Q is accepted by a nondeterministic pushdown automaton (PDA) if and only if V_Q is a *semilinear* subset of \mathbb{N}^2 . Since one of the motivations for using automata to classify quantifiers was to bring a procedural perspective to natural language semantics, it makes sense, as noted by van Benthem (1986), to investigate the effect of imposing *determinism* on automata, since nondeterministic automata do not correspond to well-defined algorithms (on a sequential model of computation).² It is known that deterministic PDAs accept a proper subclass of the context-free languages, known as the *deterministic context-free languages* (DCFL), which is closed under complementation, but not union. It is certainly interesting to obtain a van Benthem-style characterization of type $\langle 1 \rangle$ quantifiers Q such that W_Q is recognized by a deterministic PDA, in terms of the associated set V_Q of vectors in \mathbb{N}^2 .

A partial result in this direction was obtained by Mostowski (1998), who gave a characterization of the class of quantifiers that are accepted by deterministic PDAs *by empty stack* using a restricted class of semilinear sets which he called *almost linear*.³ This result does not cover all quantifiers that are accepted by deterministic PDAs by final state (and arbitrary stack), including such mundane quantifiers as Q^R (*more-than-half-of-all-things*).

In this paper, I give a complete characterization of the semilinear sets of vectors in \mathbb{N}^2 that correspond to type $\langle 1 \rangle$ quantifiers recognized by deterministic PDAs (by final state and arbitrary stack). These semilinear sets are finite unions of linear sets with at most two generators each, which are taken from a common three-element set of the form $\{(k, 0), (0, l), (m, n)\}$. One

²It would be unrealistic to assume that the human cognitive process of evaluating quantified sentences under normal circumstances even remotely resembles computation on finite or pushdown automata, whose access to the input (the string representation of the described situation) is limited to one-time, one-way scan. Nevertheless, these automata can make finer distinctions than standard models of computation that are used to define computational complexity classes, and are sometimes useful to classify problems that are already of very low computational complexity. For experimental studies of the difficulty of evaluating sentences with different quantifiers, see Szymanik and Zajenkowski 2011, Zajenkowski and Szymanik 2013, and references cited therein.

³Mostowski's definition of acceptance by empty stack is not altogether clear, and seems to be different from the standard one given in Section 2.2 below. Presumably, he works with a model that allows both pushing onto empty stack and testing stack for emptiness. At any rate, not all DCFLs are accepted by empty stack under Mostowski's model, even when restricted to commutative languages over $\{a, b\}$, as Mostowski notes.

consequence of this characterization is that a type ⟨1⟩ quantifier Q is recognized by a deterministic PDA if and only if W_Q is *real-time* recognizable by a (deterministic) *one-counter machine* (Fischer et al., 1968).

2 Preliminaries

2.1 Semilinear Sets

A subset of \mathbb{N}^n is *linear* if it is of the form

$$L(\vec{v}_0; \{\vec{v}_1, \dots, \vec{v}_r\}) = \{ \vec{v}_0 + k_1 \vec{v}_1 + \dots + k_r \vec{v}_r \mid k_i \in \mathbb{N} \ (1 \leq i \leq r) \}, \quad (\dagger)$$

where $\vec{v}_0, \vec{v}_1, \dots, \vec{v}_r$ are elements of \mathbb{N}^n . The vectors $\vec{v}_1, \dots, \vec{v}_r$ are called the *generators* of this set, and the vector \vec{v}_0 is called its *offset*.⁴

A *semilinear* set is a finite union of linear sets. It is well known (Ginsburg and Spanier, 1966) that the semilinear subsets of \mathbb{N}^n are precisely those definable by formulas of *Presburger arithmetic*, the first-order language with addition as its only function symbol. It is known that every semilinear set $S \subseteq \mathbb{N}^2$ can be expressed as a finite union of linear sets $S_1 \cup \dots \cup S_q$ each of which has at most two generators (Abe, 1995).

Let $\Sigma = \{a_1, \dots, a_n\}$. If $L \subseteq \Sigma^*$, then the *Parikh image* of L is $\#(L) = \{\#(w) \mid w \in L\}$, where $\#(w) = (\#_{a_1}(w), \dots, \#_{a_n}(w))$. Parikh's theorem (Parikh, 1966) states that every context-free language has a semilinear Parikh image.

2.2 Pushdown Automata

We adopt a fairly standard definition of the pushdown automaton given by Hopcroft and Ullman (1979). A *pushdown automaton* (PDA) is a system $M = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$, where Q is a finite set of *states*, Σ is finite set called the *input alphabet*, Γ is a finite set called the *stack alphabet*, $q_0 \in Q$ is the *initial state*, $Z_0 \in \Gamma$ is the *start symbol*, $F \subseteq Q$ is the set of *final states*, and δ is a mapping from $Q \times (\Sigma \cup \{\varepsilon\}) \times \Gamma$ to finite subsets of $Q \times \Gamma^*$.⁵ A *configuration* of M is a triple $(q, w, \gamma) \in Q \times \Sigma^* \times \Gamma^*$. An initial configuration is (q_0, w, Z_0) , and the *transition relation* \vdash_M between configurations is defined by

$$\vdash_M = \{ ((q, aw, Z\alpha), (p, w, \beta\alpha)) \mid (p, \beta) \in \delta(q, a, Z) \},$$

where a ranges over $\Sigma \cup \{\varepsilon\}$. We write \vdash_M^* for the reflexive transitive closure of \vdash_M . We say that M *accepts* a language $L \subseteq \Sigma^*$ by *final state* when

$$L = \{ w \in \Sigma^* \mid (q_0, w, Z_0) \vdash_M^* (p, \varepsilon, \gamma) \text{ for some } p \in F \text{ and } \gamma \in \Gamma^* \}.$$

The PDA M *accepts* L by *empty stack* if

$$L = \{ w \in \Sigma^* \mid (q_0, w, Z_0) \vdash_M^* (p, \varepsilon, \varepsilon) \text{ for some } p \in Q \}.$$

A PDA M is *deterministic* if every configuration allows transition to at most one configuration. Formally, $M = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ is deterministic if (i) $\delta(q, \varepsilon, Z) \neq \emptyset$ implies $\delta(q, a, Z) = \emptyset$ for all $a \in \Sigma$; and (ii) $\delta(q, a, Z)$ contains at most one element for all $a \in \Sigma \cup \{\varepsilon\}$.

⁴Note that the generators and the offset of a linear set depend on its representation in the form of (†).

⁵We write ε for the empty string and A^* for the set of strings over the alphabet A .

A language is a *deterministic context-free language* (DCFL) if some deterministic PDA accepts it by final state.

With nondeterministic PDAs, acceptance by empty stack is equivalent to acceptance by final state. This equivalence does not hold for deterministic PDAs, however. If some deterministic PDA accepts L by empty stack, then there is a deterministic PDA that accepts L by final state, but the converse does not hold. Since no transition to a new configuration is possible once the stack becomes empty, determinism means that whenever a string w is accepted by empty stack, no string of the form wv with $v \neq \varepsilon$ can be so accepted. Thus, any language L that some deterministic PDA accepts by empty stack is *prefix-free* in the sense that no proper prefix of an element of L belongs to L .⁶

It is known that infinite loops and blocking can be eliminated from deterministic PDAs, which makes it possible to use computation on a deterministic PDA as a recognition algorithm for the language it accepts (either by final state or by empty stack).⁷ For this reason, when a deterministic PDA M accepts a language L by final state, we say that M *recognizes* L . (When the acceptance is by empty stack, we say “recognizes by empty stack”.)

3 Main Result

We say that a deterministic PDA M *recognizes* a type $\langle 1 \rangle$ quantifier Q if M recognizes W_Q .

Theorem 1. *A type $\langle 1 \rangle$ quantifier Q is recognized by a deterministic PDA if and only if there exist natural numbers k, l, m, n such that V_Q is a finite union of linear sets each of which has one of the following as its set of generators:*

$$\emptyset, \quad \{(k, 0)\}, \quad \{(0, l)\}, \quad \{(m, n)\}, \quad \{(k, 0), (m, n)\}, \quad \{(0, l), (m, n)\}.$$

Example 2. Examples of quantifiers that satisfy the condition in Theorem 1 are

- *more than two thirds*, with the associated set of vectors $L((0, 1); \{(0, 1), (1, 2)\})$;
- *there are an odd number more Ps than non-Ps* (i.e., the Ps outnumber the non-Ps by an odd number), with the associated set of vectors $L((0, 1); \{(0, 2), (1, 1)\})$;
- *either there are three more than twice as many Ps as non-Ps or there are less than twice as many Ps as non-Ps*, with the associated set of vectors $L((0, 3); \{(1, 2)\}) \cup L((1, 0); \{(1, 0), (1, 2)\}) \cup L((1, 1); \{(1, 0), (1, 2)\})$.

In contrast, a semilinear quantifier like *more than one third but less than two thirds*, whose associated set of vectors $L((1, 1); \{(2, 1), (1, 2)\}) \cup L((2, 2); \{(2, 1), (1, 2)\})$ involves two non-trivial ratios, is excluded by the theorem.

The proof of the theorem in one direction relies on the following corollary of the pumping lemma for DCFLs:

Lemma 3 (Harrison 1978). *Let $L \subseteq \Sigma^*$ be a DCFL. There exists a positive integer p satisfying the following property: for every $w \in L$ with $|w| \geq p$, there exist x_1, x_2, x_3, x_4, x_5 such that*

- (i) $w = x_1 x_2 x_3 x_4 x_5$;

⁶The languages that some deterministic PDA accepts by empty stack coincide with the languages generated by *LR(0)* grammars (see Hopcroft and Ullman, 1979). Mostowski (1998), who gave a characterization of type $\langle 1 \rangle$ quantifiers recognized by deterministic PDAs by empty stack, seems to have a different conception of PDA, since he allows transition from a configuration with empty stack. See footnote 3.

⁷This also implies that the class of DCFLs is closed under complementation.

- (ii) $x_2x_4 \neq \varepsilon$;
- (iii) for every $z \in \Sigma^*$ and $n \in \mathbb{N}$, $x_1x_2x_3x_4z \in L$ if and only if $x_1x_2^n x_3 x_4^n z \in L$.

Another simple but useful lemma is the following:

Lemma 4. Suppose that a semilinear set $S \subseteq \mathbb{N}^2$ is bounded in the first component, i.e., there is a p such that $S \subseteq [0, p] \times \mathbb{N}$. Then there is an l such that S is a finite union of linear sets each of which has \emptyset or $\{(0, l)\}$ as its set of generators. (Analogously for the second component.)

To prove the “only if” direction of Theorem 1, suppose that W_Q is recognized by a deterministic PDA. By Parikh’s theorem, V_Q is semilinear. If W_Q is finite, then V_Q is a finite union of singletons, which are linear sets with \emptyset as the set of generators. If W_Q is infinite, then there must be $w = x_1x_2x_3x_4x_5 \in W_Q$ that satisfy the conditions (i)–(iii) of Lemma 3. Let $\vec{u}_0 = \#(x_1x_3)$ and $\vec{u}_1 = \#(x_2x_4)$. Define⁸

$$V_0 = \{ \vec{v} \in V_Q \mid \vec{u}_0 \leq \vec{v} \}, \quad V_1 = \{ \vec{v} \in V_Q \mid \vec{u}_0 < \vec{v}, \vec{u}_0 + \vec{u}_1 \leq \vec{v} \}.$$

Since the semilinear sets are closed under intersection, it is not difficult to see using Lemma 4 that there are $k, l \geq 1$ such that V_0 and V_1 are both finite unions of linear sets whose set of generators is one of \emptyset , $\{(k, 0)\}$, and $\{(0, l)\}$.

Now we claim

$$V_Q - V_0 = \bigcup \{ L(\vec{t}, \{\vec{u}_1\}) \mid \vec{t} \in V_1 \}. \quad (\ddagger)$$

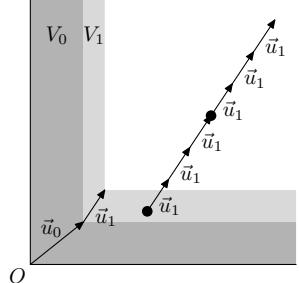
To see the \subseteq direction of (\ddagger) , suppose $\vec{v} \in V_Q$ and $\vec{u}_0 \leq \vec{v}$. Then there must be an $n \geq 0$ and $\vec{t} \in \mathbb{N}^2$ such that $\vec{u}_0 \leq \vec{t}$, $\vec{u}_0 + \vec{u}_1 \not\leq \vec{t}$, and $\vec{v} = \vec{t} + n \cdot \vec{u}_1$. Since $\vec{u}_0 \leq \vec{t}$, there is a $z \in \{a, b\}^*$ such that $\vec{t} = \#(x_1x_3z)$. Since $\vec{v} = \#(x_1x_2^n x_3 x_4^n z)$, we get $x_1x_2^n x_3 x_4^n z \in W_Q$, which implies $x_1x_3z \in W_Q$, by Lemma 3. So $\vec{t} \in V_1$ and $\vec{v} \in L(\vec{t}, \{\vec{u}_1\})$.

To see the \supseteq direction of (\ddagger) , let $\vec{t} \in V_1$. Then there is a $z \in \{a, b\}^*$ such that $\vec{t} = \#(x_1x_3z)$. Since $\vec{t} \in V_1$, $x_1x_3z \in W_Q$, and this implies $x_1x_2^n x_3 x_4^n z$ for all $n \geq 0$, by Lemma 3. So $\vec{t} + n \cdot \vec{u}_1 \in V_Q$ for all $n \geq 0$, i.e., $L(\vec{t}, \{\vec{u}_1\}) \subseteq V_Q$. Since $\vec{t} \notin V_0$, it is clear that $L(\vec{t}, \{\vec{u}_1\}) \subseteq V_Q - V_0$.

Let $(m, n) = \vec{u}_1$. Now we can show that $V_Q - V_0$ is a finite union of linear sets whose set of generators is one of $\{(m, n)\}$, $\{(k, 0), (m, n)\}$, and $\{(0, l), (m, n)\}$. Suppose $V_1 = L(\vec{t}_1, G_1) \cup \dots \cup L(\vec{t}_q, G_q)$, where each G_i is one of \emptyset , $\{(k, 0)\}$, and $\{(0, l)\}$. Then $V_Q - V_0 = L(\vec{t}_1, G'_1) \cup \dots \cup L(\vec{t}_q, G'_q)$, where $G'_i = G_i \cup \{(m, n)\}$. Clearly, G'_i is among $\{(m, n)\}$, $\{(k, 0), (m, n)\}$, $\{(0, l), (m, n)\}$. Since V_0 is a finite union of linear sets whose set of generators is one of \emptyset , $\{(k, 0)\}$, and $\{(0, l)\}$, we have shown that V_Q is as specified in Theorem 1.

For the “if” direction of Theorem 1, assume that V_Q is as specified in the theorem. We can construct a deterministic PDA recognizing W_Q as follows. We use part of the finite control as buffers to store bounded numbers of a s and b s, where the bound for the a -buffer is the maximal a -component (i.e., first component) of the offsets plus m , and likewise for the b -buffer. When scanning an a when the a -buffer is full, we push an a onto the stack, and likewise when scanning a b when the b -buffer is full. When both buffers become full, we take out m a s and n b s from their respective buffer, and move the symbols on the stack to the appropriate buffer until

⁸The inequality \leq between vectors in \mathbb{N}^2 is defined by: $(x, y) \leq (x', y')$ iff $x \leq x'$ and $y \leq y'$. The strict inequality $(x, y) < (x', y')$ holds iff $(x, y) \leq (x', y')$ and $(x, y) \neq (x', y')$.



the buffer becomes full or the stack becomes empty (using bottom-of-stack symbol to test for emptiness), whichever comes first. We also count the number of *as* (or *bs*) on the stack modulo k (or l) and keep this information in the finite control. Given this much information in the finite control, inspection of a bounded portion of the stack near the top suffices to determine whether the part of the input scanned so far is in W_Q .

This deterministic PDA always keeps its stack uniform—the stack always contains just one kind of symbol (besides the bottom-of-stack symbol). Using a flag in the finite control to indicate which symbol is in the stack, we can easily turn it into a *one-counter machine* (1-CM) (Fischer et al., 1968), which is like a deterministic PDA but has a counter instead of a pushdown stack, which can hold any natural number and can be tested for zero.

Let us see the working of our 1-CM in more detail. If either $m = 0$ or $n = 0$, it is easy to see that W_Q is regular, so we may assume $m > 0$ and $n > 0$. Since a regular set can be recognized using just the finite control, we may also discard linear sets of the form $L(\vec{u}, \emptyset)$, $L(\vec{u}, \{(k, 0)\})$, or $L(\vec{u}, \{(0, l)\})$, which correspond to regular subsets of W_Q . Writing $L(O; G)$ for $\bigcup\{L(\vec{u}; G) \mid u \in O\}$, we assume

$$V_Q = L(O_1; \{(m, n)\}) \cup L(O_2; \{(k, 0), (m, n)\}) \cup L(O_3; \{(0, l), (m, n)\}),$$

where O_1, O_2, O_3 are finite subsets of \mathbb{N}^2 . Let⁹

$$\begin{aligned} \text{max_offset_a} &= \max\{x \mid (x, y) \in O_1 \cup O_2 \cup O_3\}, \\ \text{max_offset_b} &= \max\{y \mid (x, y) \in O_1 \cup O_2 \cup O_3\}, \\ C &= \max(n \cdot \lfloor \text{max_offset_a}/m \rfloor, m \cdot \lfloor \text{max_offset_b}/n \rfloor). \end{aligned}$$

Our 1-CM is an implementation of the following pseudocode:

```

a_buffer ← 0; b_buffer ← 0; count ← 0; rem ← 0; counted ← a
loop
    accept ← CHECKFORACCEPTANCE()
    if ENDOFINPUT() then
        return accept
    else
        c ← GETNEXTSYMBOL()
        if c = a then
            if a_buffer = max_offset_a + m then
                counted ← a; count ← count + 1; rem ← (rem + 1) mod k
            else
                a_buffer ← a_buffer + 1
            end if
        end if
        if c = b then
            if b_buffer = max_offset_b + n then
                counted ← b; count ← count + 1; rem ← (rem + 1) mod l
            else
                b_buffer ← b_buffer + 1
            end if
        end if
        if (a_buffer, b_buffer) = (max_offset_a + m, max_offset_b + n) then
            (a_buffer, b_buffer) ← (a_buffer, b_buffer) - (m, n)
            if counted = a then

```

⁹If x is a real number, $\lfloor x \rfloor$ is the integer part of x .

```

while count > 0  $\wedge$  a_buffer < max_offset_a + m do
    count  $\leftarrow$  count - 1; rem  $\leftarrow$  (rem - 1) mod k; a_buffer  $\leftarrow$  a_buffer + 1
    end while
end if
if counted = b then
    while count > 0  $\wedge$  b_buffer < max_offset_b + n do
        count  $\leftarrow$  count - 1; rem  $\leftarrow$  (rem - 1) mod l; b_buffer  $\leftarrow$  b_buffer + 1
        end while
    end if
end if
end if
end loop

procedure CHECKFORACCEPTANCE()
    if count  $\leq$  C then
        if (counted = a  $\wedge$  (a_buffer + count, b_buffer)  $\in$   $V_Q$ )  $\vee$  (counted = b  $\wedge$  (a_buffer, b_buffer + count)  $\in$   $V_Q$ ) then
            return true
        end if
    else
        for all  $(x, y) \in O_2$  do
            if  $(x, y) \leq (a\_buffer, b\_buffer)$  then
                 $(x_1, y_1) \leftarrow (a\_buffer - x, b\_buffer - y)$ 
                if counted = a  $\wedge$   $y_1 \equiv 0 \pmod{n}$   $\wedge$   $x_1 - m \cdot (y_1/n) + rem \equiv 0 \pmod{k}$  then
                    return true
                end if
            end if
        end for
        for all  $(x, y) \in O_3$  do
            if  $(x, y) \leq (a\_buffer, b\_buffer)$  then
                 $(x_1, y_1) \leftarrow (a\_buffer - x, b\_buffer - y)$ 
                if counted = b  $\wedge$   $x_1 \equiv 0 \pmod{m}$   $\wedge$   $y_1 - n \cdot (x_1/m) + rem \equiv 0 \pmod{l}$  then
                    return true
                end if
            end if
        end for
    end if
    return false
end procedure

```

The length of any legal sequence of ε -transitions of this 1-CM is bounded by a constant. Using the compression technique of Fischer et al. (1968, proof of Theorem 1.1), we can eliminate ε -transitions from the machine to make it operate in *real time* (Fischer et al., 1968).

Corollary 5. *A type $\langle 1 \rangle$ quantifier is recognized by a deterministic PDA if and only if it is real-time recognized by a (deterministic) 1-CM.*

4 Conclusion

I have characterized the type $\langle 1 \rangle$ quantifiers recognized by deterministic PDAs in terms of the associated semilinear sets of vectors and showed that exactly the same type $\langle 1 \rangle$ quantifiers

are recognized by (deterministic) 1-CMs (in real time). Let me make a few more related observations.

A theorem of Fischer et al. (1968, Theorem 2.2) says that a language of the form $\{w \in \Sigma^* \mid \#(w) \in S\}$ is a finite Boolean combination of languages real-time recognizable by 1-CMs if and only if S is semilinear. So we have

Corollary 6. *A type $\langle 1 \rangle$ quantifier is recognized by a nondeterministic PDA if and only if it is a finite Boolean combination of type $\langle 1 \rangle$ quantifiers recognized by deterministic PDAs.*

It is relatively straightforward to show that whenever $S \subseteq \mathbb{N}^2$ is semilinear, $\{w \in \{a,b\}^* \mid \#(w) \in S\}$ is accepted by a *nondeterministic one-counter machine*, which gives us

Proposition 7. *A type $\langle 1 \rangle$ quantifier is accepted by a nondeterministic PDA if and only if it is accepted by a nondeterministic one-counter machine.*

By a similar pumping argument, we can also give a simple characterization of the semilinear sets associated with type $\langle 1 \rangle$ quantifiers recognized by finite automata:

Proposition 8. *A type $\langle 1 \rangle$ quantifier Q is recognized by a finite automaton if and only if there exist natural numbers k, l such that V_Q is a finite union of linear sets each of which has one of the following as its set of generators:*

$$\emptyset, \quad \{(k,0)\}, \quad \{(0,l)\}, \quad \{(k,0), (0,l)\}.$$

References

- Naoki Abe. Characterizing PAC-learnability of semilinear sets. *Information and Computation*, 116:81–102, 1995.
- Patrick C. Fischer, Albert R. Meyer, and Arnold L. Rosenberg. Counter machines and counter languages. *Mathematical Systems Theory*, 2:265–283, 1968.
- Seymour Ginsburg and Edwin H. Spanier. Semigroups, Presburger formulas, and languages. *Pacific Journal of Mathematics*, 16:285–296, 1966.
- Michael A. Harrison. *Introduction to Formal Language Theory*. Addison-Wesley, Reading, MA, 1978.
- John E. Hopcroft and Jeffrey D. Ullman. *Introduction to Automata Theory, Languages, and Computation*. Addison-Wesley, Reading, MA., 1979.
- Marcin Mostowski. Computational semantics for monadic quantifiers. *Journal of Applied Non-Classical Logics*, 8:107–201, 1998.
- Rohit J. Parikh. On context-free languages. *Journal of the ACM*, 13:570–581, 1966.
- Stanley Peters and Dag Westerståhl. *Quantifiers in Language and Logic*. Oxford University Press, Oxford, 2006.
- Jakub Szymanik and Marcin Zajenkowski. Contribution of working memory in parity and proportional judgments. *Belgian Journal of Linguistics*, 25:176–194, 2011.
- Johan van Benthem. *Essays in Logical Semantics*. Reidel, Dordrecht, 1986.
- Marcin Zajenkowski and Jakub Szymanik. MOST intelligent people are accurate and SOME fast people are intelligent. Intelligence, working memory, and semantic processing of quantifiers from a computational perspective. *Intelligence*, 41:456–466, 2013.

An experimental investigation of interrogative syntax/semantics*

Hadas Kotek and Martin Hackl

Massachusetts Institute of Technology
hkotek@mit.edu, hackl@mit.edu

Abstract

Current theories of interrogative syntax/semantics adopt two strategies for the interpretation of in-situ *wh*-phrases: covert movement and in-situ interpretation. The covert movement strategy is traditionally assumed to be all-or-nothing: the in-situ *wh*-phrase covertly moves to C or else stays in-situ and is interpreted in its base-generated position at LF. In this paper we argue that neither approach to *wh*-in-situ can be maintained as is. We present evidence from real-time processing of English multiple *wh*-questions that in-situ *wh*-phrases require *both* covert movement and in-situ interpretation for their computation: an "in-situ" *wh*-phrase undergoes a short movement step, parallel to the behavior of traditional quantifiers such as *every*, immediately upon integration into the structure. Following that step, the *wh*-phrase *can* but *need not* move any further unless independent factors are at play. To account for this pattern we propose the partial movement approach to *wh*-in-situ: A *wh*-phrase must be interpreted at a position with propositional type at LF. *Wh*-phrases in object position are thus never interpreted in their base-generated position; instead some covert movement is always required.

1 Introduction and background

The goal of this paper is to investigate the possible position(s) of in-situ *wh*-phrases at LF. We will argue that both current approaches to *wh*-in-situ—the covert movement approach and the in-situ approach—are insufficient to derive the correct syntax-semantics of *wh*-questions, leading us to propose a new *partial movement* approach to interrogative syntax/semantics, where *wh*-in-situ may be interpreted at any node with propositional type at LF.

In this section we introduce the two traditional approaches to the interpretation of in-situ *wh*-phrases: covert movement and in-situ interpretation and the predictions they make for the position of in-situ *wh*-phrases at LF. We then briefly discuss the experimental methodology that we will use in our experiments to probe for the LF-position of in-situ *wh*-phrases.

1.1 Two approaches to *wh*-in-situ

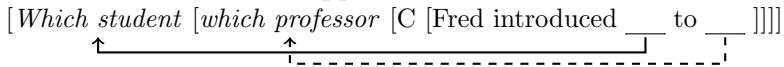
1.1.1 The covert movement approach

Under the covert movement approach to questions, *wh*-phrases must be adjacent to C in order to be able to make their contribution to the meaning of the question. Consequently, no

*For helpful comments, discussions and support, we would like to thank Danny Fox, Irene Heim, Shigeru Miyagawa, David Pesetsky, Micha Breakstone, Alexandre Cremers, Michael Yoshitaka Erlewine, Yasutada Sudo, and the audiences of NELS 44 and MIT Ling-Lunch. This material is based upon work supported by the National Science Foundation under Grant No. 1251717.

wh-phrase may remain in situ at LF; instead, all *wh*-phrases occur syntactically next to the complementizer, regardless of where they are pronounced (Karttunen, 1977; Huang, 1982; Lasnik and Saito, 1992; Hornstein, 1995; Pesetsky, 2000; Richards, 2001; Cable, 2007, a.o.).¹

(1) **The covert movement approach to *wh-in-situ***



The covert movement approach thus predicts pervasive covert movement in multiple *wh*-questions. Moreover, movement is always triggered for one and the same reason—the semantic needs of the *wh*-phrases themselves, and it always targets the same syntactic position at LF—C. This approach makes the prediction in (2).

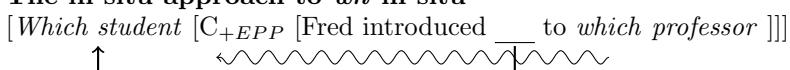
(2) **A prediction of the covert movement approach**

All *wh*-phrases in a question must (overtly or covertly) move to C for interpretation.

1.1.2 The *in-situ* approach

Under the *in-situ* approach to questions, no (overt or covert) movement is required in order to assign interrogative meaning to a structure containing *wh*-elements (Hamblin, 1973; É Kiss, 1986; Cheng, 1991; Chomsky, 1995; Reinhart, 1998; Kratzer and Shimoyama, 2002, a.o.). The meaning of *in-situ* *wh*-phrases can be calculated through a mechanism that passes the meanings of *wh*-words up the structure until they reach C, where they can be interpreted. From this perspective, there is no reason to expect any instances of *wh*-movement that are caused by the semantic needs of the *wh*-words themselves. Even the fact that English questions require overt fronting of one *wh*-phrase is unexpected. To explain this fact, a purely syntactic mechanism must be invoked, unrelated to interrogative semantics, for example an ‘EPP’ feature requiring C to have a filled specifier (Chomsky, 1981).

(3) **The *in-situ* approach to *wh-in-situ***



The *in-situ* approach to questions thus makes no predictions about the position of *wh*-phrases at LF. Following common assumptions in the theoretical literature that the simplest syntactic structure for a sentence is always preferred to a less simple one (Epstein, 1992; Chomsky, 1995; Kitahara, 1997; Collins, 2001; Fox, 2000; Reinhart, 2006, a.o.), it is predicted that *wh*-phrases occupy the position at which they were merged into the syntactic structure. No covert movement occurs for semantic interpretation of *wh*-phrases, (4).

(4) **A prediction of the *in-situ* approach**

Wh-phrases in a question can be interpreted in situ and do not require any movement.

1.2 Self-paced reading and the Hackl et al. (2012) paradigm

Hackl et al. (2012) develop a paradigm that is able to detect the presence and extent of covert movement in a structure by examining the interaction between quantifiers in object position and Antecedent Contained Ellipsis (ACE). Quantifiers in object position must QR locally to

¹Straight arrows indicate overt movement, dashed arrows indicate covert movement, and squiggly arrows indicate areas of *in-situ* interpretation.

the first node of propositional type in the structure to resolve a type mismatch (May, 1977; Heim and Kratzer, 1998, a.o.). ACE requires a constituent containing the ACE-site to QR at least as high as the antecedent (Williams, 1974, 1977; Sag, 1976, a.o.). ACE can be hosted both by quantifiers such as *every* and by non-quantificational determiners such as *the*.

In example (5), ACE can be either ellipsis of a local VP₂ (<treat...>) or non-local VP₁ (<BE reluctant to treat...>). The auxiliary verb unambiguously determines which VP is a possible antecedent: the lower VP₂ with *did*, and the higher VP₁ with *was*. To undo antecedent containment in the structure and resolve the ellipsis, the object relative clause containing the gap site must QR at least as high as its antecedent. Therefore, we expect movement above VP₂ in (5a) and above VP₁ in (5b). In online sentence processing, ACE resolution requires *reanalysis* of the structure assumed by the parser. This operation causes difficulty detectable as a slowdown in Reading Times (RTs) following the detection of ACE by the parser.

- (5) The doctor [VP₁ was reluctant to [VP₂ treat ...
 a. **the/every** patient that the recently hired nurse ***did*** local ACE
 b. **the/every** patient that the recently hired nurse ***was*** non-local ACE
 ... after looking over the test results. (Hackl et al., 2012)

In sentences with *every*, but crucially not in sentences with *the*, we expect the quantifier to QR immediately upon integration into the structure. We expect this movement to be *local*, and therefore target a position above VP₂ and no higher. As a consequence, local ACE (but not non-local ACE) should be facilitated: movement of the quantifier in object position preemptively undoes antecedent containment in the case of local ACE (5a), and therefore an antecedent is easy to retrieve when the parser reaches *did*. With non-local ACE (5b), however, the parse must be reanalyzed upon encountering *was*. The *every*-DP must move above the higher VP₁ to create an appropriate antecedent, resulting in increased processing cost compared to (5a).

Hackl et al. (2012) tested paradigms as in (5) in a self-paced reading study: participants read sentences that appeared on the screen one word at a time in a moving window display. Residual Reading Times (RRTs) were analyzed for each word in the sentence. Results support the predictions made by QR-based theories of quantifier integration: resolution of local ACE with *every* is facilitated compared to parallel sentences with *the*. Furthermore this effect does not extend to cases with non-local ACE, where *every* no longer has an advantage over *the*.

2 Experiments 1 and 2

In addition to *the* and *every*, *wh*-phrases such as *which* can also host ACE. We can thus use the methodology in Hackl et al. (2012) to experimentally test whether the processing of *wh*-in-situ involves covert movement. All three experiments we present in this paper contained 28 target sentences and 48 filler sentences. Each sentence was followed by a comprehension question. All experiments were conducted online and hosted on IBEX.² Participants were recruited through Amazon Mechanical Turk and were paid \$1.5 for their participation. Participants were asked about their native language but were told that payment was not contingent on their response. To further ensure that only native speakers of English participated in the experiments, IP addresses of participants were restricted to the US using Amazon Mechanical Turks user interface.

Experiment 1 presented participants with (embedded) *wh*-questions headed by a subject *wh*-phrase. Two factors were crossed: (a) DETERMINER: whether the embedded question contained the quantificational determiner *every*, yielding a simplex *wh*-question, or a second *wh*-phrase,

²IBEX: Internet Based Experiments, created and maintained by Alex Drummond, accessible at <http://spellout.net/ibexfarm/>.

yielding a multiple *wh*-question; and (b) ELLIPSIS SIZE: whether the sentence contained a small ellipsis marked by *did*, where the antecedent of the ACE site is the embedded VP₂, or large ellipsis marked by *was*, where the antecedent of the ACE site is the higher VP₁. A sample item is given in (6) below:

- (6) The conductor asked [_{CP} **which** soloist [_{VP₁} was willing to [_{VP₂} perform ...
 a. **which/every** concerto that the brilliant protégé ***did*** local ACE
 b. **which/every** concerto that the brilliant protégé ***was*** non-local ACE
 ... and restructured the rehearsal accordingly.

The two approaches to *wh*-in-situ make the following predictions:

- (7) a. **Covert movement:** The *which*-DP (with the relative clause, including the ACE-site) moves non-locally to C when *wh* is encountered. Ellipsis resolution is relatively easy for both local and non-local ACE compared to *every*.
 b. **In-situ interpretation:** No movement. The structure must be reanalyzed when the parser encounters the ACE-site. Both local and non-local ACE incur relatively high processing costs compared to *every*.

61 native speakers of English participated in this study. Twenty participants were excluded from the analysis for various reasons.³ In addition, two target sentences and no filler sentences were excluded from the analysis because of low accuracy (<60% across participants). Questions across the full experiment (targets and fillers) were answered correctly 87.5% of the time across participants; questions for experimental items were answered correctly on 83.3% of trials. Less than 1% of the data was trimmed using a criterion trimming.⁴ Figure 1 shows the mean residual reading times (RRTs) for the two regions of interest for the four target conditions.

Results: A linear mixed effects model with random intercepts for subjects and items was fit to the data using R and the R package lme4. The model predicted logRT from the two factors of interest: DETERMINER (every vs. which) and ELLIPSIS SIZE (small ellipsis marked by *did*, vs. large ellipsis marked by *was*). Results show a main effect of determiner at the slot at which the determiner appeared in the sentence (log likelihood tests, $p < 0.05$). This result is driven by

the fact that reaction times in the *which* condition were slower than the reaction times in the *every* condition, across both ellipsis conditions. The results additionally show a main effect of ellipsis size two words and three words after the auxiliary site (log likelihood tests, $p < 0.05$).

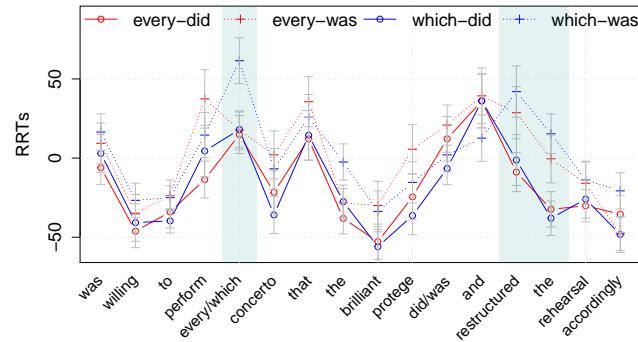


Figure 1: RRTs in target items in Experiment 1

³Participants who held the spacebar continuously pressed instead of reading the sentences one word at a time as instructed, who participated in the study more than once, who submitted the entire survey in less than 10 minutes, with an average reaction time of over 700ms, and with low accuracy rates in response to comprehension questions (<75% on filler trials and <75% on target trials) were excluded from the analysis.

⁴RRTs from the first and last words of all items, RTs faster than 90ms or slower than 2000ms, and any RTs that were more than 2 standard deviations faster or slower than the average RTs for each subject (calculated per condition) were excluded from the analysis.

This result is driven by the fact that the resolution of small ellipsis is faster than the resolution of large ellipsis for both *every* and *which*. There were no differences between the two determiners at these slots, and there were no other significant effects in the results.

This result is not predicted by both the covert movement approach and the in-situ approach to *wh-in-situ*. Recall that the covert movement approach predicts facilitation of both small and large ellipsis, because the in-situ *wh*-phrase must move to C for interpretation, preemptively undoing antecedent containment in both ellipsis conditions. We would thus expect that RTs for both *which* conditions, irrespective of whether they involve local or non-local ACE, would be lower than those for the *every* conditions. The in-situ approach, on the other hand, assumes that the *wh*-phrase can be interpreted without any movement at all, and hence predicts no facilitation by an upstream *which* for either the small or large ellipsis conditions. However, the results of Experiment 1 show an effect of ellipsis size but no effect of determiner, which is not explained under both approaches to *wh-in-situ*.

Based on the results of Hackl et al. (2012), we may take the fact that there is no difference between *every-did* and *which-did* to indicate that local ACE resolution is facilitated in both cases. More specifically, if *which* where a determiner that could be interpreted in situ like *the* we should have seen relatively longer RTs for *which-did*, just like Hackl et al. (2012) did for *every-did*. The fact that we didn't shows that *wh*-phrases behave like traditional quantificational determiners such as *every* with regard to ACE resolution in real time.

In Experiment 2, we compare the behavior of *which* and *every* against *the* in the environment of Experiment 1. **Results.** Figure 2 shows the comparison of RRTs following the ACE site for *the* and *which* on the left and *the* and *every* on the right. For both comparisons, we find a significant main effect of ELLIPSIS SIZE two and three words after the auxiliary verb site and a main effect of DETERMINER

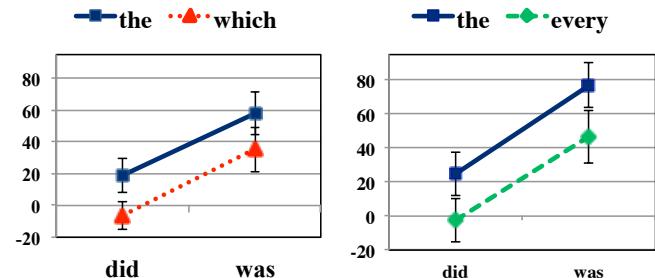


Figure 2: RRTs after the ACE site in Experiment 2

two words after the auxiliary verb site (log likelihood tests, all $p < 0.05$). This effect is driven by the fact that *the* is slower than *which* and *every* for all ellipsis conditions. The results for *every* and *the* are in line with the results of Hackl et al. (2012) in that they show that ACE resolution is more difficult when the ACE site is hosted by a definite DP. Importantly, this is also true when *the* is compared with *which*.

3 Proposal

We have seen in both Experiments 1 and 2 that *which* patterns with *every* in its effects on the resolution of ACE. This is not predicted by either approach to *wh-in-situ*. We propose that the results of Experiments 1 and 2 are explained if *wh*-phrases are quantifiers that must QR locally just like conventional quantifiers such as *every*. Similar proposals have been previously made by several researchers for a variety of languages (Baker, 1970; Dornisch, 2000; Huang, 1995; Kim, 1991; Rullmann and Beck, 1998, a.o.). We propose that once the *wh*-phrase has been integrated into the structure, it does not require any further movement: it is interpreted at the first position where it is interpretable. We propose that the movement step assumed by the parser is the smallest that can produce an interpretable structure:

- (8) **The status of *wh*-phrases**
Wh-phrases are existential quantifiers.
- (9) **The partial movement approach**
A *wh*-phrase must be interpreted at a position with propositional type at LF.

There is a possible alternative interpretation of Experiments 1–2: the difference between *the* on the one hand and *every* and *which* on the other hand is perhaps not the result of a difference in QR assumed for the interpretation of these determiners, but instead is contributed by some property of *the*, unrelated to its structural integration, which makes definite DPs relatively poor hosts for ACE. If this is true, we can no longer infer from the differences between *every* (and *which*) and *the* that upstream QR has occurred with *every* (and *which*) but not with *the*. The main effect of ELLIPSIS SIZE could be explained as an effect of the *complexity of the antecedent*: integrating a smaller, simpler antecedent into the structure is easier than integrating a larger, more complex one. The main effect of *determiner* would be a consequence of the as yet to be identified property of *the* that makes ACE resolution with it difficult.

Experiment 3 will address this alternative explanation by considering a prediction of the *partial movement* approach: that although *wh*-phrases, in principle, need not move any further than the closest node with propositional type, they may be forced to move higher on independent grounds. This prediction allows us to distinguish between the behavior of *every* and *which* in certain environments, where long-distance *wh*-movement may be necessary, leading us to expect a larger region in the question in which ACE facilitation effects are expected with *which* but not with *every*. If the alternative explanation of the results sketched above is on the right track—that is, if all Experiments 1 and 2 are showing a complexity effect, then we should not see an increased domain of ACE facilitation effects, because no long-distance QR is assumed when the parser reaches *which*. However, if the *partial movement* approach is on the right track, then additional covert movement may occur in a question and we expect to find facilitation effects of ACE resolution for any ellipsis that is smaller than the landing site of this extended covert movement step. This prediction is summarized below and tested in Experiment 3.

- (10) **Predictions of the partial movement approach:** *Wh*-phrases are quantifiers that QR to the nearest propositional node for interpretation. No additional movement is required for the interpretation of a *wh*-phrase, but movement may be forced by external interpretability considerations. *Wh*-movement may target other positions beside C.

4 Intervention effects and Experiment 3

Our proposal implies that longer-distance covert movement of a *wh*-phrase can be caused by material that triggers so called *intervention effects* (Beck, 2006, a.o.). The term intervention effect describes a situation in which a question is rendered ungrammatical because an in-situ *wh*-phrase is c-commanded at LF by an *intervener*, for instance a focus sensitive operator such as *only* or negation. Beck argues that in-situ interpretation of a *wh*-phrase is impossible when an intervener occurs above the *wh*-phrase: in that case, the *wh*-phrase must move above the intervener for interpretation.

- (11) **The configuration of an intervention effect (Beck 2006):**
 - a. $[CP \ C \ ... \ wh \ ... \ \underset{\uparrow}{\text{intervener}} \ ... t_{wh} \ ...]$
 - b. * $[CP \ C \ ... \ \underset{\leftarrow\searrow\swarrow\rightarrow}{\text{intervener}} \ wh \ ...]$

In Experiment 3 we exploit intervention effects to force long-distance covert movement, using *also* as an intervener in the items of Experiment 1. We place *also* in one of two positions: above the lower VP ('was willing to also perform...') or above the higher VP ('was also willing to perform...'). We expect *also* to force covert *wh*-movement above it because *wh*-phrases are uninterpretable when they are in-situ below an intervener. Hence, for the lower placement we expect movement to a position above the lower VP and for the high placement we expect movement above the higher VP. We furthermore expect ACE resolution to be facilitated in the entire domain of movement because the movement will preemptively undo antecedent containment in the structure, making an appropriate antecedent easier to retrieve.

As in Experiment 2, DETERMINER was a between-subject factor: in Experiment 3a participants only saw target items with *every*, and in Experiment 3b they only saw items with *which*. For Experiment 3a we expect to find no sensitivity to the presence and position of *also* because regular quantifiers are not sensitive to intervention effects. Hence, we expect to find a main effect of ELLIPSIS SIZE and nothing else. For Experiment 3b we expect to find an effect of the POSITION OF ALSO: for the high placement of *also*, we expect to find facilitation effects for the resolution of both local and non-local ACE; for the low *also*, we predict that only the resolution of local ACE is facilitated.

Results. Figure 3 shows a comparison of RRTs following the ACE site for *every* on the left and *which* on the right, for the two *also* conditions. We find a significant main effect of ELLIPSIS SIZE for *every*, but a significant interaction of ELLIPSIS SIZE and POSITION OF ALSO for *which* ($p < 0.05$).

The results partially confirm the predictions of the *partial movement* approach: the *wh*-phrase is interpreted in the lowest position at which it is interpretable. If an intervener forces long-distance movement, ACE resolution is facilitated in the entire domain of movement. That is, when *also* is placed above the higher VP with *which*, the resolution of non-local ACE is facilitated. For *every* and for the low placement of *also* with *which*, only the resolution of local ACE is facilitated, as expected, because movement will only have targeted a position above the lower VP but below the higher VP.

What is unexpected under the *partial movement* approach is that local-ACE is relatively more difficult when *also* is attached high and thus forces non-local movement of the *wh*-phrase. We propose that the relative difficulty of local ACE resolution in this case is the result of a *scope-matching preference effect* observed in both off-line judgments (Hardt and Romero, 2004) and real-time sentence processing (Breakstone et al., 2011): there is a preference for the scope of an expression that hosts an ACE gap to match the size of an elided constituent in the same sentence. The resolution of local ACE is relatively difficult in the case of *which-high also* because more movement has happened than is necessary in order to resolve the local ACE.

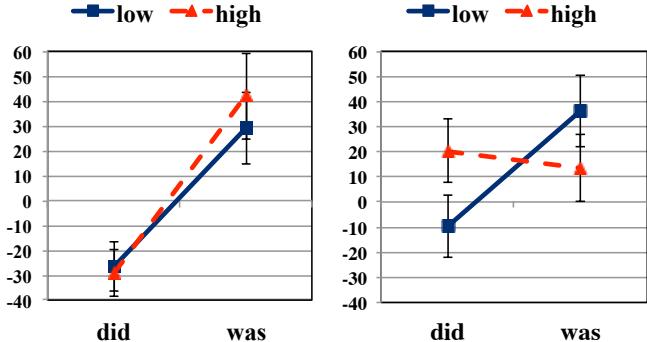


Figure 3: RRTs after the ACE site in Experiment 3

5 Conclusion

The results of Experiments 1–3 shed light on the interaction between properties of in-situ *wh*-phrases and ACE resolution. Our findings suggest the following distribution of *wh*-phrases at

LF: in-situ *wh*-phrases cannot be interpreted in their base position, but also do not necessarily move to C for interpretation. Instead, they can be interpreted at any propositional node in the structure, just like regular quantifiers. Unlike regular quantifiers, however, in-situ *wh*-phrases are subject to intervention effects. Thus, the presence of an element like *also*, which projects a domain of intervention, can force in-situ *wh*-phrases to move higher than a regular quantifier in order to escape the intervention effect.

References

- Baker, Carol L. 1970. Notes on the description of English questions: the role of an abstract question morpheme. *Foundations of Language* 6:197–217.
- Beck, Sigrid. 2006. Intervention effects follow from focus interpretation. *NLS* 14:1–56.
- Breakstone, Micha, Alexandre Cremers, Danny Fox, and Martin Hackl. 2011. On the analysis of scope ambiguities in comparative constructions. In *Proceedings of SALT 21*, 712–231.
- Cable, Seth. 2007. The grammar of Q. Doctoral Dissertation, MIT.
- Cheng, Lisa Lai-Shen. 1991. On the typology of *wh*-questions. Doctoral Dissertation, MIT.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Foris.
- Chomsky, Noam. 1995. *The minimalist program*. MIT Press.
- Collins, Chris. 2001. Economy conditions in syntax. In *Handbook of syntactic theory*. Blackwell.
- Dornisch, Ewa. 2000. Overt quantifier raising in Polish. In *Proceedings of GLiP 1*, 47–58.
- É Kiss, Katalin. 1986. Against the LF-movement of *wh*-phrases. Manuscript, Budapest.
- Epstein, Sam. 1992. Derivational constraints on A' chain formation. *Linguistic Inquiry* 23.
- Fox, Danny. 2000. *Economy and semantic interpretation*. MIT Press.
- Hackl, Martin, Jorie Koster-Hale, and Jason Varvoutis. 2012. Quantification and ACD: Evidence from real-time sentence processing. *JoS* 29:145–206.
- Hamblin, Charles. 1973. Questions in Montague English. *Foundations of Language* 10:41–53.
- Hardt, Daniel, and Maribel Romero. 2004. Ellipsis and the structure of discourse. *JoS* 21.
- Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Blackwell.
- Hornstein, Norbert. 1995. *Logical form: from GB to minimalism*. Blackwell.
- Huang, C.T. J. 1995. Logical form. In *Government and binding theory and the minimalist program*, ed. G. Webelhuth, 127–173. Blackwell.
- Huang, C.T. J. 1982. Move *wh* in a language without *wh* movement. *The Linguistic Review* 1.
- Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* 1:3–44.
- Kim, Soowon. 1991. Chain scope and quantification structure. PhD thesis, Brandeis University.
- Kitahara, Hisatsugu. 1997. *Elementary operations and optimal derivations*. MIT Press.
- Kratzer, Angelika, and Junko Shimoyama. 2002. Indeterminate pronouns: the view from Japanese. In *Proceedings of the Third Tokyo Conference on Psycholinguistics*.
- Lasnik, Howard, and Mamoru Saito. 1992. *Move α, conditions on its application and output*.
- May, Robert. 1977. The grammar of quantification. Doctoral Dissertation, MIT.
- Pesetsky, David. 2000. *Phrasal movement and its kin*. MIT Press.
- Reinhart, Tanya. 1998. *Wh-in-situ* in the framework of the Minimalist Program. *NLS* 6.
- Reinhart, Tanya. 2006. *Interface strategies*. MIT Press.
- Richards, Norvin. 2001. *Movement in language: Interactions and architectures*. Oxford.
- Rullmann, Hotze, and Sigrid Beck. 1998. Reconstruction and the interpretation of *which*-phrases. In *Proceedings of the 1997 Tübingen Workshop*.
- Sag, Ivan Andrew. 1976. Deletion and logical form. Doctoral Dissertation, MIT.
- Williams, Edwin. 1974. Rule ordering in grammar. Doctoral Dissertation, MIT.
- Williams, Edwin. 1977. Discourse and logical form. *Linguistic Inquiry* 8:101–139.

Tense and Aspect in Swing Conditionals *

Fabienne Martin

University of Stuttgart
fabienne.martin@ling.uni-stuttgart.de

1 Introduction

This paper focuses on past subjunctive conditionals (**PSCs**) in French. In French, PSCs do not have a subjunctive but rather a *conditionnel 2* in the **consequent**, cf. (1).¹

- (1) Si on avait réfléchi, on n' aurait pas signé.
If we have-IMP. thought, we NEG have-COND NEG signed
'If we had thought about it, we wouldn't have signed.'

It is often assumed that French (like Greek) requires imperfective aspect as a counterfactual (CF) marker in the **antecedent** of PSCs (cf. e.g. [Iatridou, 2000]). This should explain why we find in the antecedent of PSCs the *plus que parfait*, combining imperfective morphology with a layer of perfect, cf. (1). Tenses without imperfective morphology, i.a. the *passé composé* (that has both simple past and present perfect uses), are said to be banned there, as confirmed by the unacceptability of (2).

- (2) * Si on a réfléchi, on n' aurait pas signé.
If we have-PRST thought, we NEG have-COND NEG signed
'If we 'have thought' about it, we wouldn't have signed.'

I start from the observation that this empirical picture should be refined. One easily finds relevant occurrences of conditionals with a *conditionnel 2* in the consequent and a *passé composé* in the antecedent in corpora, cf. e.g. (3)-(6), all taken from the internet, and judged acceptable by my informants and myself.

- (3) Si un missile sol-air a effectivement été utilisé, il aurait été tiré
If a missile ground-air has indeed been used, it have-COND been launched
partir d' un bateau au large de Long Island.
from a boat at the coast of Long Island
'If a surface-to-air missile 'has indeed been' used, it would have been launched from a boat off the Long Island coast.'
- (4) Si vos vacances n'ont pas été réussies, il aurait été simple de venir
If your vacations NEG have NEG been succeeded, it HAVE-COND been simple to come
nous en parler.
PRN-DAT-1PL of-it speak

*This work is part of the B5 project 'Polysemy in a conceptual system' of the Collaborative Research Center 732 hosted by the University of Stuttgart. I would like to thank Hamida Demirdache, Atle Grønn, Hans Kamp and Christopher Piñón for discussion, as well as the reviewers and audiences of CLS 49 and LSRL 43 for their comments. None of them are responsible for the claims made in this paper. Many thanks to C. Piñón for his careful proofreading.

¹I nevertheless follow the tradition and call them PSCs despite of the absence of subjunctive morphology, because alternative names (*would have* conditionals, counterfactual conditionals...) are problematic, too.

'If your vacations 'haven't worked out' well, it would have been simple to come to us to speak about it.'

- (5) Si l'ancien proprio a tapé [le moyeu], il aurait ruiné la jante et
If the previous owner has hit [the hub], he have-COND destroyed the wheel and
il l'aurait donc changée, il y a un truc que je ne comprends pas.
he it have-COND therefore changed, there is a thing that I NEG understand NEG
'If the previous owner 'has hit' the hub, he would have destroyed the wheel and he would have therefore changed it, there is something I don't understand.'
- (6) Si le chef d'état-major a réellement tenu les propos rapportés par la presse,
If the Chief of Staff has really made the comments reported by the press,
il aurait commis un acte grave.
he have-COND committed a act serious
'If the Chief of Staff really 'has made' the comments reported by the press, he would have committed a serious act.'

The context of uses of these examples makes clear that they are not confined to a substandard variant of French, even if they are banished by some prescriptive grammars. Since conditionals like (3)-(6) mix the morphologies typical of PSCs and **past indicative conditionals (PICs)**, I call them '**swing**' PSCs. The paper is organized as follows. I show how swing PSCs differ from standard PSCs in section 2, and from PICs in section 3. Section 4 provides arguments for the claim that swing PSCs are well and truly a non-standard subtype of PSCs rather than a non-standard subtype of PICs. Section 5 points to the conclusions that swing PSCs enable one to draw on the respective role of tense/aspect morphology in the antecedent vs. the consequent of PSCs. Finally, in section 6, I briefly sketch two potential analyses of the way morphology contributes to the interpretation of swing PSCs.

2 Swing PSCs vs. standard PSCs

Swing PSCs differ from standard PSCs in at least two properties. Firstly, swing PSCs are systematically odd if the antecedent p or $\neg p$ follows from the context C (the set of worlds currently taken to be epistemically accessible by all participants, cf. [Stalnaker, 1978]): they require p to be *undecided* relative to C . This suffices to explain the problem of (2), since there, C most probably entails either p or $\neg p$. Also, if the examples in (3)-(6) are preceded by an assertion of $\neg p$, they become odd:

- (7) # L'ancien proprio n'a pas tapé le moyeu. S'il l'a tapé, il aurait
The previous owner NEG has NEG hit the hub. If he it has hit, he have-COND
ruiné la jante.
destroyed the wheel.
'The previous owner 'hasn't hit' the hub. If he 'has hit' it, he would have destroyed the wheel.'

By contrast, standard PSCs are, of course, unproblematic in a context where p is taken to be counterfactual, since standard PSCs regularly presuppose their antecedent as false.² Secondly,

²The 'subjunctive inference' of counterfactual antecedent falsity has been analysed as a presupposition, an implicature, or an antipresupposition. I adopt here von Fintel 1998's view that the counterfactual antecedent falsity is only one of the possible instantiations of the 'subjunctive inference' (see Section 4 below), and that this inference is a presupposition.

swing PSCs are typically used when p is contextually salient but not yet accepted nor rejected in the context C — p is on the Table/at issue ([Farkas and Bruce, 2010]). An evidence for this is the frequent presence in corpora of anaphorical adverbials like *effectivement/vraiment* ‘indeed/really’ in their antecedent. Asserting a swing PSC can therefore be seen as a way to address the issue p .

Interestingly, doing so through the assertion of a swing PSC projects a different *projected set* (ps) than through the assertion of a standard PSC. The projected set of an assertion characterizes the speaker’s proposal when she makes her assertion: it is the set of future common grounds relative to which the issue on the Table is decided ([Farkas and Bruce, 2010]). Accepting an assertion amounts to accepting its projected common grounds. (Note that an assertion adds on the Table not only its literal content but also its implicated content). Let me illustrate the difference between the projected set of standard PSCs and swing PSCs through the dialogue in (8)-(10). The crucial point concerns the way we interpret Marie’s confirmation (10) of Pierre’s reaction (9).

- (8) *Marie.* Peut-être que le Boeing 747 a été détruit par un missile.
‘Perhaps the Boeing 747 was destroyed by a missile.’
- (9) a. *Pierre.* S’il avait été détruit par un missile, il aurait été lancé par l’US Navy!
‘If it had been destroyed by a missile, it would have been launched by the US Navy!’
- b. *Pierre.* S’il a été détruit par un missile, il aurait été lancé par l’US Navy!
‘If it ‘has been’ destroyed by a missile, it would have been launched by the US Navy!’
- (10) *Marie.* Tu as raison.
‘You’re right.’

Both of Pierre’s reactions (9a) and (9b) have the same literal content $p \rightarrow q$. They also both presuppose that q is false or at least unlikely in the current (input) context C_1 . But they differ through the way they project p in the future common grounds, as reflected in the way we interpret Marie’s confirmation *Tu as raison* ‘You are right’. As an answer to (9a), (10) is easily understood as a confirmation of $p \rightarrow q$ but also of $\neg p$, because the rule of *modus tollens* is applied (implicated and literal contents are not kept apart here):

$$(11) \quad ps \text{ of (9a)} = \{C_1 \cup p \rightarrow q \cup \neg p\}$$

As an answer to (9b), (10) *cannot* be interpreted as an acceptance of the ps in (11). In fact, the reaction (9b) to Marie’s proposal (8) to add p to the future common grounds is inconclusive: (9b) feels like a *question* — an invitation to think more about what to do about p . More precisely, it invites one to choose between (i) rejecting p and (ii) challenging the presupposition $\neg q$ (or at least raising its probability from unlikely to likely) and accepting both p and q . In other words, through (9b), Pierre is suggesting: ‘Either you retract your proposal p , or here is the price to pay’ (namely, accepting the false/unlikely proposition q in the future common ground). The context state after a swing PSC is thus *inquisitive* wrt to p : its ps contains two future common grounds, and consequently does not help to settle the issue p :

$$(12) \quad ps \text{ of (9b)} = \{C_1 \cup p \rightarrow q \cup \neg p, C_1 \cup p \rightarrow q \cup p\}$$

By reacting through *Tu as raison* ‘You are right’, Marie only signals that she accepts the implication $p \rightarrow q$ and the imposed choice, *not* that she accepts one of the two alternatives. She can then go on after this acceptance and signal which future common ground she goes for (or signals she cannot choose neither of them):

- (10) Tu as raison...
...You are right.
- (i) *anti-conspiracist reaction* (Marie withdraws p and goes for $\neg p$)
...La théorie des missiles est après tout très improbable.
...You are right. The missile theory is after all very unlikely.
 - (ii) *conspiracist reaction* (Marie goes for p)
...Après tout ce n'est pas la première fois que l'US Navy est impliquée dans de tels désastres.
...After all it isn't the first time the US Navy is involved in such disasters.
 - (iii) *agnosticist reaction*
...C'est difficile de trancher.
...It is difficult to decide.

Note that the continuation (ii) would not be a felicitous way to assent to (9a), given that (9a) implicates that p is false.³

In sum, I have argued for the following points. Firstly, the **antecedent** p of swing PSCs is at issue/undecided relative to the (input) context C_1 and relative to the projected set. On this point, swing PSCs differ from standard PSCs, which typically presuppose that p is false. Secondly, swing PSCs resemble standard PSCs in that they typically present their **consequent** q as false or unlikely.⁴

3 Swing PSCs vs. past indicative conditionals (PICs)

Swing PSCs also differ from PICs in three respects. Firstly, PICs can sometimes be used as a rhetorical device when p follows from C , as observed by e.g. [Dancygier, 1998], cf. (13). This is not possible with swing PSCs, cf. (14).

- (13) Il a plu. S'il a plu, le match a été annulé.
It has rained. If it has rained, the match has been cancelled
'It 'has rained'. If it 'has rained', the match 'has been' cancelled.'
- (14) # Il a plu. S'il a plu, le match aurait été annulé.
It has rained. If it has rained, the match have-COND been cancelled
'It 'has rained'. If it 'has rained', the match would have been cancelled.'

The second difference concerns past conditionals *à la* Anderson ([Anderson, 1951]). Andersonian PSCs are illustrated in (15). They are used to argue for the truth of p . As Anderson emphasizes, the existence of such conditionals shows that PSCs do not systematically presuppose that their antecedent is false.

- (15) If Jones had taken arsenic, he would have shown just exactly those symptoms which he does in fact show. [So, it is likely that he took arsenic.]

It is well-known that Andersonian PICs (e.g. (16)) are odd (cf. [von Fintel, 1998] for an account in terms of uninformativeness). This is also true in French, cf. (17):

³On the contrary, the continuation (iii) would be felicitous as a way to assent to (9a), because (iii) can then mean that Marie does not want to choose between her initial proposal to adopt p and Peter's suggestion to adopt $\neg p$.

⁴Except in Andersonian cases, as we will see in the next section.

- (16) # If Jones took arsenic, he shows exactly those symptoms which he does in fact show.
- (17) # Si John a pris de l'arsenic, il a montré exactement les symptômes qu'il
If John has taken of-the arsenic, he has shown exactly the symptoms that he
a maintenant.
has now

'If John 'has taken' arsenic, he 'has shown' exactly the symptoms that he has now.'

By contrast, Andersonian swing PSCs are natural, as shown by the acceptability of (18). We can account for it the same way von Fintel 1998 explains the acceptability of the subjunctive in (15) if we admit that swing PSCs are a subvariant of PSCs, cf. the next section.

- (18) Si John a pris de l'arsenic, il aurait montré exactement les symptômes
If John has taken of-the arsenic, he have-COND shown exactly the symptoms
qu'il a maintenant.
that he has now

'If John 'has taken' arsenic, he would have shown exactly the symptoms that he has now.'

The third difference between PICs and swing PSCs is that except in Andersonian cases, the latter tend to presuppose that their consequent *q* is false/unlikely in *C*. This is not the case of PICs. Let us, for instance, compare the previous example (7) above with its (shortened) PIC variant (19):

- (19) Si l'ancien proprio a tapé [le moyeu], il a ruiné la jante et
If the previous owner has hit [the hub], he have-INDIC destroyed the wheel and
il l'a donc changée.
he it have-PRST therefore changed

'If the previous owner 'has hit' the hub, he 'has destroyed' the wheel and he therefore
'has changed' it.'

In example (7), the continuation 'there is something I don't understand' makes clear that the speaker can hardly believe *q* to be true (*p* has previously been proposed in one of the previous posts of the same forum), and thereby suggests that adopting *p* in the future common ground has a certain cost (namely, adopting the unlikely proposition *q*). This justifies the use of the *conditionnel 2*. The PIC variant (19) would make a very different contribution: it simply makes the consequence *q* of the proposal *p* under discussion explicit, without presenting this consequence as unlikely. In favour of the claim that swing PSCs differ from PICs in that they presuppose *q* as false/unlikely, one observes that in some of the examples of section 1, replacing the *conditionnel 2* by the corresponding past indicative (i.e. the *passé composé*) brings about an inappropriate variant, precisely because *q* is taken to be false in the context of the original example. Compare e.g. the previous example (4) with its PIC variant (20). Example (4) was found in a forum of a vacation club; the organizer replies to a client complaining about his stay.

- (20) # Si vos vacances n'ont pas été réussies, il a été simple de
If your vacations NEG have NEG been succeeded, it have-INDIC been simple to
venir nous en parler.
come PRN-DAT-1PL of-it speak

'If your vacations 'haven't worked out' well, it 'has been simple' to come to us to speak
about it.'

The predicate *être simple de* ‘be simple to’ used in the consequent *q* resembles French ‘enough constructions’ studied in e.g. [Hacquard, 2006] in that it entails the truth of its complement with the *passé composé*, but not with imperfective morphology.⁵ Given the *passé composé* in its consequent *q*, (20) leaves open the possibility that the complement of *être simple de* is true in *C*. This clashes with the context of the original example (4), where it is taken for granted that the client did not previously contact the vacation club.

4 Swing PSCs are subjunctive conditionals

At this point, one might still wonder why one should endorse my claim that swing PSCs are subjunctive rather than (a strange subkind of) indicative conditionals. After all, they do not implicate that *p* is false, and their morphology only partly matches the one of PSCs.

I adopt here [von Fintel, 1998]’s view according to which the difference between PSCs and PICs mainly lies in the kind of domain ($D(w)$) the conditional quantifies over. According to von Fintel and others, the natural default pragmatic constraint on quantification over worlds performed by conditionals is that $D(w)$ is entirely in *C*. The indicative being *unmarked*, it does not signal anything against this constraint $D(w) \subseteq C$. The subjunctive is *marked* and indicates a violation: SCs presuppose that $D(w)$ is partly outside *C* ($D(w) \not\subseteq C$). This explains why standard PSCs are used when the antecedent *p* is taken to be counterfactual. But it also explains why we find PSCs when $D(w)$ needs to be widened for some other reason, for instance if *p* and *q* follow from *C* but $D(w)$ contains $\neg q$ -worlds, as in von Fintel’s analysis of Andersonian PSCs. The facts described above allow to conclude that swing PSCs are well and truly PSCs: their $D(w)$ contains either counterfactual/implausible *q* worlds (cf. e.g. (3)-(6)), or counterfactual/implausible $\neg q$ worlds (cf. the Andersonian swing PSC (18)).

5 The role of the imperfective

One of the interests of swing PSCs is that they allow one to better tease apart the semantic contribution of aspect/tense morphology in the antecedent and the consequent of PSCs. Their properties point to the two following conclusions. Firstly, the ‘subjunctiveness’ of the conditional (that we equate with $D(w) \not\subseteq C$) directly depends on the *conditionnel 2* morphology in the consequent, common to swing and standard PSCs, rather than on the layer of imperfective morphology in the antecedent. This is additionally confirmed by the fact that one cannot obtain swing PSCs by combining a *plus que parfait* in the antecedent and a non-*conditionnel* indicative morphology in the consequent. Sentences of this type are either out, cf. (21a), or force a temporal interpretation of the *plus que parfait* and are PICs, cf. (21b):

- (21) a. *Si on avait réfléchi, on n’ a pas signé. (ill-formed PSC)
If we have-IMP thought, we NEG have NEG signed
‘If we had thought about it, we haven’t had signed.’
- b. S’ il l’ avait vue la veille, il lui a raconté l’histoire.
If he her see-PQP the day before, he her has told the story
‘If he had seen her the day before, he told her the story.’

⁵In fact, differently from ‘enough’ constructions, *être simple de P* even presupposes the truth of *P* with the *passé composé*; but this difference between the ‘enough’ constructions and *être simple de* is irrelevant here, because this presupposition is not projected in the consequent of conditionals. For instance, *S’il n’a pas pris son téléphone, alors il n’a pas été simple de lui parler* ‘If he ‘hasn’t taken’ his phone, then it ‘has not been simple’ to speak with him’ does not presuppose that one spoke with him.

In other words, French swing PSCs suggest that it is the tense/aspect marking in the *consequent* that is decisive for the subjunctivehood/counterfactuality of the conditional. Secondly, the properties of swing PSCs described above allow one to conclude that the presupposition of ‘counterfactual *antecedent* falsity’ regularly triggered by PSCs directly depends, in French, not only on the *conditionnel 2* in the consequent, but also on the imperfective morphology in the antecedent. Given that the *conditionnel* can be analysed as the morphological spell-out of the imperfective plus the future ([Iatridou, 2000]), this is compatible with the view that in French, imperfectivity in the consequent **and** the antecedent is necessary to signal *counterfactual antecedent falsity*, rather than counterfactuality *per se*, also found with swing PSCs which regularly present their *consequent* as counterfactual/unlikely.

6 Analysis of tense/aspect morphology in swing PSCs

I still have to explain how tense/aspect morphology in swing PSCs should be analysed, and what the lack of the expected imperfective morphology in their antecedent indicates. I will briefly and very roughly sketch two potential analyses.

Analysis 1. Let us first look at the role of tense/aspectual morphology in their *consequent*. The easiest way to look at it consists in simply extending previous analyses of the morphology in the consequent of standard PSCs to swing PSCs. According to ‘past-as-past’ approaches of standard PSCs ([Ippolito, 2003], [Arregui, 2005]), PAST does not localize the described eventualities, but rather contributes to the interpretation of the modal. Under some of these analyses, the past tense morpheme in the main clause is used to go back to a time where the proposition could still be true. A way to implement this is to have the past tense outscope the modal (NOW PAST(MOD($p \rightarrow q$))). PAST has been said to be provided by *would* in the matrix clause, cf. e.g. [Arregui, 2009]. (In French, the imperfective morphology *-ai-* in the *conditionnel* is the correspondant of the past morphology in *would*.) I follow [Grønn and von Stechow, 2011] who argued that the shift towards the past is not done by *would* (or *-ai-* in French), but rather by the auxiliary *have* (*avoir* in French). Since the relevant past possibility is no longer available at utterance time, counterfactuality can then be pragmatically derived ([Condoravdi, 2002]). In principle, these proposals can be extended to the main clause of swing PSCs.

Tense morphology *in the antecedent* of standard PSCs has been analysed as a case of (sequence of tense) agreement with the past tense in the matrix clause ([von Fintel, 1998], [Arregui, 2005], [Anand and Hacquard, 2009]). For French, agreement is only partial, since the imperfective (*-ai-*) but not the future morphology (*-r-*) is present in the antecedent of standard PSCs. But Anand and Hacquard observe that the agreement is complete in Québécois French, where both the antecedent and consequent show conditional morphology. Also relevant is the fact that conditional morphology typically appears in the antecedent of conditionals in Child French. According to Analysis 1, swing PSCs can then be conceived as a case where agreement *fails* to hold. I propose that through this agreement failure, the speaker wants to indicate that subjunctivehood is obtained through another way than the counterfactuality of *p*, i.e. that *it is not* because *p* is counterfactual that D(w) reaches outside of the context set, but rather through the counterfactuality/unlikeliness of *q*. This may serve a diplomatic purpose, if *p* has been put on the Table by another participant to the discourse.

Analysis 2. According to a second potential analysis of swing PSCs, *avoir* provides a past/perfect used to locate the described eventualities, in the scope of MOD (we then have NOW MOD (HAVE-*p* → HAVE-*q*)). We then predict an absence of shift in the temporal reference of the antecedent. Since the possibility is still open at NOW, we expect not to derive counterfactuality *stricto sensu*. But these past conditionals are still expected to be subjunctive conditionals, and

therefore to indicate a greater uncertainty wrt the past/perfect propositions expressed in their clauses than the corresponding indicative conditionals.

That swing PSCs are acceptable, as we saw with (4), in a context where q is taken to be *false* (rather than simply unlikely) *prima facie* militates against Analysis 2. However, one can observe that even PICs are in fact not so unacceptable in a context where q is taken to be counterfactual (contrary to what I concluded earlier from (20)), cf. the example below.

- (22) Pierre ne lui a pas téléphoné. S'il a eu un accident, il lui a
 Pierre NEG to-him has NEG called. If he has had an accident, he to-him has
 téléphoné.
 called
 'Pierre didn't call her. If he had an accident, he *called* her.'

Concluding that data like (4) invalidate Analysis 2 might then be too hasty. I believe that the analysis of swing PSCs with future adverbials (or ‘mismatched’ swing PSCs, cf. [Ippolito, 2003]) might help to see what it the right strategy to pursue. I leave this for future research.

References

- [Anand and Hacquard, 2009] Anand, P. and Hacquard, V. (2009). The Role of the Imperfect in Romance Counterfactuals. In Prinzhorn, M., Schmitt, V., and Zobel, S., editors, *Proceedings of Sinn und Bedeutung 14*, pages 37–50. University of Vienna, Vienna.
- [Anderson, 1951] Anderson, A. (1951). A note on subjunctive and counterfactual conditionals. *Analysis*, 12:35–38.
- [Arregui, 2005] Arregui, A. (2005). *On the accessibility of possible worlds: The role of tense and aspect*. PhD thesis, University of Massachusetts at Amherst.
- [Arregui, 2009] Arregui, A. (2009). On similarity in counterfactuals. *Linguistics and Philosophy*, 32:245–278.
- [Condoravdi, 2002] Condoravdi, C. (2002). Temporal interpretation of modals: Modals for the present and for the past. In Beaver, D., Kaufmann, S., CLark, B., and Casillas, L., editors, *The Construction of Meaning*, pages 59–88. CSLI, Stanford.
- [Dancygier, 1998] Dancygier, B. (1998). *Conditionals and Prediction: Time, Knowledge, and Causation in Conditional Constructions*. Cambridge University Press, Cambridge.
- [Farkas and Bruce, 2010] Farkas, D. and Bruce, K. (2010). On Reacting to Assertions and Polar Questions. *Journal of Semantics*, 27:81–118.
- [Grønn and von Stechow, 2011] Grønn, A. and von Stechow, A. (2011). Tense in subjunctive conditionals across languages. Ms, University of Oslo and University of Tuebingen.
- [Hacquard, 2006] Hacquard, V. (2006). Aspects of too and enough constructions. In Georgala, E. and Howell, J., editors, *Proceedings of SALT 15*, pages 20–97. CLC Publications, Ithaca, NY.
- [Iatridou, 2000] Iatridou, S. (2000). The grammatical ingredients of counterfactuality. *Linguistic Inquiry*, 31:231–270.
- [Ippolito, 2003] Ippolito, M. (2003). Presuppositions and implicatures in counterfactuals. *Natural Language Semantics*, 11:145–186.
- [Stalnaker, 1978] Stalnaker, R. (1978). Assertion. In Cole, P., editor, *Syntax and Semantics 9: Pragmatics*, pages 315–332. Academic Press, New York/San Francisco/London.
- [von Fintel, 1998] von Fintel, K. (1998). The Presupposition of Subjunctive Conditionals. In Sauerland, U. and Percus, O., editors, *The interpretive tract (MIT Working Papers in Linguistics 25)*, pages 29–44.

An Analysis of Quantifier Scope Restrictions in Dependence Logic

Ralf Naumann and Wiebke Petersen

University of Düsseldorf
`{naumann,petersen}@phil.uni-duesseldorf.de`

Abstract

In our contribution we will present a strictly semantic approach to scope phenomena which is based on (Dynamic) Dependence Logic (DL) [Vää07, Gal13] and ideas from Dynamic Semantics [Fer93]. Formulas as well as quantifiers are interpreted as relations between sets of assignments. Scopal ambiguities are analyzed as a form of non-determinism: processing a formula in an input context X can lead to different output contexts Y . These outputs are constraint by imposing dependence relations on Y .¹

1 Introduction: the data

One of the most serious challenges for formal theories of (quantified) NPs in natural language is the scopal behaviour of singular indefinites like ‘a’ or ‘some’ as well as numerical indefinites like ‘two’ or ‘three’ (see [WR11] for a recent survey). They not only give rise to wide-scope (WS) and intermediate scope (IS) readings (‘upward unboundedness’) but they can in addition escape both clausal and island boundaries, though their scope need not be maximal, resulting in the phenomenon of scopal ambiguity [Sza10, BF11].

- (1) a. Every man loves a woman.
 b. There is a woman who is loved by every man.
- (2) a. If some relative of mine dies I will inherit a house.
 b. [some relative of mine]_i [if x_i dies, I will inherit a house]
- (3) a. A referee read three abstracts.
 b. There are three abstracts and each of them was read by a referee.

Example (4) shows that such indefinites can even have intermediate scope if the sentence contains three quantifier phrases.

- (4) a. Every student read a_x paper that every_y professor recommended.
 b. Every _{x} student read every_y paper that a_z professor recommended.

Whereas in (4a) ‘a paper’ can semantically scope over ‘every student’, the universal ‘every professor’ in the relative clause cannot scope over the indefinite ‘a paper’. By contrast, for (4b) the indefinite ‘a professor’ in the relative clause can scope over both universals in the matrix clause, resulting in the following three readings where ‘ $>$ ’ indicates scopal order: (i) $x > y > z$ (narrow scope NS); (ii) $x > z > y$ (intermediate scope IS) and (iii) $z > x > y$ (wide scope WS). However, in simple sentences in English a universal NP in object position can scope over an indefinite in subject position, witness example (5) (for this sentence, the inverse scope reading is even more prominent than the direct scope reading [WR11]).

¹The research was supported by the German Science Foundation (DFG) funding the Collaborative Research Center 991.

- (5) Some inhabitant of every city participated.

'Upward unboundedness' is not valid for all indefinites: modified numerals like 'at least n', 'fewer than n' or 'exactly n' do not allow such a WS reading, as shown by the following examples.

- (6) a. Two referees read more than three abstracts.
 b. *There are more than three abstracts such that each of them was read by two (possibly different) referees.
- (7) a. At least three men read few books, possibly different ones.
 b. *Few books are such that at least three (possibly different) men read them.

Second, a distinction between existential and distributive scope has to be made ([WR11], [Sza10, 92f.]).

- (8) If three relatives of mine die, I will inherit a house.

(8) can only have the interpretations that (i) if the number of my dead relatives reaches three, I inherit a house (whichever relatives of mine passes away) or that (ii) there are three particular relatives such that if they all die, I will inherit a house. On both interpretations the total of houses inherited is one and not three, as expected on a distributive reading: there are three relatives of mine such that if any of them dies I will inherit a house (a total of three houses).

In the literature many of the above examples are analyzed using choice functions, i.e. 0-ary Skolem functions. For example, (2) can be analyzed as (9).

- (9) $\exists F_{<0>} [\text{if } F_{<0>}(\lambda x x \text{ is a relative of mine}) \text{ dies, I will inherit a house}]$

Instead of using a variable over elements of the domain, a higher-order variable over choice functions with maximal scope is used [Sza10, 93], [Sch06, 300]. As shown in [Sza10, Sch06, Chi01], this mechanism is in general not expressive enough. [Chi01] observes that (10a) can be denied by using (10b) so that the quantifier over a choice function does not have widest scope.

- (10) a. Every linguist studied every solution that some problem might have.
 b. Not every linguist studied every solution that some problem might have.

A second problem is discussed in [Sza10, 95], showing that there is an inseparability of existential scope and the property denoted by the NP. If (11a) is analyzed as (11b), using material implication, a choice-functional analysis makes wrong predictions because the existence of a non-invited philosopher makes (11a) true.

- (11) a. If we invite a certain philosopher, Max will be offended.
 b. $\exists F[\text{invite}(F(\text{philosopher}), \text{we}) \rightarrow \text{offended}(\text{m})]$

2 Scope ambiguity as differences in dependence relations

In traditional approaches the 'upward unboundedness' of indefinites is explained in terms of linear quantifier scope relations. This is achieved either by using type-shifting rules, the effect of which consists in changing the linear order of quantifiers or by interpreting indefinites in situ by generalized Skolem functions in the lexicon and thereby not as 'bona fide' quantifiers [BF11].

According to [Sza97, 109], theories of scope which are based on mechanisms like quantifier raising, storage or type shifting are *semantically blind* because they use a single *syntactic* rule of scope assignments: $\alpha[D\dots\beta\dots] \rightarrow \alpha$ scopes over β . According to this rule, an expression α is

‘prefixed’ to a domain D in order to assign scope to it over D , irrespective of what α means and of what operator β may occur in D .

In contrast to those approaches, we will develop a strictly semantical approach which builds on the following assumptions: (i) formulas are interpreted as (possibly dynamic) relations between sets of assignments, (ii) following Dynamic Dependence Logic [Gal13] and [Fer93], quantifier expressions too are interpreted as relations between sets of assignments. On this perspective, two types of relational properties can be distinguished which are related to the cardinality information expressed by the quantifier and the notion of spawning a set of processes (reachability sets in the sense of [Fer93]). either conjunctively or disjunctively, (iii) different types of relational properties of a quantifier give rise to different forms of dependencies, which can be expressed in Dynamic Dependence Logic and (iv) a quantifier imposes these relational properties as *strategies* in the sense of Game Theory (Logic), which are formulated in terms of various dependence relations from Dynamic Dependence Logic and Database Theory.

Using such a dynamic formalism makes it possible to express dependence relations that hold either between variables (or attributes) of a single relation or between combinations of relations. Non-determinism can arise at the level of a single relation (inverse scope readings) or at the level of combinations of relations (IS readings).

In our semantic-driven approach verb stems are interpreted as sets of events. Each event denoted by a verb stem is related to a tuple of objects which can be singled out, e.g., in terms of thematic roles like actor or theme. On this perspective, the elements of a team are events and the values of the attributes are objects involved in the events so that each verb determines a team at the semantic level. Given a set of events of the same type (say readings), there can be various dependency relations between the roles defined by events of that type. For example, in case of a set of reading events, at least three different dependency relations can be distinguished. By way of illustration, assume the set of students to be $\{s_1, s_2\}$ and the set of papers read to be $\{p_1, p_2\}$. Consider first the sentence ‘Every student read a paper’ with quantifier scope $\forall\exists$.

(12) Every student read a paper.

For \exists , there are two different strategies: either \exists is assigned a constant value or not. If the first strategy applies, this corresponds to a WS reading, which is shown in Table 1(a). In Dependence Logic, this is a special case of functional dependence, which can be expressed by $=t$ (see definition (21b) below) if $t = \text{theme}$ (or ‘paper’ in the present case) because the values of the attribute THEME are constant with respect to the values of the attribute ACTOR.

(a) WS	(b) NS1	(c) NS2
$F^{\text{paper}}(s_1) = \{p_1\}$	$F^{\text{paper}}(s_1) = \{p_1, p_2\}$	$F^{\text{paper}}(s_1) = \{p_1, p_2\}$
$F^{\text{paper}}(s_2) = \{p_1\}$	$F^{\text{paper}}(s_2) = \{p_2\}$	$F^{\text{paper}}(s_2) = \{p_1, p_2\}$

Table 1: Different combinations for the scope sequence $\forall\exists$

The other two tables are instances of non-(strict-)functional dependencies, corresponding to a NS-reading. These two examples differ with respect to a weak form of dependence, which can be expressed using the tensor operator.

(13) $M \models_{(X,Y)} \phi \otimes \psi$ iff there are teams X_1, X_2 s.t. $X = X_1 \cup X_2, M \models_{(X_1,Y)} \phi$ and $M \models_{(X_2,Y)} \psi$.

In (13) the following additional constraint on the subteams X_1 and X_2 must be imposed: if the variables (attributes) of the team X are x_1, \dots, x_n (in that order), then the values for x_1 must

be the same for X , X_1 and X_2 . The idea behind using the tensor is that a team can be split into n subteams such that for each subteam a FD between the attributes ‘student’ and ‘paper’ holds. This form of dependence will be called *weak functional dependence* (wFD). Table 1(c) is an example of a wFD. This table also shows that the scope sequence $\forall\forall$ is an instance of wFD. Consider next a sentence with surface order $\exists\forall$.

- (14) A student read every paper.

For (14), the same argument applies as for (12). If for \exists a strategy is chosen which assigns to the corresponding argument a constant value, one gets the surface scope order $\exists\forall$, otherwise this yields the WS reading corresponding to $\forall\exists$.

Using the above examples, one arrives at the following preliminary thesis: *a WS reading is already possible if a quantifier Q admits both of a constant and a non-constant strategy (i.e. it shows (weak) variation).* This thesis is too weak because it also holds for quantifiers like ‘at least’, which do not admit of a WS reading in object position.

- (15) Every student read at least two papers.

Suppose there are four students $\{s_1, s_2, s_3, s_4\}$ and three papers $\{p_1, p_2, p_3\}$. A possible distribution between students and papers is given in (16).

- (16)
- a. $F^{paper}(s_1) = \{p_1, p_2\}$
 - b. $F^{paper}(s_2) = \{p_1, p_3\}$
 - c. $F^{paper}(s_3) = \{p_2, p_3\}$
 - d. $F^{paper}(s_4) = \{p_1, p_2, p_3\}$

What is needed, therefore, is an additional requirement on the (non-)determinism of the cardinality information.

- (17)
- a. $\forall : \underbrace{\alpha \otimes \dots \otimes \alpha}_{n-times}$ if $card([[N]]) = n$
 - b. *at least n* : $\underbrace{\alpha \otimes \dots \otimes \alpha}_{\geq n-times}$. The number of subteams is not determined in order to fix a unique dependence relation.

The upshot of the above discussion is that two types of non-determinism must be distinguished: (i) non-determinism with respect to the cardinality, called *cardinality (non-) determinism* and (ii) non-determinism with respect to the object assigned to an attribute (\exists vs. *john*, random vs. non-random assignment), called *value (non-) determinism*. Cardinality (non-) determinism can be defined using wFD. In this case the cardinality of $F^x(y_i)$ must be the same for each y_i . Value (non-) determinism can be defined as FD. In this case only outputs are admissible for which the value of the variable (attribute) of the quantifier, say x , is constant so that $=x$ holds. Using these two types of non-determinism, one can define the notion of *weak variation* as follows.

- (18) A quantifier Q shows *weak variation* iff it satisfies (i) cardinality determinism and (ii) value determinism.

Using this definition of weak variation, one arrives at the following three theses: (i) *The surface order of quantifiers is always admissible*, (ii) *if the relational properties of a quantifier Q admit of weak variation, a (weak or strong) FD is an admissible strategy so that an inverse (WS) reading is possible* and (iii) *the set of dependence relations that can hold between two*

quantifiers is dependent on the relational properties of those quantifiers in terms of cardinality (non-)determinism and value (non-)determinism.

2.1 Intermediate readings and syntactic islands

The theory outlined so far is unable to account for the difference in scopal behaviour between indefinites and universal quantification.

- (19) a. Every_x student read a_y paper that every_z professor recommended.
b. Every_x student read every_y paper that a_z professor recommended.

The following two problems arise: (i) unlike the cases considered so far, there is no weak FD for the IS-reading of (19b). Second, for both (19a) and (19b) an MVD is possible (see below (21c) for the definition). In order to solve these problems one has to take into consideration that in both cases *two* relations are combined: the ‘read’ and the ‘recommend’ relation.

Beginning with example (19a) and assuming a set of two students $\{s_1, s_2\}$, a set of two papers $\{p_1, p_2\}$ and three professors $\{t_1, t_2, t_3\}$, one gets for the function $F^t(s_i, p_j) : F^t(s_i, p_j) = \{t_1, t_2, t_3\}$. For (19a), $F^t(s_i, p_j)$ is therefore always the whole set of professors so that no variation is possible. By contrast, for (19b) there are three different patterns, each corresponding to one of the three admissible scope readings, yielding the three patterns shown in Table 2.

(a) NS	(b) WS	(c) IS
$F^t(s_1, p_1) = \{t_1\}$	$F^t(s_1, p_1) = \{t_1\}$	$F^t(s_1, p_1) = \{t_1\}$
$F^t(s_2, p_1) = \{t_2\}$	$F^t(s_1, p_2) = \{t_1\}$	$F^t(s_1, p_2) = \{t_1\}$
$F^t(s_1, p_2) = \{t_3\}$	$F^t(s_2, p_1) = \{t_1\}$	$F^t(s_2, p_1) = \{t_2\}$
$F^t(s_2, p_2) = \{t_1\}$	$F^t(s_2, p_2) = \{t_1\}$	$F^t(s_2, p_2) = \{t_2\}$

Table 2: Different combinations for the surface scope sequence $\forall\forall\exists$

The example in Table 2(a) corresponds to a NS reading. In this case there is no dependence between papers and professors. The second example (Table 2b) shows a WS reading for ‘professor’ since the value for the attribute is constant. The third example (Table 2c) corresponds to an IS reading. The t_i are distributed in such a way that they are constant with respect to the first (student) argument. In our theory this difference between \forall and quantifiers corresponding to indefinites can be explained as follows. First, the following thesis about the join of two relations is assumed, using the notion of *totality* defined below. The second condition in (20) excludes \forall because it always denotes the top element of the lattice of plural objects using a Link-style representation.

Join of two relations: In the composition of two relations R_1 and R_2 a quantifier Q occurring in R_2 is *upward unbounded* only if (i) it admits of weak variation, (ii) there is a MVD between Q and a different quantifier occurring in R_1 and (iii) this MVD is not *total*.

- (20) A quantifier Q is *total* iff (i) Q shows weak variation and (ii) the output Y is cardinality non-deterministic.

2.2 An outline of the formal theory

Dependence Logic, DL, [Vää07, Gal13] extends the syntax of FOL with dependence formulas which make it possible to express dependence relations between variables. E.g. the formula

$= (x, y)$ has the intuitive meaning that y is determined by x . Such an assertion only makes sense for sets of assignments and not for a single assignment. Therefore, in DL formulas are interpreted w.r.t. sets of assignments which are called *teams*.

[Gal13] observes that DL can be thought of as a logic of imperfect information and shows that it is possible to develop a dynamic variant, called *Dynamic Dependence Logic (DDL)*, in which expressions are interpreted as transitions between teams. This variant is based on the following two observations. First, in DL quantifiers are already implicitly interpreted as transitions $X \rightarrow \Phi_M$ where Φ_M is a local quantifier. In case this quantifier possibly leads to different outputs, one gets what is called a spawning of processes (transitions) or events running in parallel ([Fer93]). Second, $\phi \sqcup \psi$ can be interpreted as the non-deterministic choice between doing either ϕ or ψ . In DDL the satisfaction relation \models is defined as a relation between pairs of teams, a DL formula and a model. FOL literals and dependence atoms are interpreted as tests. The difference between FOL literals like $R(x_1, \dots, x_n)$ and a dependence atom like $= (x, y)$ is that the former are interpreted in the usual Tarskian sense. Multi-valued dependence $\Rightarrow (x, y)$ is defined in (21c), whereas sequential composition ‘;’ is defined in (21d).

- (21) a. $M \models_{(X,Y)} R(x_1, \dots, x_n)$ iff $X = Y$ and for all $s \in X : M \models_s R(x_1, \dots, x_n)$ where \models_s is the satisfaction relation in the usual (Tarskian) sense.
- b. $M \models_{(X,Y)} = (x, y)$ iff $X = Y$ and for all $s, s' \in X$ if $s(x) = s'(x)$ then $s(y) = s'(y)$.
- c. $M \models_{(X,Y)} \Rightarrow (\bar{x}, \bar{y})$ iff $X = Y$ and for all $y \in \bar{y}$, F^y only depends on the values of \bar{x} where \bar{x} and \bar{y} are sequences of variables.
- d. $M \models_{(X,Y)} \phi; \psi$ iff there exists a team Z s.t. $M \models_{(X,Z)} \phi$ and $M \models_{(Z,Y)} \psi$.
- e. $\phi \sqcup \psi := \exists x_1 \exists x_2 (= (x_1) \wedge (= (x_2) \wedge ((x_1 = x_2 \wedge \phi) \otimes (x_1 \neq x_2 \wedge \psi))))$.

Constancy of the value of a variable x is expressed by the dependence formula $= (x)$. The formula $=()$ holds for all teams. Below the definitions for the quantifiers \exists , \forall and $\exists^{\geq 5}$ are given (see [Eng12], [Gal13] for details).

- (22) a. $M \models_X \exists x \phi$ iff there is a function $F : X \rightarrow \exists_M$ such that $M \models_{X[F/x]} \phi$, where \exists_M is the local existential quantifier defined by $\{A \subseteq M \mid A \neq \emptyset\}$ and $X[F/x]$ is the team $\{s[a/x] \mid s \in X, a \in F(s)\}$.
- b. $M \models_X \forall x \phi$ iff there is a function $F : X \rightarrow \forall_M$ such that $M \models_{X[F/x]} \phi$, where \forall_M is the local universal quantifier defined by $\{A \subseteq M \mid A = M\}$.
- c. $M \models_X \exists^{\geq 5} x \phi$ iff there is a function $F : X \rightarrow \exists_M^{\geq 5}$ such that $M \models_{X[F/x]} \phi$, where $\exists_M^{\geq 5}$ is the local cardinality quantifier defined by $\{A \subseteq M \mid |A| \geq 5\}$.

According to (22a), non-determinism is built into the definition of the existential quantifier because it allows the choice of an arbitrary non-empty set of witnesses for an existentially quantified variable. This is called *lax* semantics in DL. The definitions in (22) are static because a quantifier is interpreted with respect to a team and not a pair of teams. In (23) the definition of the dynamic variant of \exists is given.

- (23) $M \models_{(X,Y)} \exists x$ iff there is a function $F \rightarrow \exists_M$ s.t. $X[F/x] \subseteq Y$.

The formula $\exists x \phi$ from DL is translated into DDL by $\exists x; \phi'$ with ϕ' the translation of ϕ . $\exists x; \phi'$ requires the whole output team Y to be input to ϕ' . On this interpretation of a quantifier, the differences with respect to the relational properties expressible in terms of cardinality- and value (non-) determinism are not taken into account. Such constraints can be imposed by using the dependence formulas corresponding to the particular type of (non-) determinism. In (24a),

the general case is given, whereas (24b) holds for \exists .

- (24) a. $\Phi x; (\psi_1 \sqcup \dots \sqcup \psi_n); \phi'$.
b. $\exists x; (=x) \sqcup =(); \phi'$.

In (24a), Φ is a quantifier and ψ_i , $1 \leq i \leq n$, is a dependence formula. According to (24b), in the case of \exists (non-deterministically) either the value of the newly introduced element is chosen to be constant, $=x$, or no constraint is imposed, $=()$. On this approach, $(\psi_1 \sqcup \dots \sqcup \psi_n)$ is part of the interpretation of a quantifier, reflecting the possible choices of dependence relations it admits. Therefore, a quantifier (possibly) triggers a non-deterministic choice of how the values of variables introduced by it are determined in the interpretation of an expression in which it occurs. Different choices correspond to what is analyzed as a difference in the linear order of quantifiers in other approaches.

Sentences like (4a) and (4b) are analyzed as the \cap -combination of two relations R_1 (say ‘read’) and R_2 (say ‘recommend’). For the sake of simplicity, \cap is defined only for two 2-place relations with $\phi = R_1(x, y)$ and $\psi = R_2(z, y)$.

- (25) $M \models_{(X, Y)} \phi \cap \psi$ iff there are Y_1 and Y_2 s.t. (i) $M \models_{(X, Y_1)} \phi$, (ii) $M \models_{(X, Y_2)} \psi$ and (iii) $(Y = \{\langle a_x, b_y, c_z \rangle \mid \langle a_x, b_y \rangle \in Y_1 \text{ and } \langle c_z, b_y \rangle \in Y_2\}$ and there is a FD between the two relating variables, here x and z).

According to (25), Y is the join of two binary relations. Since both relations are (possibly) non-deterministic, the output team Y is not uniquely determined. The output team Y can be classified according to the FD’s and MVD’s that are possible according to the surface order of quantifiers. Consider the surface order $\forall \exists$ (= (19b)). For R_1 (= ‘read’), one has (26).

- (26) $\{(M, R_1) \mid \exists A \in Q_1, \exists B \in Q_2, A \times B \subseteq R_1\}$.

Thus, the output team for R_1 simply is $M_{\text{student}} \times M_{\text{paper}}$, where M_{student} and M_{paper} are the restrictions of the domain to students and papers, respectively. As a consequence, R_1 is a total relation between M_{student} and M_{paper} . For R_2 , one has the three principal possibilities depicted in Table 3.

(a) WS $F^{\text{professor}}(p_1) = \{t_1\}$ $F^{\text{professor}}(p_2) = \{t_1\}$	(b) NS1 $F^{\text{professor}}(p_1) = \{t_1, t_2\}$ $F^{\text{professor}}(p_2) = \{t_1, t_3\}$	(c) NS2 $F^{\text{professor}}(p_1) = \{t_1, t_2\}$ $F^{\text{professor}}(p_2) = \{t_1, t_2\}$
--	---	---

Table 3: Different combinations for *A professor recommended a paper*

If the output team of R_2 satisfies the FD $=(\text{professor})$, Table 3(a), i.e. the value of the variable *professor* is constant, one gets a WS reading for the corresponding quantified NP ‘a professor’. If R_2 does not impose the constraint $=(\text{professor})$, two principle cases have to be distinguished, which are depicted in Table 3(b) and Table 3(c). If the output team is like that in Table 3(b) (NS1), one gets a NS reading, corresponding to Table 2(a). By contrast, if the output team is like that in Table 3(c), one gets the IS reading depicted in Table 2(c) if clause (iii) in (25) is used with $= (x, z)$ s.t. $F(s_i) = t_i$, with F a function expressing the relation between students and professors. Thus, for the join of two relations, the different scope possibilities simply follow from the strategies imposed by the quantifiers occurring in the relations R_1 and R_2 and the way these relations are combined by \cap .

In our theory the relational properties of quantifiers depending on both cardinality- and value (non-) determinism give rise to different functional dependence relations which can be used by

a speaker to interpret a sentence, resulting in different scope sequences according to traditional theories. It is important to note that such a strategy is only possible if quantifiers are interpreted as relations between sets of assignments because only then is it possible to impose dependence relations on the output team. In this respect, our approach is similar to that of [BF11]. In contrast to approaches based on (generalized) Skolem functions, there is no separation between the existential scope and the property denoted by the NP so that the problem related to (11) does not arise. The fact that existential and distributional scope are different is explained as follows. A sentence like (8) is analyzed as the join of two relations. For bare numerical indefinites like ‘three’, both a distributive and a collective reading are possible. The thesis, now, is that (8) has to be analyzed in a way similar to the examples in (19). Recall that in the former case an IS reading is possible only if the quantifier triggering this reading shows strong variation. This type of variation is guaranteed if the quantified NP (‘three relatives’) gets a collective reading (only one plural object). By contrast, on a distributive interpretation, strong variation fails if one assumes that bare numerical indefinites are total, similarly to \forall . One way of establishing this thesis consists in arguing that on a distributive reading for a bare numerical indefinite like ‘three relatives’ the function F_{relative} always yields the whole set of three relatives as output, similarly to \forall (see full paper for details).

It goes without saying that a lot of work remains to be done. First, the formal details must be worked out in a rigorous way (see full paper for details). Second, there are at least the following empirical questions which have to be answered: (i) How is the fact explained that ‘exactly’ does not show upward unboundedness?, (ii) How can such principles as the Binder Roof Constraint and the No Skipping Constraint be incorporated into the theory (see [BF11])? and (iii) The definition of the join operator \cap does not take into account semantic differences between, say, relative clauses and conditionals. How can \cap be used in a compositional way? This requires, at least to incorporate a component to handle both intra- and intersentential anaphora into the theory. We are currently working on integrating the current theory with the sequential semantics of Vermeulen and Van Eijck’s Incremental Dynamics.

References

- [BF11] A. Brasoveanu and D. Farkas. How indefinites choose their scope. *Linguistics and Philosophy*, 34:1–55, 2011.
- [Chi01] G. Chierchia. A puzzle about indefinites. In C. Cecchetto and et al., editors, *Semantic Interfaces*, pages 51–90. CSLI Publications, Stanford, 2001.
- [Eng12] F. Engström. Generalized quantifiers in dependence logic. *Journal of Logic, Language and Information*, 21:299–324, 2012.
- [Fer93] T. Fernando. The donkey strikes back: Extending the dynamic interpretation “constructively”. In *EACL 1993*, pages 130–138, 1993.
- [Gal13] P. Galliani. Transition semantics: the dynamics of dependence logic. *Synthese*, 2013. Published online: 11 August 2013.
- [Sch06] P. Schlenker. Scopal independence: a note on branching and wide scope readings of indefinites and disjunctions. *Journal of Semantics*, 23:281–314, 2006.
- [Sza97] A. Szabolcsi. Strategies for scope taking. In A. Szabolcsi, editor, *Ways of Scope Taking*, pages 109–154. Kluwer, Dordrecht, 1997.
- [Sza10] A. Szabolcsi. *Quantification*. CUP, Cambridge, 2010.
- [Vää07] J. Väänänen. *Dependence Logic*. CUP, Cambridge, 2007.
- [WR11] Y. Winter and E. Ruy. Quantifier scope in formal linguistics. In Gabbay D., editor, *Handbook of Philosophical Logic*. 2nd edition edition, 2011. to appear.

Attaching NRCs to Plural Quantificational Heads

Claudia Poschmann

Goethe-Universität, Frankfurt a.M., Germany
poschmann@lingua.uni-frankfurt.de

Abstract

This paper presents the results of two experiments in German investigating the acceptability and interpretation of non-restrictive relative clauses (NRCs) with (plural) quantificational heads. Contrary to standard assumptions, NRCs with strong quantificational head are grammatical in sentence-internal subject position and even can get a reading in this position according to which the relative pronoun is referring to the intersection of restrictor and scope of the quantification. This observation does not only give interesting insights into the nature of NRCs, but might have far reaching consequences for existing dynamic approaches to plural anaphora in discourse.¹²

1 Introduction

As is well known, non-restrictive relative clauses (NRCs) pattern quite consistently with standard discourse anaphora. Just as anaphoric pronouns, NRCs are felicitous with referential (1), but heavily restricted with singular quantificational antecedents (2). Plural quantificational heads, by contrast, allow for anaphoric reference as well as for attachment of a NRC (3).

- (1) a. Paul invited **Nick**, who lives next door.
- b. Paul invited **Nick**. He lives next door.
- (2) a. Paul invited **every boy**, *who lives next door.
- b. Paul invited **every boy**. *He lives next door.
- (3) a. Paul invited **most boys**, who live next door.
- b. Paul invited **most boys**. They live next door.

Due to this striking similarity, most accounts assume that NRCs are interpreted (at least in a certain way) like independent propositions involving a discourse pronoun (Sells 1985 [10], Del Gobbo 2003 [3], Schlenker 2010 [9]). However, it has been argued that this parallel between NRCs and their matrix clause paraphrases should not be straightened too far. Del Gobbo (2003) [3] for example observed that (4b), where the NRC is attached to a quantificational antecedent in subject position, sounds strange while (4a), where the NRC is attached to a similar antecedent in object position, is perfectly acceptable. No contrast can be found for the corresponding matrix clause paraphrases in (5).

- (4) a. Paul invited **most students**, who came very late.
- b. ***Most students**, who arrived late, came to the party.
- (5) a. Paul invited **most students**. They came very late.
- b. **Most students** came to the party. They arrived late

¹Thanks to Adrian Brasoveanu, Rick Nouwen, Caroline Reinert and T. E. Zimmermann for fruitful discussions.

²This work was supported by a grant from the German Organisation for Scientific Research DFG.

Based on contrasts such as (4), Del Gobbo (2003) [3] concludes that position affects the grammaticality of NRCs such that NRCs with plural quantificational head are ungrammatical if attached in subject (sentence-internal) and not in object (sentence-final) position.³ Nouwen (2007) [7], by contrast, argues that the position of a NRC does not affect its grammaticality but its interpretation. According to Nouwen (2007), both sentences in (4) are grammatical, but differ in interpretation. In (4)a, the NRC can refer to the intersection of restrictor and scope of the quantification (intersection-set), giving rise to an interpretation according to which all of the invited students arrived late. In (4)b, by contrast, such an intersection-set reading is unavailable. The NRC can only refer to the maximal set of all students in the discourse (restrictor-set), which leads to a contradiction. Whereas the matrix clause implicates that most but not all students came to the party, the NRC tells us that all students (in the discourse) arrived very late at the party. A simple explanation for this contrast could be that at the time of evaluation of the NRC in (4)b, the intersection of restrictor and scope of the quantification is not yet specified and hence not yet available for anaphoric reference. In fact, this is quite what one would expect following standard theories of plural discourse anaphora. In DRT, for example, strong quantifiers introduce so-called "duplex-conditions". Discourse referents introduced within these duplex conditions are inaccessible from outside of the quantifier's scope, which explains the ungrammaticality of (2). Only after the duplex condition is evaluated, a separate antecedent formation process, called "abstraction" (Kamp & Reyle 1993 [4]), can be applied. This abstraction process allows to introduce a new discourse referent, which represents the set of all individuals satisfying restriction and scope of the preceding quantification (reference-set), making it available for plural reference. The crucial point, however, is that "abstraction" can only be applied after the quantification itself has been processed. Interestingly enough, the latter assumption has been motivated by closely related examples with restrictive relative clauses (RRCs), such as (6). According to Kamp & Reyle (1993) [4], sentence (6) only has a reading according to which both matrix and relative clause are interpreted distributively. In particular, (6) has no reading in which the plural pronoun "they" refers to the intersection-set of the quantification.⁴

- (6) Most lawyers hired a secretary they liked.

Note that all this holds for strong (non-intersective) quantifiers only. In DRT, weak (intersective) quantifiers do not introduce duplex-conditions but set-denoting (plural) discourse referents, which are available even before the quantifier's scope is specified. Thus, from a DRT perspective, it would be expected that the contrast in examples such as (4) shows up with strong (non-intersective) quantifiers only.

2 Experiments

We set-up an online-questionnaire in German (survey monkey), which 106 native speakers completed. The questionnaire consisted of two parts, an acceptability test and an interpretation test.

2.1 Part 1: Acceptability Test

The first experiment was designed to test the general acceptability of plural NRCs with quantificational heads. It contained 9 items which were presented each in one and the same condition. These included 3

³Del Gobbo (2003) [3] provides a syntactic explanation for this contrast: For the NRC to be interpreted as a independent sentence, it has to rise to a text node higher than the matrix sentence. This can be achieved by a restructuring process, but only if the NRC is attached in (sentence-final) object position. Note that this explanation is based on the assumption that the NRC is attached at DP level. If we would attach the NP higher in the tree, at IP or CP level, we would not expect such a contrast.

⁴What Kamp & Reyle (1993) [4] do not mention in this context is that in case of examples such as (6) the intersection set reading is unavailable for independent reasons. RRCs, in contrast to NRCs, contribute to the scope of a quantification. That's why any attempt to derive an intersection set reading of a pronoun inside a RRC ends up in circularity. We will come back to this problem at the end of the paper.

plural NRCs with quantified heads (test-items), all with strong quantifiers, 2 of them in sentence-internal position (subject-attachment) and 1 in sentence-final position (object-attachment). These test-items were contrasted with two types of fillers: 3 NRCs with singular quantificational heads, which were expected to be judged as clearly unacceptable (false fillers), and 3 RRC with singular quantificational heads, which were expected to be judged as fully acceptable (correct fillers). (7) provides an example for each condition.⁵

- (7) a. Die meisten Mütter, die ja nur das Beste für ihr Kind wollen, kaufen Bio-Produkte. (TI)
(*Most mothers, who PART only want the best for their child, buy organic products.*)
- b. Kein Kind, das übrigens auf der Rutsche sitzt, trägt eine blaue Jacke. (FF)
(*No child, which is PART sitting on the slide, is wearing a blue jacket.*)
- c. Jeder Student, der am Tutorium teilgenommen hat, hat die Prüfung bestanden. (CF)
(*Every student which participated at the tutorial passed the exam.*)

The participants were asked to judge the acceptability of the items on a scale from 0 (completely unacceptable) to 5 (fully acceptable). Note that, in this part, the test-items were designed to be neutral with respect to a possible restrictor-set or intersection-set reading.

PREDICTIONS: According to Del Gobbo (2003) [3], the test-items are expected to pattern with wrong fillers, Nouwen (2007)[7] predicts the test-items to get a rating comparable to correct fillers.

RESULTS: Using the lme4 package in R, we fitted a model of mixed logistic regression with CONDITION (Test-Item, Correct Filler, False Filler) as fixed effect and random effects for subjects and items including the corresponding slopes, first for the whole data set and then for the subset with sentence-internal position of the RC. In the full data set, the test-items rated nearly as high (4.02 on a scale from 0 to 5) as the corresponding correct fillers (4.26) and significantly higher than wrong fillers (0.72) ($t=-9.946$ for the comparison between test-items and false fillers).⁶ At least in our model, there was no significant difference between test-items and correct fillers ($t=0.432$). Only in the subset of items with sentence-internal position of the RC, we could find a difference between test-items and correct fillers. As expected, the test-items rated significantly lower than the correct fillers ($t=4.37$), but still very high, especially in comparison with the false fillers in the data set ($t=-12.11$), namely at 3.87 compared to 0.70 for the wrong fillers and 4.27 for the correct fillers. Only 3 out of 106 participants rejected all the test-items they were confronted with (ratings below 3).

2.2 Part 2: Interpretation Test

The second experiment was designed to investigate whether the availability of intersection-set readings depends on the NRC's POSITION in the matrix clause (internal versus final) and/or on the STRENGTH of the head's quantifier (strong versus weak). We presented the NRCs in contexts in which the intersection-set reading was explicitly ruled out. The participants were asked to judge whether the whole context seemed plausible or clearly contradictory to them. If the participants didn't get the intersection-set reading they were expected to judge the context as contradictory. The following is an example for the condition with strong quantifier and sentence internal position of the NRC.

⁵In German, unlike in English, both, NRCs and RRCs, are separated by comma. To mark the test-items as NRCs, we introduced discourse-particles (e.g. *übrigens/by the way*), which should rule out a restrictive interpretation.

⁶According to Baayen et al. (2008), we can be confident that the comparison is significant if the absolute value of the t-value is bigger than 2 (or: 1.96).

- (8) Das Lego-Set City umfasst über 300 Steine **verschiedener Farben und Größen**, unter anderem für eine Poststelle und eine Polizeistation. Die meisten Steine, **die natürlich alle rot sind**, gehören zu einem Feuerwehrhaus. (*The Lego Set City includes more than 400 bricks of different colour and size, for example for a post office and a police station. Most bricks, which are of course all red, belong to a big fire station*).

In this part, we tested 12 items in 4 conditions, varying the NRC's POSITION in the matrix clause (internal/final) and the quantifier's STRENGTH (strong/weak). In total we tested six different quantifiers, 3 STRONG quantifiers (*die meisten/most, die wenigsten/few, ein Drittel/a third of*) and 3 WEAK quantifiers (*mehr als n/more than n, weniger als n/less than n, genau n/exactly n*). Conditions and quantifiers were equally distributed over a latin square design.⁷ Additionally, the test-items were interspersed with fillers (18 similar constructed texts, 8 without an intended contradiction (Correct Fillers), 8 with an intended contradiction outside of a relative clause (False Fillers)).

PREDICTIONS: According to Nouwen (2007) [7] and Kamp & Reyle (1993) [4], it is expected that both the quantifier's strength and the position of the NRC in the matrix clause have an effect on the availability of an intersection-set reading. First, the intersection-set reading of NRCs with plural quantificational heads is expected to be better available with weak than with strong quantifiers. Notably, it is predicted that with strong quantifiers in internal position the intersection-set reading is not available at all. With strong quantifiers in final position the intersection-set reading is predicted to be available but still expected to rate lower than with weak quantifiers. In the case of strong quantifiers, the restrictor-set is presupposed to be contextually given and hence available for anaphoric reference. Weak quantifiers, by contrast, do not carry such a presupposition and as a consequence the restrictor-set is not available for anaphoric reference. Thus, whereas in case of strong quantifiers there are two competing readings available for the NRC in final position, weak quantifiers allow for the intersection-set reading only, independently of the NRC's position in the matrix clause.

RESULTS: Using the lme4 package in R, we fitted a model of mixed logistic regression for the interaction of STRENGTH and POSITION as fixed effects and random effects for subjects and items including the corresponding slopes. As expected, we found a highly significant effect of the quantifiers STRENGTH (weak versus strong) ($p < 0.001$), but we didn't find any significant effect of POSITION (final versus internal) ($p=0.979$), or of the interaction of STRENGTH and POSITION ($p=0.887$). Moreover, although conditions with strong quantifiers rated significantly lower than those with weak quantifiers, their overall acceptability was surprisingly high (the relative frequency of accepted intersection-readings was 0.9 for strong quantifiers versus 0.98 for weak quantifiers). For comparison, the correct fillers rated at 0.97 and the false fillers at 0.03. However, we found differences within the subset of items with strong quantifiers. Firstly, the quantifier "ein Drittel" (a third of) rated nearly as high as the average of the weak quantifiers (namely at 0.96 and significantly different compared to "die meisten" (most) with 0.87 and "die wenigsten" (few) with 0.86) ($z=-2.964, p=0.00304$ for the comparison between "ein Drittel" (third of) and "die meisten" (most)). Moreover, "die meisten" (most) and "die wenigsten" (few) seem to be more sensitive to POSITION than the rest of the quantifiers we tested. Whereas with "die meisten", final position of the NRC rated higher than internal position (0.91 versus 0.83), with "die wenigsten" final position rated lower than internal position (0.82 versus 0.89). This looks like an effect of monotonicity. But at least in our model, it didn't turn out to be significant ("die meisten": $z=-1.29, p= 0.196$ / "die wenigsten" $z= 1.862, p =0.062$).

⁷Note that in German, the word order can be changed quite freely and subjects can appear as well in internal as in final position. To keep the parameters involved as minimal as possible, we tested only NRCs with subject-attachment.

2.3 Discussion

Besides of some minor relativizations considering monotonicity, the main results of the two experiments are rather clear. (i) Against Del Gobbo's (2003) [3] predictions, NRCs attached to plural quantificational heads are grammatical even at sentence-internal (subject) position in German. (ii) And against Nouwen's (2007) predictions [7], NRCs in German can get an intersection-set reading even if attached to strong (non-intersective) quantifiers in sentence-internal (subject) position. (iii) As expected, the quantifiers strength does influence the availability of intersection-set readings, with NRCs attached to strong quantifiers rating lower than those attached to weak quantifiers, both in internal and in final position of the NRC. (vi) Quite unexpected is that, besides of the reported monotonicity effect, we didn't find an effect of position on the availability of the intersection-set reading with strong quantifiers. This does not only run counter the expectations of Kamp & Reyle (1993) [4], but also counter the intuitions reported by Del Gobbo (2003) [3] and Nouwen (2007) [7]. There are at least two options to explain this lack of position effect in our experiment. First, one could argue that these unexpected readings are the result of a discourse-level repair strategy. This is what Nouwen (2007) [7] suggests for similar effects in nominal appositions. From our point of view, however, the acceptability rates are too high for such an explanation, especially since the participants were hunting for contradictions and the comments the participants were invited to give showed that most of them didn't even realize the existence of a competing restrictor-set reading. The other option is to assume that, in case of strong quantificational head and internal position of the NRC, the intersection-set reading is not missing but only less salient than the restrictor-set reading, since the former in contrast to the latter is not yet specified. This assumption would be compatible with both, the intuitions behind Del Gobbo's (2003) [3] and Nouwen's (2007) [7] examples and the results of our experiment.⁸ If we follow the second option, however, we need a more fine-grained approach of plural anaphora. In the remainder of this paper, we will make a proposal how the plural anaphora account of Brasoveanu (2010) [2] can be modified to make it compatible with both the intuitions behind Del Gobbo's (2003) [3] original example and the results of our experiments.

3 Alternative Approaches to Plural Anaphora

Brasoveanu (2010) [2] offers a very fine-grained dynamic view of discourse anaphora based on plural information states in the sense of van den Berg (1996)[1]. In contrast to standard information states (which represent single assignments i,j etc.) these plural information states are modeled as sets of variable assignments I, J, which are able to store not only the values of the variables/discourse referents introduced (quantificational domains) but also the (quantificational) dependencies established between them. More importantly (for the interest of this paper), Brasoveanu (2010) [2] assumes that quantifiers introduce two separate referents, one for the restrictor-set and one for the intersection-set, with the intersection-set being defined as proper subset of the restrictor-set. What will matter as well for our argumentation is the fact that the proposal of Brasoveanu (2010) [2] is couched into a C(ompositional)DRT framework in the sense of Muskens (1996) [5].⁹ By consequence, in Brasoveanu (2010) [2] quantification is defined

⁸Note that our experiment only tested the availability of the intersection-set reading in contexts in which the competing restrictor-set reading was explicitly ruled out. We didn't test for salience of a particular reading.

⁹Dynamic info states I, J, etc are plural: they are sets of variable assignments, i.e., terms of type st. An individual dref x stores a set of individuals with respect to a plural info state I, abbreviated as uI def = $u_s e i_s : i_s \in I_{st}$, i.e., uI is the image of the set of assignments I under the function u . A sentence is interpreted as a Discourse Representation Structure (DRS), which is a relation of type (st)(st) between an input state I_{st} and an output state J_{st} , as shown in (a) below. A DRS requires, the input info state I to differ from the output state J at most with respect to the new drefs and all the conditions to be satisfied relative to the output state J. The definition of dref introduction (a.k.a. random assignment) is given in (b) below.

(a) DRS: $[newdrefs|conditions] = \lambda I_{st}. \lambda J_{st}. I[newdrefs]J \wedge I[conditions]J$
(b) New Dref: $[u] = \lambda I_{st}. \lambda J_{st}. \forall i_s \in I (\exists j_s \in J (i_s \in J)) \wedge \forall j_s \in J (\exists i_s \in I (i_s \in J))$

as quantification over DRT-like (plural) discourse referents. In (9), you see the definition of a quantifier as it is defined in Brasoveanu (2010) [2].

$$(9) \quad \text{DET (Brasoveanu 2010)} \\ \lambda P. \lambda P'. \max^u(\text{dist}_u(P(u))); \max^{u' \subseteq u}(\text{dist}_{u'}(P'(u'))); [\text{DET}u, u']$$

According to this, a quantificational determiner DET is of type $(et(et(t)))$, as usual, but it does not compare two sets of assignments, but two sets of individuals, the maximal set of all individuals that distributively satisfy the restrictor property P (restrictor-set) and the maximal set of all individuals that distributively satisfy the nuclear scope property P' (intersection-set). For each of these two sets the quantifier introduces a (plural) discourse referent, u and u' , with the intersection-set referent u' being defined as a proper (structured) subset of the restrictor-set-referent u .¹⁰ The quantification itself is defined as a static test over these two referents. The maximization and distributivity operators \max_u and dist_u are defined as selective maximization and selective distributivity operators in the sense of van den Berg (1996) [1], which allow to dynamize λ -abstraction over both values (i.e., quantifier domains) and structure (i.e., quantificational dependencies). Very roughly, the \max -operator introduces a new discourse-referent u and assigns to it the maximal set of all individuals which satisfy the condition D, whereas the dist -operator ensures that we previously split up the plural information state in its single information states (i,j), guaranteeing that we sum up with the \max -operator the maximal set of individuals which distributively satisfy D. Note that in contrast to the \max -operator, the dist -operator is not superscripted but subscripted with u . This represents the fact that the \max -operator introduces a new discourse referent for u , while the dist -operator is only making anaphoric reference to it.

$$(10) \quad \max^u(D) = \lambda I_{st}. \lambda J_{st}. ([u]; D)IJ \wedge \forall K_{st}(([u]; D)IK \rightarrow uK_{u \neq \#} \subseteq uJ_{u \neq \#})$$

$$(11) \quad \text{dist}_u(D) = \lambda I_{st}. \lambda J_{st}. uI = uJ \wedge \forall x_e \in uI (DI_{u \neq x} J_{u \neq x})$$

Note that in the actual version of Brasoveanu (2010) [2], the discourse referents are introduced in the scope of the respective \max -operator. As a consequence, the discourse referent for the intersection-set is only available after the nuclear scope set is specified by P' . Thus, although this approach is much more fine grained than the standard DRT-account of Kamp & Reyle (1993) [4], the account of Brasoveanu (2010) [2], as it stands, makes the same predictions with respect to our data.¹¹

3.1 Modifying Brasoveanu (2010)

In the following, we will argue, that the account of Brasovenau (2010) [2] can be slightly modified such that the intersection-set can be made available for reference in sentence-internal position. The main idea is rather simple: Since in Brasoveanu (2010) [2] u and u' are discourse referents, we can easily pull the discourse referents out of the scope of the \max -operators and introduce them even before they are specified. To achieve this, we only have to define the \max -operator as anaphoric, as is represented by the fact that the \max -operator is now subscripted (and not superscripted) by u and u' . The equivalent modified definition of the \max -operator is given in (13) below.

¹⁰To guarantee that subsets not only inherit the value but also the structure of their supersets, Brasoveanu (2010) [2] defines $u' \subseteq u$ as "structured inclusion" (i), with $\#$ representing dummy-individuals that are used as a tag for the cells in the matrix that should be discarded in order to obtain a structured subset u' of a superset u .

(i) $u' \subseteq u = \lambda I_{st}. (u' \subseteq u) I \wedge \forall i_s \in I_{st} (ui \in u'I_{u' \neq \#} \rightarrow ui = u'i)$

¹¹Note that similar holds for the closely related account of Nouwen (2003). Since in Nouwen (2003) quantification is defined over sets of assignment and the intersection-set-assignment is defined only if the nuclear scope property P is defined, intersection-set-reference is available only after the quantification itself has been processed.

- (12) DET (First Modification):
 $\lambda P.\lambda P'.[u];[u'];[DET u,u'];max_u(dist_u(P(u)));max_{u' \subseteq u}(dist_{u'}(P'(u')))$
- (13) (Modified *max*-operator):
 $max_u(D) = \lambda I_{st}.\lambda J_{st}. DIJ \wedge \forall K_{st}(([u]; D)IK \rightarrow uK_{u \neq \#} \subseteq uJ_{u \neq \#})$

By this little trick, we can make the intersection-set referent u' available even before the nuclear scope-property P' is specified. Thus, it would be in principle possible to attach the NRC either in internal or in final position. (14) represents the equivalent translations of the test-item represented in (8) with internal position of the NRC.¹²

(14) **sentence-internal attachment**

Most bricks, which are (of course) all red, belong to a big fire-station.
 $[u] ; [u'];[MOST u,u'];max_u(dist_u(bricks(u)));red(u');$
 $max_{u' \subseteq u}(dist_{u'}(belong_to_firestation(u')))$

Note that this mimics quite closely the surface word order of the sentence. First, the quantificational determiner "most" introduces two separate discourse referents u and u' and tells us that we have to test whether these two discourse-referents stand in the "most"-relation. Secondly, the noun phrase "bricks" specifies our discourse referent u (our restrictor-set referent). Before we get the specification of the corresponding intersection-set referent u' by the VP, the NRC is introduced. Nevertheless, the NRC can make reference to the (still underspecified) u' . (14) captures quite nicely the intuitions behind Nouwen's (2007) [7] and Del Gobbo's (2003) [3] examples: Although the intersection-set referent is available for anaphoric uptake in sentence-internal position, it is not yet specified (unlike the corresponding restrictor-set-referent). Hence, it is not surprising that in internal position of the NRC the restrictor-set-reading is more prominent. To make this idea compositional, however, is a little bit more complicated.¹³ The problem is that the discourse referents u and u' are literally spoken "invisible" in our syntactic tree.¹⁴ To make the intuition behind (12) compositional, it would be necessary to make two assumptions: (i) The NRC is attached at DP-level (which is quite unproblematic) and (ii) the quantificational DP itself has to be split up syntactically. We assume that the quantificational determiner is a four-place determiner of type $(et(et(e(e(t))))$, with two predicative and two referential arguments (λ -abtracts), which function as sort of place-holder for the discourse-referents u and u' , by which they are substituted at VP- (or IP-) level. The corresponding definition of the determiner and the necessary construction rules for NRC-Attachment and Discourse-Referent-Introduction are given in (16) and (17) below.¹⁵

- (15) Modified Definition of the Determiner:
 $\lambda P.\lambda P'.\lambda x.\lambda x'.[DET x,x'];max_x(dist_x(P(x)));max_{x' \subseteq x}(dist_{x'}(P'(x')))$
- (16) NRC Attachment Rule:
If C is a branching node consisting of two sister nodes A and B, A with the translation α being of type $(et(e(e(t))))$ and B with the translation β of type (et) , C has the following translation γ :
 $\lambda P'.\lambda x.\lambda x'.\alpha(P')(x)(x');\beta(v)^{16}$

¹²Since in Nouwen 2003 [6] quantification is defined as quantification over (partial) sets of assignments, it might be the case that this trick cannot be applied to this approach. To investigate this, however, would go beyond the scope of this paper.

¹³Thanks to Caroline Reinert and T.E.Zimmermann for pointing this out to me.

¹⁴If we attach the NRC at DP-level (Del Gobbo 2003 [3]), we only have access to the max-set-interpretation. If we attach the NRC by contrast at IP-level or higher, we lose the possibility to differentiate between internal and final position of the NRC and would predict that max-set-readings and intersection-set readings are equally available.

¹⁵This construction rule is only used to show that attachment of NRCs is generally possible. It is not intended as a full-blown analysis of NRCs. For a detailed discussion of the attachment of NRCs in this framework cf. to Reinert (forthcoming) [8].

¹⁶v can be set equal with either x or x' depending on whether the NRC is referring to the restrictor- or intersection-set of the quantification. Note that after the application of this rule C is still of the same type $(et(e(e(t))))$. This makes the rule recursive, allowing for example for stacking of NRCs.

- (17) Introduction of Discourse-Referents:

If A is a node of type ((et (e (e t)))) with the translation α , A can be replaced by a node B of type ((et t) with the following translation: $\lambda P'.[u];[u'];\alpha(P')(u)(u')$

With these minor modifications, Brasoveanu (2010) [2] enables us to account for the observed general availability of intersection-set readings in sentence-internal position. But are these predictions not too strong with respect to the corresponding examples with RRCs such as (6)b, which originally motivated the assumption that intersection-set readings should be unavailable in sentence-internal position? A look at the corresponding translation in Brasoveanu's framework shows that this is not the case.

- (18) Most lawyers hired a secretary they liked.

$[u];max_u(dist_u(\text{lawyer}(u)));[u'];max_{u' \subseteq u}(dist_{u'}([v]|\text{secretary}(v),\text{like}(u',v),\text{hire}(u',v)));$
 $[MOSNu,u']$

The crucial point is that RRCs (in contrast to NRCs) are attached at NP level and hence contribute to the scope of the quantification. That's why the RRC (and with it the plural pronoun it contains) is introduced inside the scope of the *dist*-operator. Inside the scope of the *dist*-operator, however, the plural information states are split up into single information states. Thus, the respective discourse referents can only refer to singular individuals and not to the maximal sum of individuals satisfying the scope property P^* (intersection-set).

CONCLUSION: Contrary to standard assumptions about NRCs and plural discourse anaphora, German NRCs attached to plural quantificational heads are grammatical in sentence-internal (subject) position and even can get an intersection-set interpretation in this position. This challenges the classical assumption that the intersection-set of a quantification should be available for anaphoric reference only after the quantification itself has been processed. As we argued however, the plural anaphora account of Brasoveanu (2010) [2], in contrast to competing accounts such as Kamp & Reyle (1993) [4], can be modified to be compatible with these observations by slight modifications.

References

- [1] M. van den Berg. *The internal structure of discourse*. ILLC Dissertation, Amsterdam, 1996.
- [2] A. Brasoveanu. Structured anaphora to quantifier domains. *Information and Computation*, 208:450–473, 2010.
- [3] F. Del Gobbo. *Appositives at the interface*. Ph.D. thesis. University of California, Irvine, 2003.
- [4] H. Kamp and U. Reyle. *From discourse to logic*. D. Reidel, Dordrecht, 1993.
- [5] R. Muskens. Combining montague semantics and discourse representation. *Linguistics & Philosophy*, 19:143–186, 1996.
- [6] R. Nouwen. *Plural pronominal anaphora in context*. Number 84. Netherlands Graduate School of Linguistics Dissertations (LOT), Utrecht, 2003.
- [7] R. Nouwen. On appositives and dynamic binding. *Journal of language and computation*, 5(1):87–102, 2007.
- [8] C. Reinert. *Die Anbindung nicht-restriktiver Relativsätze in Brasoveanu PCDRT*. Master Thesis, University of Frankfurt a.M., forthcoming.
- [9] P. Schlenker. Supplements without bidimensionalism. *Proceedings of the Amsterdam Colloquium*, 2009.
- [10] P. Sells. Restrictive and non-restrictive modification. *Stanford University, CSLI Report*, 85(28), 1985.

A Fregean Semantics for Number Words

Susan Rothstein

Bar-Ilan University, Ramat Gan Israel

susan.rothstein@biu.ac.il

Abstract

This paper proposes a Fregean semantics for cardinal numbers , analysing them as properties (Chierchia 1985). A cardinal numeral can occur at the predicative type denoting a set, and at an argument type denoting the individual correlate of a set. Lexical powers like *hundred* and *thousand* denote multiplicands and are of a different type from other cardinals.

1 Cardinals as singular terms

A semantics for number words must account for numerical expressions in a number of different constructions. These include prenominal modifier position, predicate position, argument position and as ‘determiners’ of multiplicative cardinal heads as in the complex numerical *two hundred*. These constructions are illustrated in (1)-(4) respectively:

- (1) a. *Two* cats were in the garden.
b. The *two* girls cooked a wonderful meal.
c. The guests were *two* girls.
- (2) a. My reasons are *two*.
b. The children were *two*.
- (3) a. *Two* plus *two* is *four*.
b. *Two* is the only even prime number
- (4) *Two* hundred people stood in line.

We want to give as simple an account of cardinals as is possible, ideally deriving all these uses from a single meaning. This requires answering the following questions: (i) What is the type of the cardinal in its standard prenominal position as in (1a)? (ii) What is the relation between the cardinal in prenominal position and its use as a singular term in (3)? (iii) What is the relation between *two* and *hundred* in (4). And we would like to extend our to account for the interpretation of numericals in constructions like (5) and (6) as well.

- (5) Two kilos of strawberries went into this jam.
- (6) Hundreds of people stood in line.

We start by addressing the second question, and arguing in favour of treating the numerical expressions in (3) as singular terms. This is not uncontroversial. Hofweber (2005) and Ionin and

Matushansky (2006) both argue that numericals do not occur as singular terms. Hofweber argues that *two* in *two boys* is a determiner at type <<e,t>,t> and Ionin and Matushansky argue that it is a predicate modifier of type <<e,t>,<e,t>>. Both argue that (7a) (= (3a)) is derived from (7b) by N ellipsis. Hofweber suggests that abstract generalisations such as (7c) are derived from (7b) by ‘cognitive coercion’. It is unclear exactly how cognitive coercion works, but it seems to be an extra-grammatical process of generalising from multiple specific instances to a general statement.

- (7) a. Two and two make four.
- b. Two things and two things make four things.
- c. Two plus two makes/#make four.

However, there are a number of pieces of grammatical evidence that apparent bare singulars like *two* are ambiguous between *two (things)*, where the numerical functions as a cardinal modifier and the modified N is possibly null, and *two* as genuine singular term. First, the verb *count* is ambiguous between two meanings: ‘count how many N there are’ and ‘name in sequence a string of natural numbers. These two uses are shown in (8a) and (8b) respectively.

- (8) a. I counted thirteen (*things, people, books*).
- b. I counted (up) to thirteen (**things*).
- c. I counted (*up to) the books.

Count (how many) in (8a) takes a direct object either *two N* or *two* modifying an elliptical N. *Count* in the second sense must be followed by the P *to* and a numeral which must be bare. Further, *count* in the first sense need not have a cardinal complement, while *count (up) to* must take a numerical complement. The contrast between (8a) and (8b) correlates with the fact that ‘counting how many’ necessarily requires counting objects under a certain description, as discussed in Rothstein (2010). This shows up explicitly in French, where *compter* used in the sense of ‘count how many’ when followed by a bare numeral obligatorily appears with the clitic *en*, ‘of them’. (9a) means “I counted thirteen *of them*”. This clitic is impossible when *compter* is used in its second sense, as shown in (9b).

- (9) a. J'*(**en**) ai compté treize.
I **of-them** AUX-PAST counted thirteen.
“I counted thirteen (of them).”
- b. J'(*en) ai compté jusqu'à treize.
I AUX-PAST counted until thirteen.
“I counted up to thirteen”

Second, statements about numbers using bare cardinals cannot be rephrased using the paradigm in (7). Numbers have second order properties which do not hold of sets of objects or pluralities.

- (10) a. *Two* is even / is a prime number.
- b. #*Two things* are even / are a prime number.

Third, two numbers *n* and *m* do not the same as the relation as two entities with cardinalities of *n* and *m*. Numbers stand in the *bigger than/smaller than* relation, while objects with cardinality properties stand in the *more than/fewer than* relation. (11a) is equivalent to (11b), not to (11c).

- (11) a. *Two* is *smaller than* three.
- b. #*Two things* are *smaller than* three things.
- c. *Two things* are *fewer than* three things.

Fourth, singular terms require singular agreement, while elliptical nouns modified by cardinals may take plural agreement.

- (12) a. Four *is/*are* bigger than three.
- b. Four (things) *are* more than three (things).
- c. "I can't take you all in the car: Five *are/is* too many"

Together, these data indicate that cardinal numerals must have interpretations as singular terms, as well as an interpretation prenominally and in predicate position.

2 Cardinals as predicates

So what account should we give of cardinals in non-argument position, as in the examples (1) and (2)? The plausible options are three. They could be cardinal predicates of the same type as adjectives i.e. type $\langle e, t \rangle$ as argued in Landman (2003); they could be determiners in classical Montague style, at type $\langle\langle e, t \rangle, t \rangle$ as argued in Barwise and Cooper (1981) and more recently Hofweber (2005); they could also be cardinal modifiers of type $\langle\langle e, t \rangle, \langle e, t \rangle \rangle$ as argued in Ionin and Matushansky (2006). A crucial factor deciding is that, as we have seen, numerals must have an interpretation at the argument type. On the assumption that there is one basic meaning for numerals, we are looking for an account in which the shift between argument type and the type licensing other uses is simple. The most straightforward account is that prenominal cardinals are predicates, since as we will see in the next section, there is an obvious way to shift between predicate and argument type. Landman (2003) gives a convincing account of cardinals as adjectives which denote sets of plural individuals with a specific cardinality: $\{x : |x| = n\}$. This explains the fact that they can be bare predicates as in (2). Like any adjective, they shift to the modifier type in prenominal attributive position. In (1c), the sentential predicate is the NP [two_{AP} girls_N] translating as the predicate $\lambda x. \text{GIRLS}(x) \wedge |x| = 2$. Landman argues that in *the two girls*, in (1b), *two* is also a prenominal adjective. *Two girls* denotes the same property as in (1c) and is the argument of the determiner *the*. When there is no determiner, as in (1a), the cardinal adjective raises to determiner position triggering existential quantification. The DP *two girls* in (1a) denotes $\lambda P. \exists x [\text{GIRLS}(x) \wedge P(x) \wedge |x| = 2]$. Landman shows that this account correctly predicts that the predicate *two girls* will have an 'exactly 2' interpretation in (1b,1c) and an 'at least 2' reading as a DP in (1a), when it denotes a generalized quantifier.

- (13) a. #The guests were two girls, and maybe even three
- b. #The two girls, and maybe even three arrived late.
- c. Two girls, and maybe even three, got 100 on the exam.

Accounts which suggest that numerals are born at the type determiners have a much harder time explaining how DP lowers to predicate position in (1c) and (2), especially since, as is well known generalized quantifiers such as *every girl* do not lower in this way. It is also unclear how to account for bare numerical predicates like (2) if *two* is a determiner with an elliptical NP. Hofweber (2005) discusses (14), suggesting that *seven* is a focussed determiner, but he does not give any account of how this would work.

- (14) The number of planets is seven.

An extensively developed account of cardinals as predicate modifiers is presented in Ionin and Matushansky (2006). They start from two desiderata. The first is that *four* should have a single interpretation in (15a) and (15b).

- (15) a. *four* cats b. *four hundred* cats

The second is that all cardinal numerical expressions are of the same type. In particular, *four* and *hundred* in (15b) should both be of the same type. They propose that all cardinal numerals are predicate modifiers at type $\langle\langle e, t \rangle, \langle e, t \rangle\rangle$. In (15a) *four* modifies *cats*. In (15b) *hundred* modifies *cats* and *four* modifies *hundred cats*. (15b) has the recursive structure in (16):

- (16) [four [hundred [cats]]]

A numerical n denotes a partition on a plural individual which has n parts:

- (17) $\lambda P \lambda x_{pl} \exists S [\text{PARTITION}(S, x) \wedge |S| = n \wedge \forall s \in S: P(s)]$ (Ionin & Matushansky 2006: (5))

Thus, (15a) denotes a partition on the set of cats into four cells each containing a single cat (see (18b)). (15b) denotes a partition on a set of pluralities of cats into 4, with each individual cell itself portioned into 100 cells each containing an individual cat. Thus the **multiplicative** effect is obtained (18c).

- | | |
|-------------------------------|--|
| (18) a. <i>four</i> : | $\lambda P \lambda x_{pl} \exists S [\text{PARTITION}(S, x) \wedge S = 4 \wedge \forall s \in S: P(s)]$ |
| b. <i>four cats</i> : | $\lambda x_{pl} \exists S [\text{PARTITION}(S, x) \wedge S = 4 \wedge \forall s \in S: \text{CAT}(s)]$ |
| c. <i>four hundred cats</i> : | $\lambda x_{pl} \exists S [\text{PARTITION}(S, x) \wedge S = 4 \wedge \forall s \in S: \exists S' \text{PARTITION}(S', s) \wedge S' = 100 \wedge \forall s' \in S': \text{CAT}(s')]$ |

Four hundred and four cats is analysed as an NP conjunction as in (19), with both occurrences of *four* analysed as predicate modifiers.

- (19) [four [hundred [cats]]]_{NP} and [four [cats]]_{NP}

This account leaves a number of facts unexplained. First Ionin and Matushansky predict that all possible combinations of complex cardinals should appear multiplicatively. In fact these complex cardinals are very constrained. Only a few lexically specified numerals can occur as multiplicands in structures like (15b), as (20) shows. These include the lexical powers, *hundred* (10^2), *thousand* (10^3), *myriad* (10^4) | (10^6), and words like *dozen*, *score*, as shown in (20a). Other combinations such as **three four cats* or **three twenty cats* are impossible, while *twenty three cats* has only an additive reading. The minimal contrast between *twenty* and *score* rules out an extralinguistic explanation.

- (20) a. three dozen/score/hundred/thousand/million cats
 b. three score cats/ *three twenty cats

Second, the structure in (16) cannot explain why with lexical powers the multiplier is obligatory, while with other numbers it is impossible, as shown in (21).

- (21) a. *hundred cats
 b. one hundred cats
 c. *one twenty cats

Third, all the lexical powers that occur in multiplicative structures like (16) can occur as approximative classifiers. No others can, except for *ten*, which arguably appears in multiplicative structures as the bound morpheme *-ty*. This indicates that multiplicands and cardinals must have different structures which both constructions exploit.

- (22) a. hundreds of cats
- b. thousands of cats
- c. scores of cats
- d. *twenties of cats.

The conclusion is that there is no good reason to treat *four* and *hundred* as being the same type, and that the semantic evidence does not support (16) or the predicate modifier theory of cardinals.

3 Numbers as properties

We aim, then to give an semantic interpretation of numericals which both account for the dual use of numbers as predicates and as bare singular arguments and which allows us to give a semantics for *four* and *hundred* which explains the data in the previous section.

We adopt (and adapt) Chierchia's (1985) property theory, which follows Frege (1892) in assuming that predicates have two modes of interpretation, an unsaturated mode in which they are predicated of arguments, and a saturated mode in which they can be arguments of predicates.

Property theory associates with predicates an applicative interpretation at type $\langle e, t \rangle$, and a corresponding property-correlate at the type of individuals, π . Two operations \cap and \cup switch between them. For the applicative interpretation $\lambda x.WISE(x)$, the individual property correlate, WISDOM is $\cap x.WISE(x)$, as in (23a). (A morphological expression of nominalisation is not necessary. *Blue* has an applicative use in *Her eyes are blue* and a saturated use in *Blue is my favourite colour*.)

- (23) a. $\llbracket wise_{\langle e, t \rangle} \rrbracket = \lambda x.WISE(x)$
- b. $\llbracket wise_{\pi} \rrbracket = \cap \lambda x.WISE(x)$
- c. $\cup \cap \lambda x.WISE(x) = \lambda x.WISE(x)$

For cardinal numerals, we start out with a standard modifier interpretation at type $\langle e, t \rangle$, as in (24a), with the cardinality function defined as in (24b) and x ranging over plural individuals.

- (24) a. $\llbracket four_{\langle e, t \rangle} \rrbracket = \lambda x. |x| = 4$
- b. $|x| = n \leftrightarrow |\{y: y \sqsubseteq_{ATOMIC} x\}| = n$

- (25) a. $\llbracket four \text{ guests} \rrbracket = \lambda x. GUESTS(x) \wedge |x| = 4$
- b. $\llbracket \text{the guests are four} \rrbracket = \lambda x. |x| = 4 \ (\sigma\{x: GUESTS(x)\})$
 $= |\sigma\{x: GUESTS(x)\}| = 4$

The singular term *four* is type n . It denotes the individual property correlate of the set in (24a)

$$(26) \quad \cap(\lambda x. |x| = 4).$$

The central equation which defines numbers is (27):

$$(27) \quad n = \cap \lambda x. |x| = n \quad (\text{for numbers in the domain of type } n)$$

Second order properties at type $\langle \mathbf{n}, \mathbf{t} \rangle$ such as *is a prime number* apply only to numerals at type \mathbf{n} .

Lexical powers, unlike the simple numerals illustrated in (24) are of type $\langle \mathbf{n}, \langle \mathbf{e}, \mathbf{t} \rangle \rangle$ and combine with a numerical expression at type \mathbf{n} , denoting a number to yield a cardinal predicate:

$$(28) \begin{array}{ll} \text{a. } [\![\text{hundred}_{\langle \mathbf{n}, \langle \mathbf{e}, \mathbf{t} \rangle}]\!] & = \lambda n \lambda x. |x| = 100 \times n \\ \text{b. } [\![\text{two hundred}_{\langle \mathbf{e}, \mathbf{t} \rangle}]\!] & = \lambda x. |x| = 100 \times 2 \\ \text{c. } [\![\text{two hundred}_n]\!] & = 200 = {}^\cup(\lambda x. |x| = 200). \end{array}$$

$$(29) \quad [\![\text{two hundred cats}]\!] = \lambda x. \text{CATS}(x) \wedge |x| = 100 \times 2$$

Lexical powers are similar to measure expressions such as *kilo* which combine with numerals on their singular term interpretation at type \mathbf{n} to give a measure predicate $\lambda n \lambda x. \text{MEAS}_{\text{weight}=\langle n, \text{KILO} \rangle}$, in examples such as (5). (Landman 2004, Rothstein 2009). Lexical powers and measure expressions, are thus construction in which bare cardinals are used at type \mathbf{n} .

Once this analysis is given, various ways of deriving complex numerals are available. One way is as follows. We derive *two hundred and four* at type \mathbf{n} from (30a) together with the expressions in (24a) and (28c) shifted to type \mathbf{n} via ${}^\cup$.

$$(30) \begin{array}{ll} \text{a. } [\![\text{and}_{\langle \mathbf{n}, \langle \mathbf{n}, \mathbf{n} \rangle}]\!] & = \lambda m \lambda n. n + m \\ \text{b. } [\![\text{two hundred and four}_n]\!] & = \lambda m \lambda n. m + n (200) \\ & = \lambda n. 200 + n (4) \\ & = 204 \\ \text{c. } [\![\text{two hundred and four}_{\langle \mathbf{e}, \mathbf{t} \rangle}]\!] & = {}^\cup 204 = \lambda x. |x| = 204 \end{array}$$

$'+'$ is defined in terms of the sum relation. It maps two numbers n and m onto the cardinality of the sum of two non-overlapping entities y and z , where y has cardinality m and z has cardinality n . This is given in (31). \oplus and \sqcup are the standard overlap and sum functions respectively. *Two plus three* denotes the number associated with the cardinality of the entity x , where x is the result of summing y , an entity with two atomic parts, with z , an entity with three atomic parts, if y and z do not overlap.

$$(31) \quad +_{\langle \mathbf{n}, \langle \mathbf{n}, \mathbf{n} \rangle \rangle} = \lambda m \lambda n. {}^\cup \lambda x. \forall y \in {}^m \mathbf{n}: \forall z \in {}^n \mathbf{n}: \neg(y \oplus z) \rightarrow |\sqcup \{y, z\}| = |x|$$

$'\times'$ can be given a similar interpretation, using the operation *DISJOINT*, which maps a set X onto $X' \subseteq X$, such that no two members of X' overlap. \times combines with n and m and yields the number of atomic parts of the sum of n non-overlapping entities each with cardinality m .

$$(32) \quad \times_{\langle \mathbf{n}, \langle \mathbf{n}, \mathbf{n} \rangle \rangle} = \lambda m \lambda n. {}^\cup \lambda x. \forall z \in \text{PLURAL}(\text{DISJOINT}({}^\cup m)) \wedge |z| = n: \\ |\{x': \forall y \sqsubseteq z: x' \sqsubseteq_{\text{ATOM}} y\}| = |x|$$

Two times three denotes the number of atomic parts of two non-overlapping entities each with three atomic parts, i.e. the number 6.

This analysis is Fregean in two senses. First, cardinals denote properties at both the applicative and argument types. As applicative, unsaturated expressions they denote cardinality properties. As saturated, argument expressions they denote numbers. This is a direct instantiation of Frege's insight (Frege 1892) that a property has 'two modes of presentation', one unsaturated in which it applied to an argument to form a sentence, and one saturated, in which it can itself be the subject of a predication. Second Frege argued that as an object, a number n denotes the equivalence class of sets with cardinality n (Frege 1884). This is captured directly. Frege treated *four* as denoting the equivalence class in (33a), i.e. the set of sets with cardinality 4, sets with four members. We have defined the

meaning of the cardinality predicate as a property of plural individuals (24a). Since a plural individual with cardinality n is the sum of a set with n atomic parts (24b), the one can be reduced to the other, for example by the **PRED** function, defined in (33b), which maps sets of sets with cardinality n onto sets of plural individuals with cardinality n , the meaning of the cardinal predicate.

- (33) a. $\{X: |X| = 4\}$
- b. For $Y = \{X: |X| = n\}$: $\text{PRED}(Y) = \{x: \exists X \in Y: x = \sqcup X\}$
- c. $\text{PRED}(\{X: |X| = 4\}) = \{x: \exists X \in Y: x = \sqcup X\} = \{x: |x| = 4\}$

4 Lexical powers as classifiers

We conclude by looking briefly at lexical powers in their classifier use. All lexical powers, i.e. numbers at type $\langle n, e, t \rangle$ occur as approximatives. As already noted, they must be marked plural (34a), they may not be preceded by a numerical (34b), and, they have the syntax of classifiers: they must be followed by *of* and a bare plural (34c):

- (34) a. hundreds of cats/hundred of cats
- b. *two hundreds of cats
- c. hundreds of cats/*hundreds cats

Only lexical powers occur in this construction. (As noted above, *ten* counts as a lexical power, although as a multiplicand it only occurs in its bound form *-ty*, as in *twen-ty*, *thir-ty* etc.)

Like measure classifiers, approximative classifiers can occur with definite determiners preferably if the nominal is modified by a relative clause:

- (35) a. the hundreds of cats (#that I saw in the garden)
- b. the ten kilos of flour (#that should have been delivered this morning)
- c. The hundreds of rabbits (that you promised me I should see) never appeared.

Since only lexical powers can be used as classifiers in this way (see the examples in (22)), we assume that the operation APPROX deriving approximative readings must exploit the $\langle n, d, t \rangle$ type of lexical powers. The morphological operation associated with the operation is pluralization: *hundred* become *hundreds* and so on. We assume the semantic operation is APPROX as in (36), which derives *hundreds* from *hundred*:

$$(36) \quad \text{APPROX}([\![\text{hundred}_{\langle n, e, t \rangle}]\!]) = \text{APPROX}(\lambda n \lambda x. |x| = n \times 100) \\ = \lambda x. \exists n [n \geq 2 \wedge |x| \geq n \times 1000]$$

The APPROX operation does three things: it existentially quantifies over the n argument, it changes the '=' to ' \geq ' and it adds the clause ' $n \geq 2$ '. APPROX($[\![\text{hundred}_{\langle n, e, t \rangle}]\!]$) gives the set of individuals which have a cardinality which is greater than $100 \times n$, where n is greater than or equal to 2, i.e. which have a cardinality greater than 200. Since classifiers are of type $\langle e, t \rangle, \langle e, t \rangle$, (36) shifts to the predicate modifier type to apply to the denotation of *cats*.

$$(37) \quad \lambda P \lambda x. \exists n [n \geq 2 \wedge |x| \geq n \times 100 \wedge P(x)] (\lambda x. \text{CATS}(x)) \\ = \lambda x. \exists n [n \geq 2 \wedge |x| \geq n \times 100 \wedge \text{CATS}(x)]$$

Approximators can stack as in *hundreds of thousands of cats*. Function composition composes *hundred* and *thousand*, as in (38a), and APPROX applies to the whole expression, marking every lexical power plural.

- (38) a. *hundred* \circ *thousand* = $\lambda n \lambda x. |x| = n \times 100 \circ \lambda n \lambda x. |x| = n \times 1000$
 $= \lambda n \lambda x. |x| = (\wedge(\lambda y. |y| = n) \times 100) \times 1000$
- b. APPROX($\lambda n \lambda x. |x| = (\wedge(\lambda y. |y| = n) \times 100) \times 1000$) =
 $\lambda x. \exists n [n \geq 2 \wedge |x| = (\wedge(\lambda y. |y| = n) \times 100) \times 1000]$
- c. *hundreds of thousands of cats*:
 $\lambda x. \exists n [n \geq 2 \wedge |x| = (\wedge(\lambda y. |y| = n) \times 100) \times 1000 \wedge \text{CATS}(x)]$

References

- Barwise, J., & Cooper, R. (1981) Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4(2), 159-219.
- Chierchia, G. (1985) Formal semantics and the grammar of predication. *Linguistic Inquiry* 16.3 (1985): 417-443.
- Frege, G. (1884) *Die Grundlagen der Arithmetik: eine logisch-mathematische Untersuchung über den Begriff der Zahl*, Breslau: W. Koebner. Translated as *The Foundations of Arithmetic: A logico-mathematical enquiry into the concept of number*, by J.L. Austin, Blackwell: Oxford, second revised edition, 1974.
- Frege, G. (1892) Über Begriff und Gegenstand. In *Vierteljahrsschrift für wissenschaftliche Philosophie*, 16: 192–205. Translated as ‘Concept and Object’ by P. Geach in *Translations from the Philosophical Writings of Gottlob Frege*, P. Geach and M. Black (eds. and trans.), Oxford: Blackwell, third edition, 1980.
- Hofweber, T. (2005) Number determiners, numbers, and arithmetic. *The Philosophical Review*, 114(2), 179-225.
- Ionin, T. and O. Matushansky (2006) The composition of complex cardinals. *Journal of Semantics* 23, 315-360.
- Landman, F. (2003) Predicate-argument mismatches and the adjectival theory of indefinites. In Coene, M., & d'Hulst, Y. (eds.) *From NP to DP: The syntax and semantics of noun phrases* (Vol. 1). John Benjamins: Amsterdam. 211-237.
- Landman, F. (2004) *Indefinites and the type of sets*. Wiley (Blackwell): Oxford.
- Rothstein, S. (2009) Measuring and counting in Modern Hebrew. *Brill's Annual of Afroasiatic Languages and Linguistics*, Volume 1. 2009. 106-145.
- Rothstein, S. (2010). Counting and the mass count distinction. *Journal of Semantics*. 27.3: pp343-397

Acknowledgements:

This work was partially funded by Israel Science Foundation grant 1345/13 to Susan Rothstein.

At least* and Quantity Implicature: Choices and Consequences

Bernhard Schwarz

McGill University, Montreal, Quebec, Canada
bernhard.schwarz@mcgill.ca

1 Introduction

Numerals modified by *at least* lack the upper-bounding implication typically associated with bare numerals ([7, 9]). In contrast to (2), for example, (1) lacks the upper-bounding implication that Al did not hire more than two cooks.

- (1) Al hired at least two cooks.
- (2) Al hired two cooks.

At least also introduces an implication of speaker ignorance ([9, 1, 12, 19]). (1) conveys that the speaker is uncertain about the exact number of cooks Al hired. In one view, pioneered by Büring ([1]), this ignorance implication is a Gricean Quantity implicature. Büring proposes to capitalize on a striking similarity between *at least* and disjunction. Just like 1, 3 below lacks the upper-bounding implication of 2, and 3 also shares the ignorance implication of 1.

- (3) Al hired exactly two cooks or Al hired more than two cooks.

Assuming that ignorance implications with *at least* and disjunction alike are Quantity implicatures, [1] proposes that in the two cases these implicatures are also calculated in a parallel fashion, suggesting that statements with *at least* in some sense *are* disjunctions. But [1] stops short of executing this idea within a general theory of Quantity implicature. As a consequence, Büring's proposal remains unclear with regard to precise relation between *at least* and disjunction, and in fact about the logic of ignorance implications in either case.

Building on [19] and [10], I will in the following formulate, and explore the consequences of, a neo-Gricean implementation of Büring's approach. This account attempts to derive ignorance implications with *at least* through a straightforward adaptation of the treatment of disjunction by Sauerland in [15]. Such a derivation is suggested to call for a departure from the standard neo-Gricean algorithm for calculating Quantity implicatures ([15, 4, 5]), requiring the Neo-Gricean correspondent of a notion of implicature maximization ([16, 4, 20]).

2 Alternatives and the Standard Recipe

Under the neo-Gricean approach, listeners enrich the semantic meaning of an utterance by reasoning about formally defined *alternatives*, semantic meanings of certain alternative utterances that the speaker could have made but did not. The Neo-Gricean theory of alternatives is the

*Thanks to two anonymous AC 2013 reviewers for detailed feedback that has helped my understanding of the issues discussed here. Previously, I also learned from audiences at the University of Göttingen, the University of Konstanz, and McGill University. Thanks, in particular, to Luis Alonso-Ovalle, Brian Buccola, Emmanuel Chemla, Yosef Grodzinsky, Irene Heim, Rick Nouwen, and Doris Penka. This research was supported, in part, by grants SSHRC 410-2010-1264, FQRSC 2012-SE-144646, and SSHRC 435-2013-0592.

topic of subsection 2.2. Section 2.1 spells out a standard elaboration of Gricean Quantity-based reasoning about alternatives, known as the *Standard Recipe* (SR)¹. In the introduction of SR in subsection 2.1, the relevant alternatives are in the meantime simply assumed as given.

2.1 The Standard Recipe

2.1.1 Primary and Secondary Implicatures

According to SR, a listener interpreting an utterance infers that the speaker lacks the knowledge to support assertion of semantically stronger alternatives. Such *primary implicatures* take the form $\neg\Box\beta$, where β is an alternative and \Box expresses speaker certainty. So under a weak semantics for numerals, 2 semantically conveys that Al hired more than one cook, [2...). According to SR, if [3...) is an alternative, a listener infers from 2 that the speaker lacks the knowledge to support the assertion of [3...), and infers the primary implicature $\neg\Box[3...)$.

- (2) Al hired two cooks.

SR posits that listeners sometimes strengthen a primary implicature $\neg\Box\beta$ to a *secondary implicature*, $\Box\neg\beta$. This happens just in case the listener takes the speaker to be *competent* with regard to the alternative β , assuming $\Box\beta \vee \Box\neg\beta$. For (2), a listener who assumes $\Box[3...)\vee\Box\neg[3...)$ will infer $\Box\neg[3...)$, which captures the attested upper-bounding implication of 2.

2.1.2 Excludability and Symmetry

SR assumes that the listener takes the speaker to respect Quality, inferring $\Box\alpha$ from an utterance with semantic content α . I will refer to the conjunction of the Quality implicature and all primary implicatures as the *implicature base*. Also, if $\neg\Box\beta$ is a primary implicature, let me say that an alternative β is *weakly excludable* just in case the potential secondary implicature $\Box\neg\beta$ is consistent with the implicature base. SR posits weak excludability of β as a necessary condition for strengthening the primary implicature $\neg\Box\beta$ to the secondary implicature $\Box\neg\beta$.

If an alternative β fails to be weakly excludable, this is due to the implicature base entailing an ignorance implication $\neg\Box\beta \wedge \neg\Box\neg\beta$ that is inconsistent with the potential secondary implicature $\Box\neg\beta$ and the competence assumption $\Box\beta \vee \Box\neg\beta$ that would be required to support it. The paradigm case illustrating this is disjunction. Suppose that for 4 below, the listener strengthens the semantic meaning **b** or **c** by reasoning about the stronger alternatives **b** and **c**. Under SR, the listener's reasoning then yields the primary implicatures $\neg\Box\mathbf{b}$ and $\neg\Box\mathbf{c}$.

- (4) Al hired Bill or Al hired Carol.

But neither **b** nor **c** is weakly excludable. The reason is that in conjunction with the Quality implicature $\Box\mathbf{b} \vee \Box\mathbf{c}$, $\neg\Box\mathbf{b}$ entails $\neg\Box\neg\mathbf{c}$ and $\neg\Box\mathbf{c}$ entails $\neg\Box\neg\mathbf{b}$. So the implicature base entails the ignorance implications $\neg\Box\mathbf{b} \wedge \neg\Box\neg\mathbf{b}$ and $\neg\Box\mathbf{c} \wedge \neg\Box\neg\mathbf{c}$. These ignorance implications are inconsistent with the competency assumptions $\Box\mathbf{b} \vee \Box\neg\mathbf{b}$ and $\Box\mathbf{c} \vee \Box\neg\mathbf{c}$, and hence with the potential secondary implicatures $\Box\neg\mathbf{b}$ and $\Box\neg\mathbf{c}$. The weak excludability condition predicts, then, that neither secondary implicature will be inferred. All of these predictions about (4) are correct. An utterance of (4) does not support the inferences that either of **b** and **c** is false. Instead, it conveys that the speaker is uncertain about the truth or falsity of both **b** and **c**.

In SR, ignorance implicatures arise when, and only when, for a Quality implicature $\Box\alpha$ and two primary implicatures $\neg\Box\beta$ and $\neg\Box\gamma$, the conjunction of $\Box\alpha$ and $\neg\Box\beta$ entails $\neg\Box\neg\gamma$ and

¹See, in particular, [15, 4, 5]. The term *Standard Recipe* is due to Geurts ([5]). The terminology introduced below is for the most part taken from, or adapted from, language established in [15, 4, 5].

the conjunction of $\square\alpha$ and $\neg\square\gamma$ entails $\neg\square\neg\beta$. The Quality implicature and the two primary implicatures will relate in this way just in case β and γ jointly exhaust α in the sense that α entails $\beta \vee \gamma$. In that case, β and γ are also said to be *symmetric*.

2.2 Horn scales and the Substitution Method

In the Neo-Gricean approach, the availability of alternatives is grammatically regulated. Horn in [7] proposes that alternatives to the asserted utterance meaning are generated from sets of alternative lexical items, the *Horn scales*. Alternative meanings are the semantic meanings of syntactic structures obtained from the syntactic structure of the asserted sentence by substituting one or more lexical item with a Horn scale mate. Under this approach, the proposition [3...]) counts as an alternative to the asserted meaning of (2), [2...), because the numerals *two* and *three* are both members of the Horn scale of numerals in (5), and because substituting *three* for *two* in (1) results in a sentence with the meaning [3...]).

- (5) Horn scale: {*one*, *two*, *three*, ...}

3 At least and The Standard Recipe

Recall that (1), repeated below, conveys that the speaker is uncertain about the exact number of cooks Al hired. Büring ([1]) proposes that such ignorance implications are Gricean inferences, hence are not semantically entailed. If so, *at least* in (1) can be analyzed as truth conditionally inert, so that the semantic meaning of (1) is the same as that of (2), namely [2...).

- (1) Al hired at least two cooks.

3.1 Ignorance via Symmetry

Pursuing the approach pioneered by Büring ([1]), Schwarz and Shimoyama ([19]) and Mayr ([10]) articulate an account of ignorance implication with *at least* under SR by positing suitable alternatives. Suppose that the alternatives to the semantic meaning of (1), [2...), include [3...]) and also [2], the proposition that Al hired exactly two cooks. Since both of these alternatives are stronger than [2...), SR derives the primary implicatures $\neg\square[3...)$ and $\neg\square[2]$. Neither of these is expected to strengthen into a secondary implicature, since neither [3...] nor [2] is weakly excludable: [3...]) and [2] are symmetric, jointly exhausting [2...). The implicature base therefore entails the ignorance implications $\neg\square[3...)\wedge\neg\square\neg[3...)$ and $\neg\square[2]\wedge\neg\square\neg[2]$, which are inconsistent with the assumption that speaker is competent with regard to [3...]) or [2], hence inconsistent with the potential secondary implicatures $\square\neg[3...)$ and $\square\neg[2]$.

This predicts that a listener will infer from an utterance of (1) that the speaker lacks an opinion on whether Al hired exactly two cooks and on whether Al hired more than two cooks. Note that the implication derived amounts to *partial* ignorance. It is consistent with the speaker having an opinion regarding the number of cooks hired that goes beyond [2...). For example, it is consistent, with the speaker being certain that Al hired no more than ten cooks. All of these predictions appear consistent with intuitions about the meaning of (1).²

²Schwarz and Shimoyama ([19]) note that ignorance implications with *at least* are obviated in cases where *at least* is interpreted under a universal quantifier. For example, *Every manager hired at least two cooks*. is consistent with the speaker having full knowledge of the number of cooks hired by each manager. They also note that this ignorance obviation is predicted under SR, since, as [4] discusses, universal quantifiers break the symmetry between alternatives. Ignorance obviation under universals (and other operators, see [17, 10]) is unaccounted for in certain analyses of *at least*, including [6] and [12].

3.2 A Scale Mate for *at least*

The proposed application of the SR to *at least* rests on the assumption that the requisite alternatives are generated by the Horn-scale based substitution method. For (1), the substitution method and the Horn scale in 5 deliver the alternative [3...). What is missing is a source for the alternative [2]. A straightforward way to fill this gap, is the addition of the Horn scale in 6, which pairs *at least* with *exactly*. Substituting *exactly* for *at least* in (1) results in a structure that has the intended double sided meaning ([17, 10]).

- (6) Horn scale: {*at least*, *exactly*}

Note that under the account arrived at, a coherent interpretation emerges of Büring's proposal that *at least* has a meaning that is in some sense disjunctive. What *at least* has in common with disjunction is that both give rise to alternatives that instantiate the configuration of symmetry, jointly exhausting the semantic meaning of the assertion.³

4 Overgeneration

Assuming the two Horn scales in (5) and (6), the substitution method applied to (1) delivers not just the alternatives [2] and [3...). The full set of alternatives generated is **ALT** = {[n]: n≥2} ∪ {[m,...): n≥3}. Mayr ([10]) notes that the additional alternatives in **ALT** support an unattested secondary implicature.⁴ Applied to example (1), Mayr's claim is that SR derives the secondary implicature $\square\neg[4\dots)$, so that (1) is predicted to carry the upper bounding implication that Al did not hire more than three students. (1) should then convey that Al hired exactly two or three students. Clearly, (1) cannot be so understood.

While Mayr is pointing to a real problem, a closer look at the predictions under SR leads to a refinement of its characterization. Determining predictions can be facilitated by displaying **ALT** in the way shown in (7), where meanings are aligned to reflect semantic strength.

(7)	$\begin{array}{ccccccc} & & \cdots & \cdots & & & \\ & & [4] & [5] & \dots &) & \\ & [3] & [4] & 5 & \dots &) & \\ [2] & [3] & 4 & 5 & \dots &) & \text{alternatives} \\ \hline [2] & 3 & 4 & 5 & \dots &) & \text{asserted meaning} \end{array}$
-----	---

SR delivers the primary implicatures $\neg\square[3]$ and $\neg\square[4\dots)$. Neither of [3] and [4 ...) pairs up with the other, or any available alternative, in a configuration of symmetry to exhaust the semantic meaning [2...). Hence SR derives no ignorance implications for these alternatives, which are therefore weakly excludable. So nothing in SR prevents strengthening of either $\neg\square[3]$ or $\neg\square[4\dots)$ to $\square\neg[3]$ and $\square\neg[4\dots)$, respectively. According to SR, then, listeners might infer from (1) that Al did not hire exactly three cooks, or that he did not hire more than three.⁵ More generally, SR allows for the secondary implicatures $\square\neg[n]$ and $\square\neg[m\dots)$, for all n≥3 and m≥4. None of these additional Quantity inferences can be detected in the meaning of (1).

³In the case of *at least*, the two alternatives posited happen to be mutually exclusive. But mutual exclusivity is not, as [1] and [8] seem to suggest, a crucial component of a neo-Gricean derivation of ignorance implications.

⁴I first learned about this problem from Brian Buccola (personal communication, commenting on [17]).

⁵In fact, SR as stated does not exclude strengthening of both $\neg\square[3]$ and $\neg\square[4\dots)$ together, which would result in an inconsistent set of implicatures. As [4] notes (with regard to different types of examples), SR is not contradiction free.

5 Adjusting the Alternative Set?

Under SR, the absence of an unwanted upper-bounding inference $\Box\neg\beta$ can have two possible sources: first, it can be credited to the assumption that β is not available as an alternative in the first place; second, it can be credited to the assumption that β and some other alternative establish a configuration of symmetry, so that neither alternative is weakly excludable. I will consider both options, starting with the second.

5.1 Closure under Disjunction?

One strategy to preempt the unwanted secondary implicatures supported by the alternatives in (7) is to enlarge the alternative set in a way that lets every alternative participate in a configuration of symmetry, thereby ensuring that no alternative is weakly excludable. A principled way of doing this is to follow a suggestion in [20] to close alternative sets under disjunction.

Closure under disjunction will expand **ALT** into $\text{ALT}' = \{\forall S: S \subseteq \{[n]: n \geq 2\}\}$. In the case at hand, closure under disjunction will, for example, add the alternative $[2,3]$, which is symmetric to $[4,...)$, so that neither is weakly excludable; likewise, $[2,4,...)$ will be added, which is symmetric to $[3]$, so that again neither is weakly excludable. This generalizes to all the alternatives in **ALT'**. Closure of the alternative set under disjunction, then, correctly preempts derivation of any secondary implicatures for (1).

However, closure under disjunction eliminates unwanted secondary implicatures by virtue of supporting additional ignorance implications. Thus the particular alternatives considered above will yield the ignorance implications $\neg\Box[4,...) \wedge \neg\Box\neg[4,...)$, $\neg\Box[2,3] \wedge \neg\Box\neg[2,3]$, $\neg\Box[3] \wedge \neg\Box\neg[3]$, and $\neg\Box[2,4,...) \wedge \neg\Box\neg[2,4,...)$. Since in fact every element of **ALT'** participates in a symmetric pair, closure under disjunction predicts (1) to carry an implication of *total* ignorance modulo the asserted meaning. That is, (1) is predicted to convey that, beyond the assertion that Al hired more than one cook, the speaker has no opinion at all regarding the number of cooks hired. As already noted, however, (1) only conveys partial ignorance, concerning the question whether Al hired exactly two cooks or more than two. Intuitions on (1) furnish no evidence for the total ignorance derived by closure under disjunction. Closing the alternative set under disjunction, then, does not appear to be a promising solution to the overgeneration problem.⁶

5.2 A one-scale analysis?

What emerges from the above is that in order for SR to apply correctly to (1), it is necessary for Quantity inferences to be drawn on the basis of $[2]$ and $[3,...)$ alone. The question is then whether a general theory of alternatives can be motivated that delivers the alternative set $\text{ALT}'' = \{[2], [3,...)\}$. Krifka's ([9]) proposal that *at least* blocks projection of alternatives generated in its scope provides one possible part of such a theory. To be sure, though, along with the unwanted alternatives $[n,...)$, for all $n \geq 4$, this also eliminates the crucial alternative $[3,...)$. But $[3,...)$ can conceivably be recovered by adopting the proposal that in generating alternatives, the possible substitutes for *at least* include the comparative operator *more than* (see [3] and [8]). In a Neo-Gricean implementation, this amounts to replacing the Horn scale in (6) above with (8) below.

- (8) Horn scale: {*at least*, *exactly*, *more than*}

⁶Closure under disjunction is also insufficient to exclude unwanted secondary implicatures across-the-board. Additional secondary implicatures are incorrectly expected to emerge in cases like *Every manager hired at least two cooks*, where a higher universal quantifier breaks symmetry and hence obviates ignorance implications (see previous footnote). Mayr ([10]) discusses a version of this problem in a related proposal.

Unfortunately, while this “one-scale analysis” analysis derives the correct set of implicatures for examples like (1), it fails to derive the intended ignorance implications in other cases with *at least*. As emphasized in [9] and [2], the syntactic distribution of *at least* resembles that of the focus sensitive particle *only*. One manifestation of this is the ability of *at least* to associate with a numeral at a distance. This is illustrated by (9), where *at least* can be read as associating with the numeral across a possibility modal.

- (9) Al is at least allowed to hire two cooks.

In analogy to (1), an utterance of (9) can be read as conveying that the speaker is uncertain about whether Al is allowed to hire more than two cooks. Writing \diamond to express deontic possibility, this ignorance implication can be captured under SR by assuming that (9) semantically conveys $\diamond[2,\dots]$, the proposition that Al is allowed to hire more than one cook, and by positing the symmetric alternatives $\diamond[2] \wedge \neg\diamond[3,\dots]$, and $\diamond[3,\dots]$, that is, the proposition that Al is only allowed to hire two cooks and the proposition that he is allowed to hire more than two. The problem for a one-scale analysis based on the Horn scale in (8) is that the distribution of comparative operators like *more than* differs greatly from that of *at least* (see [6]). In particular, the result of substituting *more than* for *at least* in (9), shown in (10), is ungrammatical. A one-scale analysis based on (8), then, seems to leave the $\diamond[3,\dots]$ alternative unaccounted for.

- (10) *Al is more than allowed to hire two cooks.

A two-scale analysis based on the Horn scales in (5) and (6) fares better in this respect. The alternative $\diamond[3,\dots]$ for the interpretation of (9) can be taken to arise in the very same way as the alternative $[3,\dots]$ for the interpretation (1), viz. by substituting the numeral *three* for the numeral *two*.⁷

This advantage of the two-scale analysis is a reason to explore a solution of the overgeneration problem that retains the two Horn scales (5) and (6) and locates the source of the problem elsewhere.

6 Strong Excludability

We are back to the assumption that the set of alternatives for the interpretation of (1) is ALT, the set of all meanings generated by the substitution method from the two Horn scales (5) and (6). The remaining analytical option is then to revise SR itself. This can be done by replacing weak excludability with a more restrictive property, call it *strong excludability*. The revised algorithm, call it the *Revised Standard Recipe* (RSR), posits strong excludability as a necessary condition for strengthening of a primary implicature to a secondary implicature.

A possible blueprint for the definition of strong excludability is given in [4] (see also [22, 16, 20]). If $\neg\Box\beta$ is a primary implicature, let us say that β is *strongly excludable* just in case the potential secondary implicature $\Box\neg\beta$ is an element of every maximal set of potential secondary implicatures that is consistent with the implicature base.

Note first that by this definition, an alternative that is not weakly excludable cannot be strongly excludable, either. If a potential secondary implicature $\Box\neg\beta$ is itself inconsistent with the implicature base, then it will not be an element of any maximal set of potential secondary

⁷The sentence **Al is exactly allowed to hire two cooks* is actually no better than (10), putting into question the viability of both Horn scales for *at least* under consideration. An amendment that suggests itself is to replace the Horn scale *{at least, exactly}* with *{at least, only}*. *Only* delivers the intended double-sided alternatives for all cases considered here, including (9). *Al is only allowed to hire two cooks* expresses the intended meaning $\diamond[2] \wedge \neg\diamond[3,\dots]$.

implicatures consistent with the implicature base, let alone every such set. For the case at hand, this guarantees that [2] and [3,...) are not strongly excludable. Just like SR, therefore, SRS preempts strengthening of the primary implicature $\neg\Box[2]$ and $\neg\Box[3,...)$.

Consider now the set of potential secondary implicatures $\{\Box\neg[4,...), \Box\neg[4], \Box\neg[5,...), \Box\neg[5], ...\}$. This set is consistent with the implicature base, but adding any further potential secondary implicature to it ($\Box\neg[2]$, $\Box\neg[3,...)$, or $\Box\neg[3]$) would render it inconsistent with the implicature base; hence it is a maximal set of potential secondary implicatures consistent with the implicature base; since it does not contain $\Box\neg[3]$, it shows that [3] is not strongly excludable. Next consider the set of potential secondary implicatures $\{\Box\neg[3], \Box\neg[5,...), \Box\neg[5], \Box\neg[6,...), \Box\neg[6], ...\}$. This set is again consistent with the implicature base, but adding any further potential secondary implicature to it ($\Box\neg[2]$, $\Box\neg[3,...)$, $\Box\neg[4,...)$, or $\Box\neg[4]$) would render it inconsistent with the implicature base; hence it is again a maximal set of potential secondary implicatures consistent with the implicature base; since it does not contain [4,...) or [4], it shows that neither [4,...) nor [4] is not strongly excludable, either.

Extrapolating from these examples, it is apparent that in fact no alternative in **ALT** is strongly excludable. RSR therefore bars strengthening of any primary implicatures, hence correctly predicts that (1) does not trigger any Quantity inference beyond the ignorance implications $\neg\Box[2]\wedge\neg\Box\neg[2]$ and $\neg\Box[3,...)\wedge\neg\Box\neg[3,...)$.

7 Conclusion

The main finding of this paper is that a neo-Gricean account of *at least* seems to call for a departure from the Standard Recipe: strengthening to secondary implicature requires strong excludability of the relevant alternative, rather than mere weak excludability. Under the resulting Revised Standard Recipe, unattested strengthening to secondary implicature may or may not be due to an ignorance implication. Failure of weak excludability results in ignorance implications, whereas mere failure of strong excludability does not.

The question that remains is whether strong excludability is indeed a plausible ingredient of a Gricean algorithm for the calculation of ignorance. It is also notable that strong excludability is transparently the neo-Gricean counterpart of Fox's ([4]) notion of *innocent exclusion*, an ingredient of his grammatical theory of scalar implicature. If strong excludability is deemed to be an unlikely component of Gricean reasoning, the results here could be taken to indicate that ignorance implications should be accommodated in a grammatical theory of Quantity implicature. In fact, the effects of the Revised Standard Recipe arrived at above are fully replicated in Meyer's ([11]) *Matrix K Theory*, a radically grammatical theory of Quantity implicature in that all kinds of implicatures, including ignorance implications, are derived in the grammar.

However, before adducing *at least* data as evidence for or against any theory of Quantity implicature, a good amount of work remains to be done to establish that an implicature account of the sort pioneered by Büring ([1]) is feasible in view of a broader set of data. It remains to be seen, in particular, whether the account successfully extends to *at most*, association of *at least* and *at most* with non-numerals, and in particular the bewildering interaction of *at least* and *at most* with modals and other operators.⁸

⁸See [1, 6, 19, 12, 14, 21, 13, 17, 18, 10, 2, 8].

References

- [1] Daniel Büring. The least *at least* can do. In Charles B. Chang and Hannah J. Haynie, editors, *Proceedings of WCCFL 26*, pages 114–120, Somerville, MA, 2008. Cascadilla Proceedings Project.
- [2] Elizabeth Coppock and Thomas Brochhagen. Raising and resolving issues with scalar modifiers. *Semantics & Pragmatics*, 6(3):1–57, 2013.
- [3] Chris Cummins and Napoleon Katsos. Comparative and superlative quantifiers: Pragmatic effects of comparison type. *Journal of Semantics*, 27(3):271–305, 2010.
- [4] Danny Fox. Free choice and the theory of scalar implicatures. In Uli Sauerland and Penka Statteva, editors, *Presupposition and implicature in compositional semantics*, pages 71–120. Palgrave MacMillan, 2007.
- [5] Bart Geurts. *Quantity Implicatures*. Cambridge University Press, Cambridge, 2011.
- [6] Bart Geurts and Rick Nouwen. *At least* et al.: the semantics of scalar modifiers. *Language*, pages 533–559, 2007.
- [7] Laurence Horn. *On the semantic properties of logical operators in English*. PhD thesis, University of California, Los Angeles, 1972.
- [8] Christopher Kennedy. A de-fregean semantics for modified and unmodified numerals. manuscript, University of Chicago, 2013.
- [9] Manfred Krifka. At least some determiners aren't determiners. *The semantics/pragmatics interface from different points of view*, 1:257–291, 1999.
- [10] Clemens Mayr. Implicatures of modified numerals. In Ivano Caponigro and Carlo Cecchetto, editors, *From grammar to meaning: the spontaneous logicality of language*, pages 139–171. Cambridge University Press., 2013.
- [11] Marie-Christine Meyer. *Ignorance and Grammar*. PhD thesis, Massachusetts Institute of Technology, 2013.
- [12] Rick Nouwen. Two kinds of modified numerals. *Semantics and Pragmatics*, 3:1–50, 2010.
- [13] Rick Nouwen. Superlative modifiers, ignorance implicatures and modals. Unpublished handout for colloquium talk at MIT, April 2011.
- [14] Doris Penka. A superlative analysis of superlative scalar modifiers. Unpublished talk handout for *Sinn und Bedeutung 15*, September 2010.
- [15] Uli Sauerland. Scalar implicatures in complex sentences. *Linguistics and Philosophy*, 27:367–391, 2004.
- [16] Katrin Schulz and Robert Van Rooij. Pragmatic meaning and non-monotonic reasoning: The case of exhaustive interpretation. *Linguistics and Philosophy*, 29(2):205–250, 2006.
- [17] Bernhard Schwarz. Remarks on Class B numeral modifiers. Unpublished handout for talk presented at *Indefinites and Beyond*, University of Göttingen, November 2011.
- [18] Bernhard Schwarz, Brian Buccola, and Michael Hamilton. Two types of class B numeral modifiers: A reply to Nouwen 2010. *Semantics and Pragmatics*, 5(1):1–25, 2012.
- [19] Bernhard Schwarz and Junko Shimoyama. Negative Islands and obviation by *wa* in Japanese Degree Questions. In David Lutz et al., editor, *Proceedings of SALT 20*, Ithaca, NY, 2010. Cornell University.
- [20] Benjamin Spector. Scalar Implicatures: Exhaustivity and Gricean Reasoning. In Maria Aloni Paul Dekker and Alastair Butler, editors, *Questions in Dynamic Semantics*, pages 225–250. Elsevier, 2007.
- [21] Benjamin Spector. *At least* at last? Unpublished handout, École Normale Supérieure, Paris, March 2011.
- [22] Robert Van Rooij and Katrin Schulz. Exhaustive interpretation of complex sentences. *Journal of logic, language and information*, 13(4):491–519, 2004.

Semantics of DP islands*

Alexandra Simonenko

McGill University
alexandra.simonenko@mail.mcgill.ca

Abstract

This paper is concerned with the role of the semantics of definite articles in the analysis of e(xtended)NP-island effects. It brings in a new sort of evidence confirming the proposal of [25] for German and [6] for English that the semantics of the so called “strong” articles and demonstratives differs qualitatively from that of “weak” articles and *the* in that the former involves a pronominal element which makes corresponding eNPs directly referential. Islandhood for wh-subextraction thus emerges as a new diagnostic for direct referentiality.

1 Introduction

This paper provides a semantic-pragmatic analysis of the ban on wh-subextraction out of a certain type of eNP. Austro-Bavarian German features two distinct definite articles which contrast with respect to the wh-subextraction ban. Namely, wh-subextraction is possible out of a weak-eNP in (1), but leads to ungrammaticality in (2) (in any context).

- (1) Vo wem host du [s Possbldl t] gsegn?
of whom have you det_w passport.photo t seen
'Who did you see the passport picture of?'
- (2) *Vo wem host du [des Possbldl t] gsegn?
of whom have you det_s passport.photo t seen
Intended: 'Who did you see that passport picture of?' (adapted from [29])

The problem of the impossibility of establishing dependencies between a wh-word and a trace embedded within certain kind of eNPs has been around at least since [4]. In [7], [19], [22], [14] the notions of specificity and definiteness has been proposed as relevant for the wh-subextraction ban. To my knowledge, there has been no investigation as to what goes wrong on the semantic side. I propose that the locus of the contrast between (1) and (2) is the interaction of the semantics of a strong definite article with requirements pragmatically imposed on the set of possible answers to the question. I demonstrate that subextraction out of strong-eNPs results in questions which cannot serve to update the inquirer's state of knowledge: both the asserted and the presupposed contents of possible answers to such questions have to be part of the inquirer's knowledge prior to uttering the question. I propose that questions with what I call zero information-seeking potential are ungrammatical, unlike rhetorical and biased questions, which, although not updating the inquirer's state of knowledge in some contexts, can serve this purpose in others. This project aligns with a number of recent studies of constraints on movement whereby such constraints are analyzed as stemming from semantic-pragmatic constraints on question formation ([9], [24] and references therein), rather than from syntactic factors. The analysis proposed in this paper can be extended onto a number of other languages

*I am very thankful to Bernhard Schwarz for all his help and ideas and to Martina Wiltschko for discussions of the Austro-Bavarian data. Thanks to Luis Alonso-Ovalle and the audience of NELS44 for discussions. I am grateful to Bettina Gruber, Walter Pedersen, and Maksim Kudrinsky for help with the data.

which feature eNPs contrasting with respect to the wh-subextraction possibilities, in particular, onto English, French, Italian, Komi, Russian, and Spanish where demonstrative-eNPs are wh-subextraction islands. Islandhood with respect to wh-subextraction thus emerges as a new diagnostic for direct referentiality.

In the next section I introduce the Austro-Bavarian strong and weak article paradigms followed by the discussion of articles' semantics based on the proposal of [25]. Then I will proceed to the analysis of the interaction of articles' semantics with the semantics of questions.

2 Weak and strong definite articles in Austro-Bavarian

Austro-Bavarian is among varieties of German that distinguish between two types of definite articles, the “weak” and the “strong” series ([23], [29]). Throughout this paper I refer to nominal expressions headed by strong and weak articles as strong-eNPs and weak-eNPs respectively.

According to the generalizations made in [25] for Standard German and [29] for Austro-Bavarian, descriptively, a strong article is used in contexts where the eNP in question has an anaphoric or deictic antecedent, as in (3). A weak article is used whenever it is part of the common ground that the eNP has only one possible referent in a given situation, as in (4).

(A points to a house (the only one in the immediate surrounding) and asks B,)

- (3) Gfoit da s'/#des Haus?
like you det_w/det_s house
'Do you like this house?'
- (4) Wie geht 's 'n da/*dea Frau?
how goes it prt det_w/det_s woman
'How is your wife doing?' ([29, 7])

[25] essentially proposes a classic Fregean treatment for the weak article, supplemented with a machinery of domain restrictions, which results in picking out a unique individual with a relevant property within a relevant domain.¹

$$(5) \quad \llbracket D_w \rrbracket = \lambda P_{<e,t>} : \exists!x[P(x)] . \iota x[P(x)] \quad (\text{based on [25]})$$

The strong article crucially, has an “enriched” LF compared to the weak one in that it includes a silent individual pronoun (here it bears the index 1).

$$(6) \quad [1 [D_s NP]] \quad (\text{based on [25]})$$

The denotation of the strong article differs from (5) in that it involves an individual argument. Namely, the function denoted by the strong article takes a property of individuals P, an individual argument y, and, if defined, returns a unique individual with property P identical to the argument individual y.

$$(7) \quad \llbracket D_s \rrbracket = \lambda P_{<e,t>} . \lambda y : \exists!x[P(x) \ \& \ x = y] . \iota x[P(x) \ \& \ x = y] \quad (\text{based on [25]})$$

Notice that (7) can be rewritten equivalently as in (8).

$$(8) \quad \llbracket D_s \rrbracket = \lambda P_{<e,t>} . \lambda y : P(y) . y$$

¹[25] implements his analysis of German articles in the situation semantics framework of [18]. In order to simplify the discussion I omit the situational aspect of Schwarz's (2009) model (undeniably a crucial one for the treatment of weak-eNPs).

In the next section I examine the semantics of wh-subextraction out of a strong-eNP, showing that the crucial property of a strong article is that it makes $\llbracket i D_s NP \rrbracket^g$, if defined, to be identical to a particular individual $g(i)$.

3 Semantics of wh-subextraction out of strong-eNP

In this section I look in detail at how the semantics of articles interacts with the interpretation of questions. I assume that the semantics of questions involves two main ingredients: the semantics of the question word and the semantics of its complement (roughly, TP). I first discuss the contribution of the strong article to the interpretation of the TP, and then show why such a TP cannot be part of a grammatical question.

Given the LF in (6), a question with a strong-eNP has the following LF, assuming that as a result of movement, the wh-phrase becomes a sister to TP, which results in lambda-abstraction over TP.²

- (9) who $\lambda_1 TP[\text{you have seen } [2 [\text{des passport picture of } t_1]]]$

Assuming (8), the denotation of the object eNP is as follows.

- (10) $\llbracket 2 \text{ des picture of } t_1 \rrbracket^{g,w}$ is defined iff $g(2)$ is a passport picture of $g(1)$ in w
if defined, $\llbracket 2 \text{ des picture of } t_1 \rrbracket^{g,w} = g(2)$

Assuming, as in [12], that presuppositions introduced by subconstituents become restrictions on the domain of the function denoted by a larger expression, the presupposition introduced by the strong-eNP that $g(2)$ is a passport picture of $g(1)$ turns into a restriction on the domain of the function denoted by the whole TP, as shown below.

- (11) $\lambda x . \llbracket \text{you have seen } [2 \text{ des picture of } t_1] \rrbracket^{g[1 \rightarrow x],w} =$
 $= \lambda x : \text{in } w, g(2) \text{ is a passport picture of } x . \text{the hearer has seen } g(2) \text{ in } w$

As a denotation of TP, we obtain a function which, if defined, maps an individual to truth in case the hearer has seen the picture $g(2)$. Notice that relative to any given possible world this is a constant function. In what follows I argue that a question has to be formed based on a function with a variable output in a given world.

I assume a version of the semantics of questions of [16], whereby a wh-word such as *what* or *who* denotes a function from open propositions (denotation of the complement of the wh-word, our TP) to a function from worlds to sets of propositions. Each proposition in this set corresponds to the function denoted by the complement of the wh-word with its argument slot filled by some individual from the domain of the wh-word. The following is an adaptation of Karttunen's (1977) wh-semantics.³

- (12) $\llbracket \text{[who]} \rrbracket = \lambda f_{<s,<e,t>>} . \lambda w . \{ p : \exists x \in D_e [x \text{ is a person in } w \ \& \ p = \lambda w' . f(w')(x)] \}$

²Both in English and in Austro-Bavarian I assume that the preposition *vo(n)/of* is semantically vacuous. It therefore does not make a semantic difference whether the preposition is stranded, as in English, or carried along with the wh-word, as in Austro-Bavarian. Also, here and throughout the paper, I use English words in LF and semantic formulae for the clarity of exposition.

³In Karttunen's (1977) original version the answer-set contained only *true* answers. In addition, the original version involves formation of a proto-question, a singleton set containing an (open) proposition corresponding to denotation of TP. The proto-question then serves as an argument to the wh-function. The original version was designed to accommodate cases of multiple wh-words. Since I am not concerned with those here, a simplified version in (12) is enough for the present purposes.

Let us see how this works for the case at hand. Let us assume that the domain D_e in (2) is a set of three individuals {Hans, Elsa, Otto}. Assuming that $\llbracket who \rrbracket$ combines with the denotation of its complement by the intensional version of Function Application spelled out in [12, 308] and given (11), we obtain the following extension of the question in a world w where there are three individuals, Hans, Elsa, and Otto, in the domain of the *wh*-word.

$$(13) \quad \llbracket who \rrbracket^g(\lambda w. (11)) = \{[\lambda w': \text{in } w', g(2) \text{ is a passport picture of } x . \text{the hearer has seen } g(2) \text{ in } w']: x \in \{\text{Hans, Elsa, Otto}\}\}$$

Answer-propositions in (13) differ only in the presupposition associated with each of them, whereas their asserted content is identical. Intuitively, it seems that possible answers to a question should differ precisely in what is not yet presupposed. We ended up with possible answers of this sort because the individuals from the domain of the *wh*-word make no semantic contribution to the asserted content of answer-propositions: whatever value is assigned to t_1 when interpreting the object eNP, it ends up denoting the individual $g(2)$, by (8).

The question is now what, given the general theory of questions, is wrong with a question all possible answers to which are identical in what they assert? The following discussion consists of two parts. First, I propose that the asserted content of the possible answers gets presupposed by such a question. Second, I demonstrate that accommodation of the presupposed content of the answers cannot be a source of new information either, because inquirers cannot introduce presuppositions they are ignorant about.

3.1 Question's “existential” presupposition

When asking the *who*-question in (2), the point of departure for the inquirer is that the hearer has seen someone's picture. Following up on similar intuitions, [13] and [16], among others, assumed that questions carry something like an existential presupposition. A *wh*-question of the form *wh TP* seems to presuppose $\exists x[\llbracket TP \rrbracket(x) = 1]$, that is, that there exists an individual which has the property denoted by TP. While it was shown that the alleged existential presupposition of an embedded question projects through negation and presuppositional holes, questions do not pass all the presupposition tests in that they are defeasible, [10]. The same is true for Austro-Bavarian. Instead of denying the presence of an existential presupposition in questions altogether, negative answers to *wh*-questions have been taken to convey that the presupposition is not met, as in [5], [20]. I follow [1] who proposed that such defeasible presuppositions are generated every time we deal with a grammatical object whose semantics makes reference to a set of alternatives. The presence of such a set in the semantic representation triggers a default presupposition that a proposition corresponding to the disjunction of the propositions in the alternative set is true. This creates the effect of an existential presupposition.⁴

Let us look from this perspective at (2). The requirement that there be a true answer is satisfied in case (14) is true (existential closure of the presuppositions of possible answers; for at least one answer to be true, the function denoted by at least one answer has to be defined) and (15) is true.

$$(14) \quad \lambda w . \exists x \in \{\text{Hans, Elsa, Otto}\} [\text{in } w, g(2) \text{ is a passport picture of } x]$$

$$(15) \quad \lambda w . \text{the hearer has seen } g(2) \text{ in } w$$

⁴The existential presupposition generated by a question is an implicit part of the analyses of *wh*-island effects which rely on Dayal's (1996) principle of the existence of a maximally true answer, such as [9], [24] and references therein. The requirement that there exist a maximally true answer entails that there must exist a true answer, which is the same as the existential presupposition in the sense of [1].

How does the “existential” presupposition relate to the felicity conditions on the use of the question in (2)? I will make use of Stalnaker’s (1974) notion of a “context set” as a set of possible worlds c where all the propositions believed to be true by conversation participants, and which form the Common Ground, are true. According to [28, 4], who develops the insight of [27] about “bridging” semantic well-formedness and pragmatic felicity, it is an “irreducible property of natural language” that an utterance is felicitous only in case the function it denotes is defined in all the context set worlds. I assume that Stalnaker’s bridging principle is extendable onto cases of “floating” presuppositions such as the one proposed by [1]. In our specific case, this means that in order for the question in (2) to be asked felicitously, (14) and (15) have to be true in all the worlds in the context set. What makes such a question ungrammatical, I argue, is a conflict between the felicity conditions and some cooperativeness principle such as a version of the Gricean informativeness maxim of the kind “don’t ask a question when no answer will provide you with new information”. This conflict is going to arise in any Common Ground. I will refer to the relevant cooperativeness principle as Question information-seeking potential. I propose that a question has a zero information-seeking potential when no proposition in its denotation can change the inquirer’s state of knowledge in any Common Ground.

We find a similar effect in simple declaratives: an utterance whose asserted content is part of what is required to be part of the Common Ground for the utterance to be felicitous sounds odd, as observed by [2]. On this approach we can explain the infelicity of (16) as arising from the presupposition (that there exist two students, triggered by *both*) entailing the assertion of existence.

- (16) *There are both students.

3.2 The presuppositions of the answers

A possible objection to the claim that a question which gives rise to answers with identical asserted content entailed by the Common Ground has zero information-seeking potential is that the informational import of an answer might be in its presuppositional part. It is well known that conversation participants are able to accommodate presuppositions (at least since [27], [15]). In the case at hand one could say that updating of the Common Ground can be achieved not by virtue of adding the asserted content of an answer as such, but a) either by virtue of accommodating the existential presupposition of the question or b) by virtue of accommodating the presuppositional content of the answers.

We can rule out the first option right away since what it means is that it is not part of the Common Ground that there is a true answer to the question. But, according to [1], an expression associated with a set of alternatives automatically makes it part of the Common Ground that one of the alternatives is true. That is, it does not depend on the state of knowledge of a particular inquirer and/or addressee. The second option, namely, updating the Common Ground with answers’ presupposition, requires some more discussion. From looking at (17), we can conclude that presuppositions of all answers have to be entailed by the Common Ground in order for a question to be felicitous: it has to be part of the Common Ground that all individuals in the relevant domain actually have a hat (by virtue of the definedness condition on the function denoted by an eNP with a Saxon genitive that there exists a unique individual with the nominal property which belongs to the possessor).

- (17) Among Peter, Bill, and Sam, whose hat do you like best?

Possible answers to (17) are *I like Peter’s hat*, *I like Bill’s hat*, and *I like Sam’s hat*. These

answers presuppose that Peter has a hat, Bill has a hat, and Sam has a hat, respectively. Since it is strange to ask this question with regard to Peter, Bill, and Sam if only the former two have a hat, while Sam does not, we can conclude that possible answers' presuppositions project universally onto the level of question semantics. What is crucial for the current discussion is that the accommodation can happen only on the side of the person to whom the question is addressed, as the infelicity of the following exchange shows where A accommodates that Peter has a hat.

- (18) A: Among Peter, Bill, and Sam, whose hat do you like best? B: Peter's. A: #Oh, I didn't know Peter had a hat.

Getting back to wh-subextraction out of strong-eNP, for the question in (13) to be felicitous in a given Common Ground, the following propositions have to hold: *that g(2) is a passport picture of Hans*, *that g(2) is a passport picture of Elsa*, and *that g(2) is a passport picture of Otto*.⁵ While the addressee can still be ignorant the presuppositional content and simply ready to accommodate it, the inquirer cannot be ignorant about a presupposition that he introduces.⁶ This rules out the possibility that answers in (13) can be informative for the inquirer by virtue of their presuppositional content. This means that a felicitously uttered question of this kind cannot update the inquirer's state of knowledge in any Common Ground. This makes (2) different from rhetorical and biased questions (e.g. *Did Peter (even) lift a finger to help you?*), which are uninformative only with respect to some Common Grounds ([3], [11]).

We also encounter questions with truly zero information-seeking potential in the case of wh-extraction out of complements of factive verbs which contain the so called "one-time-only" predicates, as discussed in [21].

- (19) #Who does Max know that Alice got married to on June 1st?

If we assume that it is presupposed that at least one answer in the answer-set is true, in the set corresponding to the denotation of (19) there is only one true answer since there is only one *assertable* answer (e.g. *Max knows that Alice got married to John on June 1st* which presupposes *Alice got married to John on June 1st*), as the presuppositions of the possible answers are mutually exclusive. That is, for (19) to be uttered felicitously, a Common Ground has to entail all the contents of its only assertable answer. [21] dubs such context-independent pragmatic infelicity non-contingent, as opposed to contingent infelicity which obtains only in specific contexts. Wh-subextraction out of strong-eNPs is another case of non-contingent infelicity.

3.3 Wh-subextraction out of weak-eNP

Now that the proposal for why it is bad with strong-eNP has been fully fleshed out, let us take a look at the grammatical wh-subextraction out of weak-eNPs in (1). Given (5), the denotation of the TP, combined with the semantics of the wh-word, gives the following denotation for the

⁵Note that in this particular case there is an additional potential source of infelicity, namely that the universal projection of answers' presuppositions entails that g(2) is a passport picture of all three, which is pragmatically implausible. However, this cannot be the main source of infelicity as wh-subextraction is still ungrammatical if *passport picture* is replaced with *picture* or with a different noun altogether.

⁶Moreover, as [26, 319] note, "attempts to answer questions using presuppositions ... are typically infelicitous".

(i) Q: *What's the weather like?* A: #Bob realizes/doesn't realize that it's raining

Even though technically *Bob realizes that it's raining* is not an answer in Hamblin/Karttunen's sense and the inquirer is not responsible for introducing the presupposition that it's raining, the addressee is still not supposed to convey novel information by making the inquirer accommodate a presupposition.

question with the weak article.

- (20) $\llbracket(1)\rrbracket = \{[\lambda w' : \text{in } w', \text{ there exists a unique passport picture of } x . \text{ the hearer has seen the unique passport picture of } x \text{ in } w'] : x \in \{\text{Hans, Elsa, Otto}\}\}$

Obviously, these answers assert different things. The requirement that one of the alternatives be true, combined with universal projection of the presuppositions of the answers, amounts in this case to the truth of the following proposition.

- (21) $\lambda w . \forall x \in \{\text{Hans, Elsa, Otto}\} [\text{there exists a unique passport picture of } x \text{ in } w] \wedge \exists x \in \{\text{Hans, Elsa, Otto}\} [\text{the hearer has seen the unique passport picture of } x \text{ in } w]$

This proposition, if entailed by the Common Ground, does not make an answer such as *I have seen the passport picture of Hans* uninformative.

This analysis makes a prediction that if under some conditions strong/demonstrative-eNPs were to behave like weak/*the*-eNPs in that their denotation would be covarying with a quantifier-bound variable, wh-subextraction should be repaired. Such conditions happen to exist, one of them being modification of the head noun by a restrictive relative clause, [17], and the prediction is borne out, as (22) shows, where wh-subextraction out of a strong-eNP is possible.

- (22) Vo wem₁ hot da Hons [des Possbündl t₁ [was a jo söwa gmocht hot]]
of whom has det_w Hans det_s passport.pict. t that he prt himself made has
brocht?
brought
‘Who did Hans bring that picture of that he made himself?’

4 Conclusions

I proposed that the key to the pathology of wh-subextraction out of strong-eNPs is the semantics of the strong article in the version of [25], which anchors the denotation of the eNP-of-extraction to a given individual and thus precludes the assertive content of answer-propositions from covarying with the wh-bound variable. From the cross-linguistic perspective, then, the impossibility of wh-subextraction out of certain eNPs can serve as a diagnostic for the semantic nature of the article involved. Namely, the impossibility of wh-subextraction out of a given eNP indicates its direct referentiality. According to my preliminary investigation, eNPs headed by demonstratives are islands for wh-subextraction in English, Italian, Spanish, Russian, French, and Komi. One immediately relevant research area in this respect are the so called referential uses of indefinites, discussed in [8]. Another highly relevant testing ground for the proposal are cases of subextraction out of eNPs in non-questions, such as topicalization and relativization-related movements.

References

- [1] Dorit Abusch. Presupposition triggering from alternatives. *Journal of Semantics*, 27(1):37–80, 2010.
- [2] Jon Barwise and Robin Cooper. Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4(2):159–219, 1981.

- [3] Ivano Caponigro and Jon Sprouse. Rhetorical questions as questions. In *Proceedings of Sinn und Bedeutung*, volume 11, pages 121–133, 2007.
- [4] Noam Chomsky. Conditions on transformations. In *Essays on Form and Interpretation*. North Holland, New York, 1973.
- [5] Veneeta Dayal. *Locality in WH quantification: Questions and relative clauses in Hindi*. Kluwer Academic Publishers, 1996.
- [6] Paul Elbourne. Demonstratives as individual concepts. *Linguistics and Philosophy*, 31(4):409–466, 2008.
- [7] Robert Fiengo and James Higginbotham. Opacity in NP. *Linguistic Analysis*, 7(4):395–421, 1981.
- [8] Janet Dean Fodor and Ivan A. Sag. Referential and quantificational indefinites. *Linguistics and Philosophy*, 5(3):355–398, 1982.
- [9] Danny Fox and Martin Hackl. The universal density of measurement. *Linguistics and Philosophy*, 29(5):537–586, 2006.
- [10] Jeroen Groenendijk and Martin Stokhof. *Studies on the Semantics of Questions and the Pragmatics of Answers*. PhD thesis, Universiteit van Amsterdam, 1984.
- [11] Elena Guerzoni. *Why Even Ask?* PhD thesis, Massachusetts Institute of Technology, 2003.
- [12] Irene Heim and Angelika Kratzer. *Semantics in Generative Grammar*. Blackwell, 1998.
- [13] Laurence R. Horn. *On the Semantic Properties of Logical Operators in English*. PhD thesis, UCLA, 1972.
- [14] Ángel L. Jiménez-Fernández. A new look at subject islands: The phasehood of definiteness. *Anglica Wratislaviensis*, 50:137–168, 2012.
- [15] Lauri Karttunen. Presupposition and linguistic context. *Theoretical Linguistics*, 1(1):181–194, 1974.
- [16] Lauri Karttunen. Syntax and semantics of questions. *Linguistics and Philosophy*, 1(1):3–44, 1977.
- [17] Jeffrey C. King. *Complex Demonstratives: A Quantificational Account*. MIT Press, 2001.
- [18] Angelika Kratzer. An investigation of the lumps of thought. *Linguistics and Philosophy*, 12(5):607–653, 1989.
- [19] Maria-Rita Manzini. A minimalist theory of weak islands. *Syntax and Semantics*, pages 185–210, 1998.
- [20] Clemens Mayr, To appear. Intervention effects and additivity.
- [21] David Y. Oshima. On Factive Islands: Pragmatic anomaly vs. pragmatic infelicity. In *New Frontiers in Artificial Intelligence*, pages 147–161. Springer, 2007.
- [22] Andrew Radford. *Minimalist syntax: Exploring the structure of English*. Cambridge University Press, 2004.
- [23] Hannes Scheutz. Determinantien und Definitheitsarten im Bairischen und Standarddeutschen. In *Festschrift für Ingo Reiffenstein zum 60. Geburtstag*, pages 231–258. Kümmerle, Göppingen, 1988.
- [24] Bernhard Schwarz and Junko Shimoyama. Negative Islands and obviation by ‘wa’ in Japanese degree questions. In *Proceedings of SALT*, volume 20, pages 702–719, 2011.
- [25] Florian Schwarz. *Two Types of Definites in Natural Language*. PhD thesis, University of Massachusetts Amherst, 2009.
- [26] Mandy Simons, Judith Tonhauser, David Beaver, and Craige Roberts. What projects and why. In *Proceedings of SALT*, volume 20, pages 309–327, 2011.
- [27] Robert Stalnaker. Presuppositions. *Journal of Philosophical Logic*, 2:447457, 1973.
- [28] Kai von Fintel. What is presupposition accommodation, again? *Philosophical Perspectives*, 22(1):137–170, 2008.
- [29] Martina Wiltschko. What does it take to host a (restrictive) relative clause? *Working Papers of the Linguistics Circle of the University of Victoria*, 21(2):100–145, 2012.

Towards a Formal Theory of Explanatory Biases in Discourse

Torgrim Solstad¹ and Oliver Bott²

¹ Department of Language and Literature, Norwegian University of Science and Technology
torgrim.solstad@ntnu.no

² SFB 833, University of Tübingen
oliver.bott@uni-tuebingen.de

Over the past four decades, psycholinguistic studies of discourse structure have revealed interpretation biases for a class of verbs called *implicit causality* (henceforth, IC) verbs. In brief, IC verbs are transitive verbs with two animate arguments which show a preference for anaphorically referring to one of the two arguments when followed by a *because* sentence, cf. (1):

- (1) a. **fascinate, NP1 bias:** *John* fascinated *Mary* because *he* danced beautifully.
b. **congratulate, NP2 bias:** *John* congratulated *Mary* because *she* won the race.

Such preferences are commonly elicited in production experiments where participants are prompted to continue sequences such as “*John congratulated Mary because ...*”. The proportion between continuations mentioning the subject or object argument first is referred to as *IC bias*. Referring to the linear order of arguments, anaphoric reference to the subject argument contributes to NP1 bias, whereas reference to the object argument contributes to NP2 bias.

Although highly consistent across production and comprehension, experimental paradigms and languages, IC bias has attracted little interest in semantic and pragmatic theory. Previous approaches have focused on the establishment of correlations between reference resolution and thematic properties of arguments (e.g. [4, 6]). For instance, it has been shown that the stimulus arguments of psychological verbs are strong bias attractors, cf. (1a). However, little is known as to *why* it is that certain thematic roles attract the bias (and others don’t).

IC verbs are special in other respects, too. In a recent paper, [7] have shown that IC verbs are prone to trigger explanations in subsequent discourse. Participants were prompted to continue sequences such as “*John congratulated Mary.*”, i.e. after a full stop. In such cases, IC verbs triggered 60% explanations, i.e. sentence continuations which were causally related to the prompt, whereas “non-IC verbs” triggered only 25% explanations. However, [7] offered no explanation for this property.

This, however, is a rather interesting property, in particular because it promises to broaden the scope of discourse-theoretic studies. Consider the mini-discourse in (2):

- (2) [John congratulated Mary.]_{s_n} [She won the race.]_{s_{n+1}}

In (2), s_{n+1} is most naturally interpreted as an explanation of s_n , cf. (1b). For such unmarked discourse relations, i.e. without a connective, discourse-theoretic studies have been concerned with the *post hoc* processing of discourse, asking how s_{n+1} is integrated into s_n . From the point of view of IC, however, one would ask already after processing s_n which discourse relation could be *expected* to be established with the next segment. In our study of IC, we investigate lexical triggers of such expectations. We will point at two well-established phenomena from semantic theory involved in such a forward-looking perspective (cf. [8] for experimental evidence).

In this paper, we propose a semantic theory of IC which incorporates the discourse coherence and reference resolution properties of IC verbs, explaining why IC verbs pattern the way

they do. Crucially, our approach links the coreference patterns to the observed preference for explanations [7]. We also present comprehensive experimental evidence in favour of the theory with important implications for future experimental research.

A Semantic Theory of Implicit Causality

IC verbs, we contend, trigger expectations for *specific explanation types*. They do so because they are underspecified with respect to certain properties of the situation described which are (causally) contingent on one of the two participants. Put differently, IC verbs carry an empty “slot” for specific explanatory content. It is this missing information which triggers explanations in full stop continuations and which also triggers primary reference to one of the two participants. IC thus reflects a general processing preference for not leaving missing content unspecified, i.e. a tendency to avoid accommodation [2, 10]. On our analysis, IC bias as a measure of coreference preferences is an epiphenomenon of specific explanatory preferences derived from verb semantics and the particular realization of its arguments.

In order to capture the relation between explanations and coreference, we need to distinguish several types of explanations. Based on [9], we distinguish three main types of explanations, (i) simple causes, (ii) external reasons, and (iii) internal reasons, cf. the examples in (3):

- (3) a. **simple cause:** *John disturbed Mary because he was making lots of noise.*
- b. **external reason:** *John disturbed Mary because she had damaged his bike.*
- c. **internal reason:** *John disturbed Mary because he was very angry at her.*

Simple causes are causes of events or (mental) states. They never involve volition or intention. In (3a), Mary feeling disturbed is understood to be a by-product, as it were, of John’s noise-emitting activity. *External and internal reasons* are causes of attitudinal states. Thus, the *because* clauses in (3b)-(3c) specify causes for John’s intention to disturb Mary. External reasons (3b) are states of affairs external to the attitude-bearer’s mind whereas internal reasons (3c) are attitudes or mental states internal to the attitude-bearer’s mind. The interdependency between explanation type and reference resolution may be seen in (3b)-(3c). The external reason is associated with, and thus makes primary reference to the object (NP2) argument by mentioning it first, whereas the internal reason is associated with the subject (NP1) argument.

We will also assume *backgrounds* as a fourth type of explanation. Backgrounds provide information which make possible, or ‘facilitate’ the situation described by the verb, cf. (4):

- (4) a. *Felix frightened Vanessa because he suddenly screamed.*
- b. *Felix frightened Vanessa because she didn’t hear him coming.*

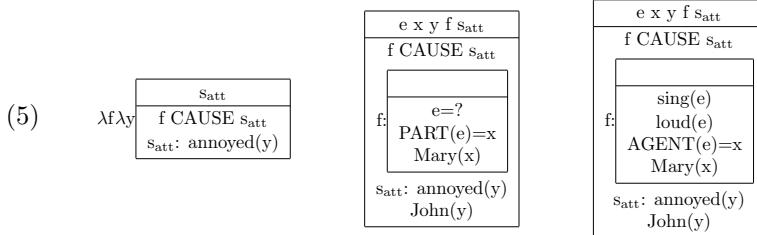
Whereas the *because* clause in (4a) specifies the simple, direct cause of Vanessa’s state of being frightened, the *because* clause in (4b) specifies the background (or: preconditions) for Vanessa being frightened, saying nothing about the actual cause of Vanessa’s fear.

Finally, it is of importance to our approach that *because* clauses introduce entities propositional in nature, cf. [9]. Thus, if *because* clauses are to suitably specify underspecified causal entities, these must likewise be of a propositional semantic type.

Two kinds of underspecified content trigger explanatory expectations. One involves arguments which are underlyingly propositional in nature. Consider the stimulus argument of the stimulus-experiencer verb *annoy*. In the sentence *Mary annoyed John*, *Mary* may be seen as a mere placeholder of a semantic entity more complex in nature. It is actually a specific property or action of Mary’s which is the cause of John being annoyed. Support for this analysis derives from the fact that stimuli in general may be realized as either noun phrases or *that* clauses:

Mary annoyed John/It annoyed John that Mary . . . Stimuli are simple causes, contributing to NP1 bias for stimulus-experiencer verbs, and to NP2 bias with experiencer-stimulus verbs.

Although we cannot go into great detail, we will show by way of example how one can conceive of our “missing content” approach to IC. Consider the DRSs in (5), which illustrate an incremental construction of *Mary annoyed John because she sang loudly*:



The left-most DRS (A.) shows a possible lexical entry for *annoy*. Crucially, the subject argument of *annoy* is associated with a discourse referent *f* which is propositional in nature. When this argument is realized by a DP introducing an entity, this leads to a reinterpretation of the entity argument, much akin to complement coercion (cf. also the mechanism of type-shifting assumed for concealed questions in [1]). A first step in this process is shown in the middle DRS (B.): It is not specified which event involving Mary (it could also be a property of hers) it is that causes annoyance in John (which is what the brief-hand notation *e=?* represents). All we know is that Mary is a participant in this event. Finally, the right-most DRS shows the result of modifying the sequence *Mary annoyed John* by means of the *because* clause *because she sang loudly*.

The other trigger of explanation expectations which we have identified involves presuppositions suitable to give an external reason for the execution of the action denoted by an agent-patient verb (these verbs have previously been assigned to the *ad hoc* class of ‘agent-evocator’ predicates). Take *congratulate*. For it to be used adequately, there must be some occasion on which the agent may congratulate the patient, e.g. an event occurring prior to the act of congratulating. It is exactly this ‘occasioning circumstance’ which is specified by means of the *because* sentence: *John congratulated Mary because she won the race*. Importantly, this is a presupposition which is – for lack of a better term – *cataphorically verifiable*, i.e. after the occurrence of the trigger. Such presuppositions are external reasons, contributing to NP2 bias.

For verbs which lack a relevant underspecified slot, such as agent-patient verbs without an “explanatory” presupposition or causative verbs other than stimulus-experiencer verbs, there is no relevant causal content missing. This basically leaves external and internal reasons as alternative explanations. For some such verbs, a strong bias may still be observed. We cannot yet predict the bias of these verbs on semantic or pragmatic grounds. However, if we are right about the bias resulting from the explanatory strategy sketched above, the bias is predicted to follow from the ratio of external to internal reasons. This is shown in Experiment 2.

Experimental evidence

Testing our semantic theory of IC, we conducted two discourse continuation experiments in German and Norwegian. In Experiment 1a/b, we tested the effects of specifying the bias-triggering missing causal content of psych verbs (Experiment 1a) and presupposition verbs (Experiment 1b) by means of a modifier in the matrix clause. By introducing the modifier, we preempt the strategy of avoiding accommodation which we argue to be behind the observed discourse continuation strategy. If we are right about this, we expect the following effects: (i) a drop in the proportion of explanations of the expected kind, and consequently, (ii) a shift

in IC bias. In Experiment 2, we show that our assumptions are cross-linguistically valid in a large-scale study of German and Norwegian.

Experiment 1a: Modification by means of *durch* ‘by’ Phrases. In this experiment, we specified the expectation-triggering, missing content of stimulus experiencer verbs by means of an adverbial modifier, cf. (6a) (adverbial modifier in italics):

- (6) a. Peter faszinierte Linda (*durch seine Reiseberichte*) a), weil b).
Peter fascinated Linda (*by/with his travel-descriptions*) a) because b) full stop
- b. Maria tötete Johann (*durch einen Schuss*) a), weil b).
Mary killed John (*with a shot*) a) because b) full stop

To control for the general effects of *durch*-modification, we also included causative a-p verbs like *kill* which can also be modified by *durch* phrases, cf. (6b). This is crucial, because the causing entity of causative verbs other than stimulus-experiencer verbs is assumed to be an event. Consequently, *because* clauses cannot specify the causing event for these predicates since they introduce causes propositional in nature:

- (7) #Mary killed John because she stabbed him in the back.

The *because* sentence in (7) cannot be taken to introduce a simple cause as in (6a). Due to this ontological restriction of *because*, the *durch* phrase modification should neither affect the distribution of explanation types nor the bias in the case of causatives such as *kill*.

Methods: 48 native German participants (mean age 23.7 years; range: 19 – 37 years; 33 female) took part in the experiment which employed a $2 \times 2 \times 2 \times 2$ design. We constructed 20 items with s-e verbs and 20 items with causative a-p verbs such as (6a) and (6b). Within each *verb type*, we manipulated the factors *modification* (*durch PP* vs. *no modification*), *connective* (*because* vs. *full stop*) and *gender* ($NP1_{fem.}NP2_{masc.}$ vs. $NP1_{masc.}NP2_{fem.}$). The latter manipulation was included to control for potential effects of gender (see e.g. [5]). The items were presented together with 90 fillers in four lists which were constructed according to a latin square design (only taking into account *verb type*, *modification* and *connective*). The experiment consisted of two blocks. In the first block, participants provided continuations after a full stop (for 20 out of 40 items and 45 fillers). This way, we made sure there was no overt bias towards a particular coherence relation. In a second block, they were prompted to continue the other items and 45 fillers with *weil* ‘because’. Participants who felt that no sensible continuation was possible, had the option to press a “no sensible continuation” button.

Corpus annotation: We removed all nonsensical and ungrammatical continuations from the corpus (3.5% of the data). Continuations after a full stop were annotated with respect to discourse relations using a *because* insertion test to identify explanations (cf. [7]). Explanations after a full stop and the continuations in the *because* conditions were annotated with respect to the causal typology described above; being categorized as *simple causes* (*SC*), *external reasons* (*ER*), *internal reasons* (*IR*) or *backgrounds*. Finally, we annotated the IC bias in all continuations which were explanations. Details on the annotation scheme can be found in [3].

Results and discussion: In the following, we will first report the results for s-e verbs and then present the results obtained for a-p verbs. For statistical analysis we computed log-linear models analyses including factors for participants and items. In the following we report the partial associations. Note that there is no distinction between independent and dependent variables within these models. $LRCS_1$ ($LRCS = \text{log-likelihood ratio } \chi^2$) refers to the analysis by subjects, $LRCS_2$ to the analysis by items.

s-e verbs: Fig. 1a shows the proportions of explanations (for each type) that were produced

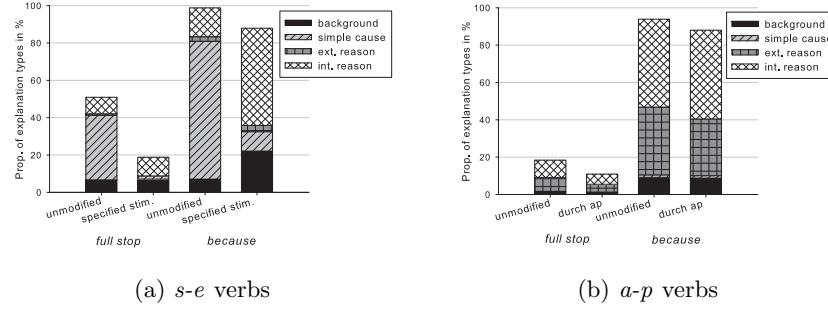


Figure 1: Types of explanations in Experiment 1a.

after full stop and *because* prompts. Trials in the *because* conditions judged not to be continuable in a sensible way were counted as non-explanations. After a full stop, participants produced explanations in 51.3% of the unmodified cases, but only in 19.2% of the modified prompts. A drop in the proportion of explanations was also observed in the *because* conditions where participants provided continuations in 99.2% of the unmodified cases. In the *durch* condition, however, they were unable to produce a continuation in 6.2% of the cases.

The statistical analysis revealed significant partial associations between *modification* and *explanation* ($LRCS_1(1) = 72.9, p < .01$; $LRCS_2(1) = 69.1, p < .01$) and *connective* and *explanation* ($LRCS_1(1) = 525.8, p < .01$; $LRCS_2(1) = 494.4, p < .01$) but no significant interaction between *modification*, *connective* and *explanation* ($LRCS_1(1) = .85$; $LRCS_2(1) = 1.5$).

The drop in explanations was accompanied by a shift in the type of explanation. Explanations in the unmodified conditions were predominantly simple causes (68.6% SC relative to other explanation types after a full stop; 75.0% after *because*). In the modified conditions, however, *internal reasons* were the default explanation type (modified full stop condition: 53.3% IR; modified *because* condition: 59.2%). Statistical analysis revealed that the shift in explanation type was significant (interaction between *modification* and *SC vs. ER/IR/other*: $LRCS_1(1) = 284.3, p < .01$; $LRCS_2(1) = 266.6, p < .01$).

As predicted, the shift in explanation type led to a shift in IC bias. Fig. 2a presents the proportion of continuations that were explanations coreferent with NP1 or NP2 and the explanation profiles of both NP1 and NP2 continuations. Simple causes refer to the stimulus argument leading to a NP1 bias (81.1% NP1) in the unmodified cases. In the modified cases the bias disappears. Continuations are balanced between NP1 and NP2 coreference (52.1% NP1). Statistical analysis showed that the observed shift in bias was reliable (interaction between *modification* and *bias*: $LRCS_1(1) = 85.2, p < .01$; $LRCS_2(1) = 72.2, p < .01$).

a-p verbs: Fig. 1b shows the proportions of explanations that were produced after full stop and *because* prompts. After a full stop, participants produced explanations in 18.8% of the unmodified cases with a slight drop to 11.0% explanations after *durch* phrases. Statistical analysis also revealed significant partial associations between *modification* and *explanation* ($LRCS_1(1) = 8.1, p < .01$; $LRCS_2(1) = 7.1, p < .01$). However, a loglinear model comparing the relative drop in explanations for the verb types revealed a significant interaction between *verb type*, *modification* and *explanation* ($LRCS_1(1) = 6.9, p < .01$) which was due to the fact that the drop in the proportion of explanation was stronger for the *s-e* than the *a-p* verbs.

As predicted, the distribution of explanation types was the same for modified and unmodified *a-p* verbs as indicated by far from significant partial associations (interaction *modification* \times *explanation type* including all four types in the analysis: $LRCS_1(3) = 1.6$; $LRCS_2(3) = .8$).

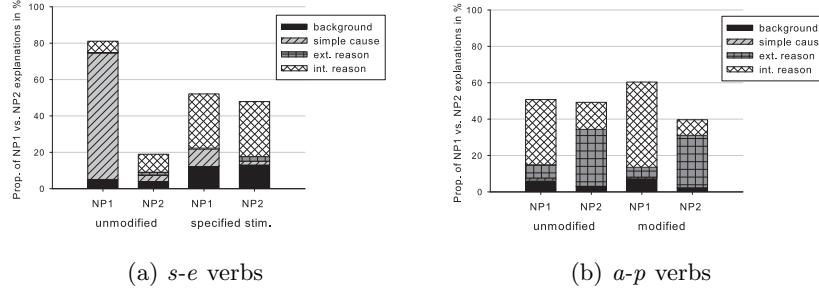


Figure 2: IC bias of explanations in Exp. 1a (aggregated over full stop and *because* conditions).

Fig. 2b presents the proportion of continuations that were explanations coreferent with NP1 or NP2 and the explanation profiles of both kinds. While the unmodified conditions were almost unbiased (50.8% NP1), *durch* modification led to a weak NP1 bias (60.4% NP1). This change was reflected by a significant interaction between *bias* and *modification* in loglinear analyses ($LRCS_1(1) = 4.8, p < .05$; $LRCS_2(1) = 3.9, p < .05$). Crucially, this shift is in the opposite direction of the *s-e* verbs. An explanation may be that in the modified conditions the salience of NP1 might have been higher than in the unmodified conditions since the simple cause provided by the *durch* phrase made implicit reference to the subject. For instance, in sentence (6b) *durch einen Schuss* may strengthen the activation of the subject because it was the agent of the shooting event. The opposite behavior of *s-e* as compared to *a-p* verbs with respect to IC bias was statistically confirmed by a significant interaction between *verb type*, *bias* and *modification* in a loglinear analysis comparing both verb types ($LRCS_1(1) = 73.1, p < .01$).

Experiment 1b: Implicit Arguments of PSP verbs. Verbs that carry a presupposition (PSP) such as *congratulate* constitute the second class verbs discussed above. The verification of the presupposition can take various forms. The presupposed external reason can be stated in a *because* clause, but it can also be specified by a prepositional object, as in (8). The experiment shows that providing an external reason in the matrix sentence shifts the bias from NP2 to NP1.

- (8) Peter dankte Maria (*für die finanzielle Unterstützung*) a), weil b).
 Peter thanked Mary (*for the financial support*) a) because b) full stop.

Methods: 20 items with PSP verbs were constructed in eight conditions manipulating the factors *modification* (*implicit argument* vs. *no modification*), *connective* (*because* vs. *full stop*) and *gender* ($NP1_{fem}, NP2_{masc}$. vs. $NP1_{masc}, NP2_{fem}$.). The methods were as in Exp. 1a.

Results and discussion: Fig. 3a presents the proportions of explanation types. Again, specifying the missing causal content led to a significant drop in the proportion of explanations from 55.8% to 30.8% after a full stop and from 94.6% to 91.7% after a *because* prompt (interaction of *modification* and *explanation*: $LRCS_1(1) = 34.3, p < .01$; $LRCS_2(1) = 31.9, p < .01$).

The types of explanations were as expected. In the unmodified conditions, external reasons were the default (unmodified full stop condition: 70.1% of all explanations were ER; unmodified *because* condition: 86.3% ER). After matrix sentences with a specified presupposition, the proportion of ERs dropped and IRs became the predominant type of explanation (unmodified full stop condition: 67.6% IR; unmodified *because* condition: 58.6% IR). Statistical analyses revealed that this shift in explanation type was significant (interaction between *modification* and *ER* vs. *SC/IR/other*: $LRCS_1(1) = 193.6, p < .01$; $LRCS_2(1) = 202.1, p < .01$).

Fig. 3b shows that this led to a reversal in IC bias. Explanations in the unmodified conditions

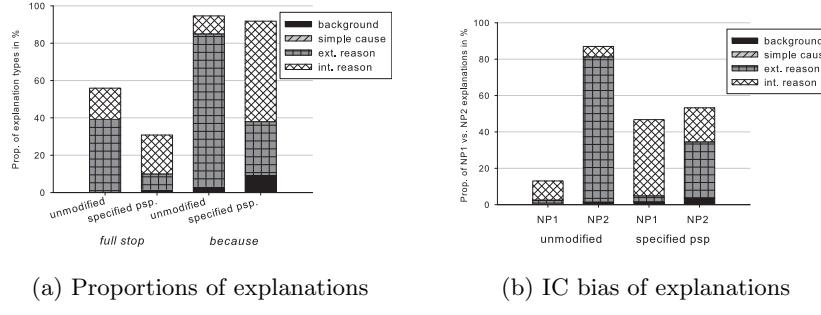


Figure 3: Explanation types and IC bias in Experiment 1b.

displayed a clear NP2 bias (full stop: 84.4% NP2; because: 88.4% NP2). By contrast, conditions that specified the PSP in the prompt, revealed no clear bias (full stop: 44.4% NP2; because: 50.8% NP2) – a highly significant shift (interaction *modification* \times *bias*: $LRCS_1(1) = 93.5, p < .01$; $LRCS_2(1) = 93.4, p < .01$). The analysis of the NP1 and NP2 explanation profiles shows that external reasons are associated with NP2 reference, whereas internal reasons are associated with NP1 reference. Therefore, IC bias can be directly read off the explanation profile lending support to our claim that it is an epiphenomenon of the explanatory strategy.

Experiments 1a/b provide evidence for our claim that specifying missing causal content minimizes the need for explanation, which in turn gives rise to a shift to another discourse relation. We have seen that the discourse processor tends to elaborate on this content probably in order to avoid accommodation costs [2]. Furthermore, the triggered explanations were of the type expected on our analysis. Considering argument associations provided us with a direct way to account for the observed bias distributions. Thus, IC bias seems to be an epiphenomenon of developing a discourse expectation on the basis of presupposed causal content.

Experiment 2 – Large-scale corpus. In this continuation experiment, we elicited 10.100 written productions for 101 near-synonymous German and Norwegian verbs (for details see [3]). 52 German and 48 Norwegian participants were prompted to continue sentences after either a full stop or *because*. Besides 16 stimulus-experiencer and 10 presupposition verbs the study included 17 experiencer-stimulus verbs and 43 agent-patient verbs without a presupposition as well as 14 ambiguous agent-patient/stimulus-experiencer verbs such as *disturb* and *hurt*.

A verb-by-verb comparison revealed that the biases were highly correlated between the two languages ($r = .92$), showing that the phenomenon is cross-linguistically stable as expected under a semantic account. Moreover, our analysis of stimulus arguments generalized from stimulus-experiencer to experiencer-stimulus verbs. Again, we observed explanations to be the default coherence relation. While agent-patient verbs without a presupposition had a mean proportion of only 39.1% explanations, experiencer-stimulus had 63.2% explanations and patterned with the stimulus-experiencer (67.6%) and presupposition verbs (61.5%). Fully consistent with our model, the explanations of experiencer-stimulus verbs were overwhelmingly of the *simple cause* type (87.9%) and led to a strong NP2 bias (14.7% NP1). Finally, even for verbs without underspecified content, i.e. agent-patient verbs without a presupposition, we were able to account for most of the variance in IC bias by correlating it with the ratio of external and internal reasons (linear regression: corrected $R^2 = .75$). Fig. 4 shows that the ratio of internal and external reasons is a good predictor of IC bias. Also in line with our predictions, the figure shows a gradient transition from 100% external reasons to 96% internal reasons which offers a straightforward explanation for the observed variability of IC bias in this class of verbs.

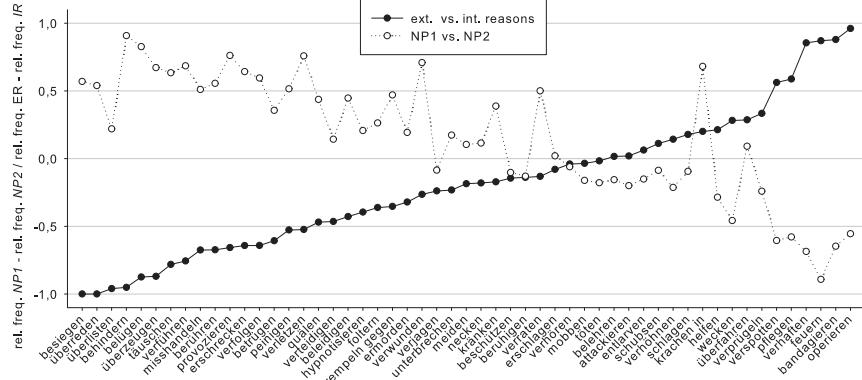


Figure 4: Unambiguous AP verbs in Exp. 2. IC bias (white dots) is shown as a function of the ratio of internal (IR) and external reasons (ER) (black dots). Bias: +1.0 corresponds to 100% NP1 and -1.0 to 100% N2 bias. Reasons: -1.0 corresponds to 100% IR, +1.0 to 100% ER.

Conclusions

The results show that IC bias strongly depends on the availability of specific explanation types and that it can be manipulated by specifying implicit explanations in the prompt. The proposed analysis offers a formally precise, predictive model that can account for a large body of processing data (off- and online) that have accumulated over the last four decades, having been largely neglected in semantic/pragmatic theory. At the same time, it yields fine-grained, novel predictions as to the incremental nature of explanatory discourse which are highly relevant for psycholinguists interested in the underlying forces of establishing discourse coherence.

References

- [1] M. Aloni and F. Roelofsen. Interpreting concealed questions. *L&P*, 34:443–478, 2011.
- [2] G. Altmann and M. Steedman. Interaction with context during human sentence processing. *Cognition*, 30:191–238, 1988.
- [3] O. Bott and T. Solstad. From verbs to discourse: a novel account of implicit causality. In B. Hemforth, B. Schmiedtová, and C. Fabricius-Hansen, editors, *Psycholinguistic approaches to meaning and understanding across language*. Springer, to appear.
- [4] R. Brown and D. Fish. The psychological causality implicit in language. *Cognition*, 14:237–273, 1983.
- [5] E. C. Ferstl, A. Garnham, and C. Manouilidou. Implicit causality bias in English: A corpus of 300 verbs. *Behav Res Methods*, 43:124–135, 2011.
- [6] J. K. Hartshorne and J. Snedeker. Verb argument structure predicts implicit causality: The advantages of finer-grained semantics. *Language and Cognitive Processes*, 28:1474–1508, 2013.
- [7] A. Kehler, L. Kertz, H. Rohde, and J. L. Elman. Coherence and coreference revisited. *JoS*, 25:1–44, 2008.
- [8] P. Pyykkönen and J. Järvikivi. Activation and persistence of implicit causality information in spoken language comprehension. *Exp Psychol*, 57:5–16, 2010.
- [9] T. Solstad. Some new observations on ‘because (of)’. In M. Aloni, H. Bastiaanse, T. de Jager, and K. Schulz, editors, *Amsterdam Colloquium 2009*, LNCS. Springer, 2010.
- [10] R. A. van der Sandt. Presupposition projection as anaphora resolution. *JoS*, 9:333–377, 1992.

Epistemic Modals, Qualitative Probability, and Nonstandard Probability

Satoru Suzuki

Faculty of Arts and Sciences, Komazawa University
1-23-1, Komazawa, Setagaya-ku, Tokyo 154-8525, Japan
bxs05253@nifty.com

Abstract

Yalcin [23] shows that Kratzer's model [5] does not validate some intuitively valid inferences and validates some intuitively invalid ones. He adopts a model based directly on a probability measure. However, as Kratzer [6] says, 'Our semantic knowledge alone does not give us the precise quantitative notions of probability and desirability that mathematicians and scientists work with', Yalcin's model seems to be unnatural as a model for comparative epistemic modals. The aim of this paper is to propose a new version of complete logic—modal-qualitative-probability logic (MQPL)—the model of the language of which has the following four merits: (i) The model reflects Kratzer's intuition above in the sense that the model is not based directly on a probability measure, but based on a qualitative probability ordering. (ii) The model does not cause Yalcin's problem. (iii) The model has no limitation of the size of the domain. (iv) The model can deal with the two-dimensional geometric probability that Kolmogorov probability theory cannot.

1 Motivation

Kratzer [5] provides comparative epistemic modals such as 'at least as likely as' with their models in terms of a qualitative ordering on propositions derived from a qualitative ordering on possible worlds. Yalcin [23] shows that Kratzer's model does not validate some intuitively valid inference schemata and validates some intuitively invalid ones. He adopts a model based directly on a probability measure for comparative epistemic modals. His model does not cause this problem. However, as Kratzer [6] says, 'Our semantic knowledge alone does not give us the precise quantitative notions of probability and desirability that mathematicians and scientists work with', Yalcin's model seems to be unnatural as a model for comparative epistemic modals. Holliday and Icard [3] prove that not only a probability measure model but also a qualitatively additive measure model and a revised version of Kratzer's model do not cause Yalcin's problem.

The aim of this paper is to propose a new version of complete logic—modal-qualitative-probability logic (MQPL)—the model of the language of which has the following four merits:

1. The model reflects *Kratzer's intuition* above in the sense that the model is not based directly on a probability measure, but based on a *qualitative probability ordering*.
2. The model does not cause Yalcin's problem.
3. The model has *no limitation of the size* of the domain.
4. The model can deal with the *two-dimensional geometric probability*, for example, of picking a point from the diagonals D_1 and D_2 of a rectangle, given that the point is on one of D_1 and D_2 , equals $\frac{1}{2}$ under *nonstandard infinitesimal* probability theory and equals $\frac{0}{0}$ (*undefined*) under *Kolmogorov* probability theory. In other words, we can provide the

following sentence with its truth condition in terms of the model of the language of MQPL, but cannot provide in terms of Kolmogorov probability measure model:

The point being on D_1 is as likely as it being on D_2 .

So the model of the language of MQPL has a wider scope than Kolmogorov probability measure model.

The structure of this paper is as follows. In Section 2, we argue on some relations between qualitative probability and a standard probability measure and a relation between qualitative probability and a nonstandard probability measure. In Section 3, we define the language $\mathcal{L}_{\text{MQPL}}$ of MQPL, define a model \mathfrak{M} of $\mathcal{L}_{\text{MQPL}}$, provide MQPL with a truth definition and a validity definition, show that MQPL justifies the (in)validity of Yalcin's formulae, provide MQPL with its proof system, and touch upon the soundness and completeness theorems of MQPL. In Section 4, we compare MQPL with some logics of Holliday and Icard [3]. In Section 5, we finish with brief concluding remarks.

2 Qualitative Probability and Nonstandard Probability

When W is a nonempty set of possible worlds, \mathcal{F} a Boolean algebra of subsets of W , and \succsim a qualitative probability ordering on \mathcal{F} , de Finetti [1] specifies necessary conditions on $\langle W, \mathcal{F}, \succsim \rangle$ for the existence of a probability measure. Kraft, Pratt, and Seidenberg [4] shows that de Finetti's conditions on $\langle W, \mathcal{F}, \succsim \rangle$ are not sufficient and presents necessary and sufficient conditions on $\langle W, \mathcal{F}, \succsim \rangle$ for the existence of a probability measure in the case that W is *finite*. Scott [8] presents much the same conditions in a general setting as follows:

Definition 1 (Nontriviality, Nonnegativity, and Scottness (S)).

- **Nontriviality:** $W \succ \emptyset$.
- **Nonnegativity:** $A \succsim \emptyset$ for any $A \in \mathcal{F}$.
- **Connectedness:** $A \succsim B$ or $B \succsim A$ for any $A, B \in \mathcal{F}$.
- **Scottness (S):** For any $A_1, \dots, A_n, B_1, \dots, B_n \in \mathcal{F}$, if, for any $i < n$, $(A_i \succsim B_i)$, then $B_n \succsim A_n$, given that

$$\sum_{i=1}^n \chi_{A_i} = \sum_{i=1}^n \chi_{B_i}$$

holds, where χ_A is a characteristic function of $A \in \mathcal{F}$. χ_A is a function from W to $\{0, 1\}$ such that

$$\chi_A(w) := \begin{cases} 1 & \text{if } w \in A, \\ 0 & \text{if } w \notin A. \end{cases}$$

Remark 1. Intuitively, **Scottness (S)** says that whenever, for any $w \in W$, w is in **exactly as many** A_i 's as B_i 's, if, for any $i < n$, $(A_i \succsim B_i)$, then $B_n \succsim A_n$.

Narens [7] shows that the same conditions as Scott are necessary and sufficient for the existence of a *nonstandard* probability measure *without the limitation of the size of W* . However, **Scottness (S)** is unpleasant because it is stated in terms of *characteristic functions* rather than in terms of union or other primitive notions. Domotor [2] states the condition in terms of intersection, union, and \succsim as follows:

Definition 2 (Scottness (D)). *Scottness (D)*: For any $A_1, \dots, A_n, B_1, \dots, B_n \in \mathcal{F}$, if for any $i < n$, $(A_i \succsim B_i)$, then $B_n \succsim A_n$, given that

$$\bigcup_{1 \leq i_1 < \dots < i_k \leq n} (A_{i_1} \cap \dots \cap A_{i_k}) = \bigcup_{1 \leq i_1 < \dots < i_k \leq n} (B_{i_1} \cap \dots \cap B_{i_k})$$

holds for any k with $1 \leq k \leq n$.

When ${}^*\mathbb{R}$ denotes the set of *nonstandard reals* (containing *infinitesimals*), we can state *Scott-Narens theorem* by using *Domotor's notation* as follows:

Theorem 1 (Representation for Qualitative Probability Ordering). *For any $\langle W, \mathcal{F}, \succsim \rangle$, there exists a finitely additive **nonstandard**-real-valued probability measure $P : \mathcal{F} \rightarrow {}^*\mathbb{R}$ satisfying*

$$A \succsim B \quad \text{iff} \quad P(A) \geq P(B)$$

iff *Nontriviality*, *Nonnegativity*, *Connectedness*, and *Scottness (D)* are met.

3 Modal-Qualitative-Probability Logic (MQPL)

3.1 Language

We define the language $\mathcal{L}_{\text{MQPL}}$ of MQPL, which is the *same language as Holliday and Icard [3]*, as follows:

Definition 3 (Language). Let \mathcal{S} denote a set of sentential variables, \diamond a unary sentential operator, and \geqslant a binary sentential operator. The language $\mathcal{L}_{\text{MQPL}}$ of MQPL is given by the following BNF grammar:

$$\varphi ::= s \mid \top \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \diamond\varphi \mid (\varphi \geqslant \varphi)$$

such that $s \in \mathcal{S}$.

- $\perp, \vee, \rightarrow, \leftrightarrow$ and \square are introduced by the standard definitions.
- $\diamond\varphi$ is interpreted to mean that **possibly** φ .
- $\varphi \geqslant \psi$ is interpreted to mean that φ is **at least likely as** ψ .
- $\varphi > \psi := (\varphi \geqslant \psi) \wedge \neg(\psi \geqslant \varphi)$.
- $\varphi > \psi$ is interpreted to mean that φ is **more likely than** ψ .
- $\Delta\varphi := \varphi > \neg\varphi$.
- $\Delta\varphi$ is interpreted to mean that **probably** φ
- The set of all well-formed formulae of $\mathcal{L}_{\text{MQPL}}$ will be denoted by $\Phi_{\mathcal{L}_{\text{MQPL}}}$.

3.2 Semantics

We define a structured model \mathfrak{M} of $\mathcal{L}_{\text{MQPL}}$ as follows:

Definition 4 (Model). \mathfrak{M} is a quadruple $\langle W, R, V, \rho \rangle$ in which

- W is a non-empty set of possible worlds,
- R is a binary accessibility relation on W ,
- V is a truth assignment to each $s \in S$ for each $w \in W$, and
- ρ is a qualitative probability space assignment that assigns to each $w \in W$ a qualitative probability space $\langle W_w, \mathcal{F}_w, \succsim_w \rangle$ in which
 - $W_w := \{w' \in W : R(w, w')\}$,
 - \mathcal{F}_w is a Boolean algebra of subsets of W_w with \emptyset as zero element and W_w as unit element, and
 - \succsim_w is a qualitative probability ordering on \mathcal{F}_w that satisfies all of **Nontriviality**, **Nonnegativity**, **Connectedness**, and **Scottness (D)** of Theorem 1.

We provide MQPL with the following truth definition at $w \in W$ in \mathfrak{M} , define the truth in \mathfrak{M} , and then define validity as follows:

Definition 5 (Truth and Validity). The notion of $\varphi \in \Phi_{\mathcal{L}_{\text{MQPL}}}$ being true at $w \in W$ in \mathfrak{M} , in symbols $(\mathfrak{M}, w) \models_{\text{MQPL}} \varphi$, is inductively defined as follows:

- $(\mathfrak{M}, w) \models_{\text{MQPL}} s$ iff $V(w)(s) = \text{true}$.
- $(\mathfrak{M}, w) \models_{\text{MQPL}} \top$.
- $(\mathfrak{M}, w) \models_{\text{MQPL}} \neg\varphi$ iff $(\mathfrak{M}, w) \not\models_{\text{MQPL}} \varphi$.
- $(\mathfrak{M}, w) \models_{\text{MQPL}} \varphi \wedge \psi$ iff $(\mathfrak{M}, w) \models_{\text{MQPL}} \varphi$ and $(\mathfrak{M}, w) \models_{\text{MQPL}} \psi$.
- $(\mathfrak{M}, w) \models_{\text{MQPL}} \Diamond\varphi$ iff, for some w' such that $R(w, w')$, $(\mathfrak{M}, w') \models_{\text{MQPL}} \varphi$.
- $(\mathfrak{M}, w) \models_{\text{MQPL}} \varphi \geqslant \psi$ iff $\llbracket \varphi \rrbracket_w^{\mathfrak{M}} \succsim_w \llbracket \psi \rrbracket_w^{\mathfrak{M}}$, where $\llbracket \varphi \rrbracket_w^{\mathfrak{M}} := \{w' \in W : R(w, w') \text{ and } (\mathfrak{M}, w') \models_{\text{MQPL}} \varphi\}$.

If $(\mathfrak{M}, w) \models_{\text{MQPL}} \varphi$ for all $w \in W$, we write $\mathfrak{M} \models_{\text{MQPL}} \varphi$ and say that φ is true in \mathfrak{M} . If φ is true in all models of $\mathcal{L}_{\text{MQPL}}$, we write $\models_{\text{MQPL}} \varphi$ and say that φ is valid.

The next corollary follows from Theorem 1 and Definition 5.

Corollary 1 (Truth Condition by Probability Measure). There exists a finitely additive nonstandard-real-valued probability measure $P : \mathcal{F} \rightarrow^* \mathbb{R}$ satisfying

$$(\mathfrak{M}, w) \models_{\text{MQPL}} \varphi \geqslant \psi \text{ iff } P(\llbracket \varphi \rrbracket_w^{\mathfrak{M}}) \geq P(\llbracket \psi \rrbracket_w^{\mathfrak{M}}).$$

iff **Nontriviality**, **Nonnegativity**, **Connectedness**, and **Scottness (D)** are met.

Yalcin (2010) presents the following list of intuitively valid formulae (V1)–(V11) and intuitively invalid formulae (I1) and (I2):

- (V1) $\Delta\varphi \rightarrow \neg\Delta\neg\varphi$,

- (V2) $\Delta(\varphi \wedge \psi) \rightarrow (\Delta\varphi \wedge \Delta\psi)$,
- (V3) $\Delta\varphi \rightarrow \Delta(\varphi \vee \psi)$,
- (V4) $\varphi \geq \perp$,
- (V5) $\top \geq \varphi$,
- (V6) $\Box\varphi \rightarrow \Delta\varphi$,
- (V7) $\Delta\varphi \rightarrow \Diamond\varphi$,
- (V8) $(\varphi \rightarrow \psi) \rightarrow (\Delta\varphi \rightarrow \Delta\psi)$,
- (V9) $(\varphi \rightarrow \psi) \rightarrow (\neg\Delta\psi \rightarrow \neg\Delta\varphi)$,
- (V10) $(\varphi \rightarrow \psi) \rightarrow (\psi \geq \varphi)$,
- (V11) $(\psi \geq \varphi) \rightarrow (\Delta\varphi \rightarrow \Delta\psi)$,
- (V12) $(\psi \geq \varphi) \rightarrow ((\varphi \geq \neg\varphi) \rightarrow (\psi \geq \neg\psi))$,
- (I1) $((\varphi \geq \psi) \wedge (\varphi \geq \chi)) \rightarrow (\varphi \geq (\psi \vee \chi))$, and
- (I2) $((\varphi \geq \neg\varphi) \wedge (\neg\varphi \geq \varphi)) \rightarrow (\varphi \geq \psi)$.

MQPL justifies the (in)validity of Yalcin's formulae as follows:

Proposition 1 (Justification of Yalcin's Formulae). MQPL validates all of (V1)–(V12) and validate neither (I1) nor (I2).

3.3 Syntax

The proof system of MQPL consists of the following:

Definition 6 (Proof System).

- All tautologies of classical sentential logic,
- $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$ (*K*),
- $(\Box(\varphi_1 \leftrightarrow \varphi_2) \wedge \Box(\psi_1 \leftrightarrow \psi_2)) \rightarrow ((\varphi_1 \geq \psi_1) \leftrightarrow (\varphi_2 \geq \psi_2))$
(*Replacement of Necessary Equivalents*),
- $\top > \perp$ (*Syntactic Counterpart of Nontriviality*),
- $\varphi \geq \top$ (*Syntactic Counterpart of Nonnegativity*),
- $(\varphi \geq \psi) \vee (\psi \geq \varphi)$ (*Syntactic Counterpart of Connectedness*),
- $$\begin{aligned} & \left(\bigvee_{\substack{1 \leq i_1 < \dots < i_k \leq n}} (\varphi_{i_1} \wedge \dots \wedge \varphi_{i_k}) \leftrightarrow \bigvee_{\substack{1 \leq i_1 < \dots < i_k \leq n}} (\psi_{i_1} \wedge \dots \wedge \psi_{i_k}) \right) \\ & \rightarrow \left(\bigwedge_{i=1}^{n-1} (\varphi_i \geq \psi_i) \rightarrow (\psi_n \geq \varphi_n) \right) \end{aligned}$$

(*Syntactic Counterpart of Scottness (D)*),

- *Modus Ponens*, and
- *Necessitation*.

A proof of $\varphi \in \Phi_{\mathcal{L}_{MQPL}}$ is a finite sequence of \mathcal{L}_{MQPL} -formulae having φ as the last formula such that either each formula is an instance of an axiom or it can be obtained from formulae that appear earlier in the sequence by applying an inference rule. If there is a proof of φ , we write $\vdash_{MQPL} \varphi$.

3.4 Metalogic

We can prove the soundness and completeness of MQPL.

Theorem 2 (Soundness). *For any $\varphi \in \Phi_{\mathcal{L}_{MQPL}}$, if $\vdash_{MQPL} \varphi$, then $\models_{\mathcal{L}_{MQPL}} \varphi$.*

Theorem 3 (Completeness). *For any $\varphi \in \Phi_{\mathcal{L}_{MQPL}}$, if $\models_{\mathcal{L}_{MQPL}} \varphi$, then $\vdash_{MQPL} \varphi$.*

4 Comparison with Holliday and Icard [3]

In this section, we would like to compare MQPL with some logics of Holliday and Icard [3]. We have adopted the same language \mathcal{L}_{MQPL} as \mathcal{L} of [3]. Holliday and Icard consider three kinds of models for \mathcal{L} :

1. measure model,
 - (a) finitely additive measure model (finitely additive probability measure model),
 - (b) qualitatively additive measure model,
2. event-ordering model,
3. world-ordering model
 - (a) Kratzer's world-ordering model,
 - (b) revised version of Kratzer's world-ordering model.

Holliday and Icard show that a complete logic **WJR** based on Kratzer's world-ordering model validate all of (V1)–(V10), (V12), (I1), and (I2), but does not validate (V11), and that a complete logic **FP** $_\infty$ based on a finitely additive measure model, a complete logic **FA** based on a qualitatively additive measure model, and a logic **WP** $_\infty$ **R** based on a revised version of Kratzer's world-ordering model each validate all of (V1)–(V12) and validate neither (I1) nor (I2). On the other hand, we have proposed a complete logic MQPL based on a kind of *event-ordering* model—*qualitative probability ordering* model. This model is a compromise between an event-ordering model and a finitely additive measure model in the sense that by *Theorem 1* (Representation for Qualitative Probability Ordering), the event-ordering model (qualitative probability ordering model) can be connected to the finitely additive measure model (finitely additive nonstandard-real-valued probability measure model) in *Corollary 1* (Truth Condition by Probability Measure). The qualitative probability ordering side of the model \mathfrak{M} of \mathcal{L}_{MQPL} can reflect Kratzer's intuition above in the sense that the model is not based directly on a probability measure, but based on a qualitative probability ordering. The finitely additive nonstandard-real-valued probability measure side of \mathfrak{M} enables MQPL to avoid Yalcin's problem.

5 Concluding Remarks

In this paper, we have proposed a new version of complete logic—modal-qualitative-probability logic (MQPL)—the model of the language of which has the four merits listed in Section 1.

This paper is only a part of a larger *measurement-theoretic* study. By means of measurement theory, we constructed or are trying to construct such logics as

1. (dynamic epistemic) preference logic [9, 11],
2. dyadic deontic logic [10],
3. vague predicate logic [14, 15]
4. threshold-utility-maximiser's preference logic [12, 13],
5. interadjective-comparison logic [18],
6. gradable-predicate logic [17],
7. logic for better questions and answers [16],
8. doxastic and epistemic logic [22],
9. multidimensional-predicate-comparison logic [19],
10. logic for preference aggregation represented by a Nash collective utility function [20], and
11. preference aggregation logic for weighted utilitarianism [21].

Acknowledgements This work was supported by Grant-in-Aid for Scientific Research (C) (23520015). The author would like to thank two anonymous reviewers of AC 2013 for their helpful comments.

References

- [1] B. de Finetti. La prévision: Ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré*, 7:1–68, 1937.
- [2] Z. Domotor. Qualitative probabilities revisited. In P. Humphreys, editor, *Patrick Suppes: Scientific Philosopher, Vol.1. Probability and Probabilistic Causality*, pages 197–238. Kluwer, Dordrecht, 1994.
- [3] W. H. Holliday and T. F. Icard, III. Measure semantics and qualitative semantics for epistemic modals. In *Proceedings of SALT 23*, 2013. to appear.
- [4] C. H. Kraft et al. Intuitive probability on finite sets. *Annals of Mathematical Statistics*, 30:408–419, 1959.
- [5] A. Kratzer. Modality. In A. von Stechow and D. Wunderlich, editors, *Semantics: An International Handbook of Contemporary Research*, pages 639–50. de Gruyter, Berlin, 1991.
- [6] A. Kratzer. *Modals and Conditionals*. Oxford University Press, Oxford, 2012.
- [7] L. Narens. Minimal conditions for additive conjoint measurement and qualitative probability. *Journal of Mathematical Psychology*, 11:404–430, 1974.
- [8] D. Scott. Measurement structures and linear inequalities. *Journal of Mathematical Psychology*, 1:233–247, 1964.

- [9] S. Suzuki. Preference logic and its measurement-theoretic semantics. In *Electronic Proceedings of the 8th Conference on Logic and the Foundations of Game and Decision Theory (LOFT 2008)*, 2008. Paper No. 42.
- [10] S. Suzuki. Measurement-theoretic foundation of preference-based dyadic deontic logic. In X. He et al., editors, *Proceedings of the Second International Workshop on Logic, Rationality, and Interaction (LORI-II)*, volume 5834 of *LNCS*, pages 278–291. Springer-Verlag, Heidelberg, 2009.
- [11] S. Suzuki. Prolegomena to dynamic epistemic preference logic. In H. Hattori et al., editors, *New Frontiers in Artificial Intelligence*, volume 5447 of *LNCS*, pages 177–192. Springer-Verlag, Heidelberg, 2009.
- [12] S. Suzuki. Prolegomena to threshold utility maximiser’s preference logic. In *Electronic Proceedings of the 9th Conference on Logic and the Foundations of Game and Decision Theory (LOFT 2010)*, 2010. Paper No. 44.
- [13] S. Suzuki. A measurement-theoretic foundation of threshold utility maximiser’s preference logic. *Journal of Applied Ethics and Philosophy*, 3:17–25, 2011.
- [14] S. Suzuki. Measurement-theoretic foundations of probabilistic model of jnd-based vague predicate logic. In H. van Ditmarsch et al., editors, *Proceedings of the Third International Workshop on Logic, Rationality, and Interaction (LORI-III)*, volume 6953 of *LNCS*, pages 272–285. Springer-Verlag, Heidelberg, 2011.
- [15] S. Suzuki. Prolegomena to salient-similarity-based vague predicate logic. In T. Onada et al., editors, *New Frontiers in Artificial Intelligence*, volume 6797 of *LNCS*, pages 75–89. Springer-Verlag, Heidelberg, 2011.
- [16] S. Suzuki. Measurement-, information-, and decision-theoretic foundations of logic for better questions and answers. In *Proceedings of the 10th Conference on Logic and the Foundations of Game and Decision Theory (LOFT 2012)*, pages 504–514, 2012.
- [17] S. Suzuki. Measurement-theoretic foundations of gradable-predicate logic. In M. Okumura et al., editors, *New Frontiers in Artificial Intelligence*, volume 7258 of *LNCS*, pages 82–95. Springer-Verlag, Heidelberg, 2012.
- [18] S. Suzuki. Measurement-theoretic foundations of interadjective-comparison logic. In A. Aguilar-Guevara et al., editors, *Proceedings of Sinn und Bedeutung 16*, volume 2, pages 571–584, Cambridge, Mass., 2012. MIT Working Papers in Linguistics.
- [19] S. Suzuki. Measurement-theoretic bases of multidimensional-predicate logic, 2013. manuscript.
- [20] S. Suzuki. Measurement-theoretic foundations of many-sorted preference aggregation logic for nash collective utility function, 2013. manuscript.
- [21] S. Suzuki. Measurement-theoretic foundations of preference aggregation logic for weighted utilitarianism. In T. Yamada, editor, *Electronic Post-Proceedings of the Third International Workshop on Philosophy and Ethics of Social Reality (SOCREAL 2013)*, 2013. to appear.
- [22] S. Suzuki. Remarks on decision-theoretic foundations of doxastic and epistemic logic (revised version). *Studies in Logic*, 6:1–12, 2013.
- [23] S. Yalcin. Probability operators. *Philosophy Compass*, 5:916–937, 2010.

Monotonicity has only a relative effect on the complexity of quantifier verification

Jakub Szymanik* and Marcin Zajenkowski†

¹ Institute for Logic, Language and Computation

University of Amsterdam

J.K.Szymanik@uva.nl

² Faculty of Psychology, University of Warsaw

Zajenkowski@psych.uw.edu.pl

Abstract

We discuss a computational model of quantifier verification. It predicts that there is *no effect of monotonicity* on the verification of numerical quantifiers but *only the interaction of monotonicity and sentential truth-values*. Moreover, it predicts no monotonicity or interaction with truth-values effects for proportional quantifiers. We present an experimental study supporting the predictions of the computational model. We argue that the role of the interaction between monotonicity and sentential truth-values as well as the differences between various quantifier classes (numerical vs. proportional) have been regularly overlooked in the literature.

1 Introduction

Monotonicity is considered to be one of the key properties of languages both in logic and linguistics. Barwise and Cooper [1] even suggested that monotonicity is one of the *semantic universals*: the simplest noun phrases of any natural language express monotone quantifiers or conjunctions of monotone quantifiers. They have also noted that monotonicity relates to the intuitive truth-value checking procedures for quantified sentences. For example, imagine a parking lot filled with cars. To verify an upward monotone (increasing) sentence ‘More than seven cars are green’, you need to find at least eight green cars (a so-called witness set in Barwise’s and Cooper’s terminology). For a downward monotone (decreasing) sentence, e.g., ‘Fewer than eight cars are red’ you must check all the cars and make sure that there are no more than seven green cars. Based on the intuitive complexity of these procedures, Barwise and Cooper predicted that ‘*response latencies for verification tasks involving decreasing quantifiers would be somewhat greater than for increasing quantifiers*’ (*p. 192*).

It seems that [1] has overlooked the truth-value of the sentence as an important aspect of the verification complexity. For instance, if the upward monotone sentence is true, then indeed one needs

*Acknowledges the support of Veni grant NWO-639.021.232

†Acknowledges the support of Polish National Science Centre grant no. 2011/01/D/HS6/01920.

to just find any witness set; otherwise, when the sentence is false, one needs to check all the cars and make sure that really not more than 7 of them are green. If you can perceptually quickly identify the set satisfying the predicate in question, e.g., the set of all green cars, then for true upward monotone quantifier it will take more counting than for the corresponding false quantifier to judge whether the set is ‘large enough’. However, for the downward monotone sentence the situation is exactly the opposite. Moreover, intuitively, false (resp. true) instances of an upward monotone sentence are equally hard as true (resp. false) instances of an downward monotone sentence. Therefore, it seems that we should rather expect *the effect of interaction between monotonicity and truth-value* than pure monotonicity to have an impact on the verification complexity.

Thinking about quantifier verification in terms of computations [2, 8] can help us to clarify the intuitions about the influence of monotonicity on verification, and its interaction with truth-values:

Sentences with upward monotone numerical quantifiers, e.g., ‘more than seven’, should take longer to process when they are true than when they are false. For true ‘more than seven’ subjects need to count up to eight while for the false sentences only up to seven. In case of the sentences containing downward monotone numerical quantifiers, like ‘fewer than eight’, the relationship is reversed. This hypothesis directly improves on [1] by taking into account the interaction between monotonicity and truth-value.

Proportional quantifiers should be more difficult, as the minimal verification procedure triggered by them is cognitively more complex [6, 7, 9–13]. It should be the case even in our experimental setting where ‘more than half’ (‘less than half’) and ‘more than seven’ (‘less than eight’) are denotationally equivalent. The reason is that the corresponding procedure is triggered automatically rather by the linguistic form than the situation to be judged. Moreover, in the case of proportional quantifiers, subjects need to always compare all elements, no matter whether the sentences are true or false, and so there should be no significant difference in the processing difficulty between the upward and the downward monotone proportional quantifiers, not even when taking into account truth-values.

2 Experiment

Sixty nine native Polish-speaking adults took part in the study (31 male and 38 female). They were volunteers from the University of Warsaw undergraduate population. The mean age was 21.42 years ($SD = 3.22$) with a range of 18–30 years. Each subject was tested individually and was given a small financial reward for participation in the study.

The task used in the study consisted of sixteen grammatically simple sentences in Polish, containing a quantifier that probed a color feature of a car in a display. Each picture contained fifteen objects in two colors. Four different quantifiers were presented to each subject in four trials. The quantifiers were: fewer than eight, more than seven (numerical of high rank), fewer than half, more than half (proportional).

For each quantifier, half of the sentences were true. The sentences were accompanied by pictures with a quantity of target items near the criterion for validating or falsifying the proposition, therefore requiring a precise judgment (e.g., seven targets in ‘fewer than half’). In each quantifier problem, first

the proposition appeared in the middle of a screen, followed by a blank screen (500 ms) and stimulus array containing 15 randomly distributed cars. Hence, within each trial, the sentence and the picture were presented separately. We recorded then the time used for reading the sentence and the time used for verification with the picture. Debriefing, that followed the experiment, revealed that none of the participants were aware that each picture consisted of fifteen objects.

The stimulus arrays were presented for 15000 ms. Within this time, the subjects were asked to decide if the proposition accurately described the presented picture. They responded by pressing the buttons referring to the first letters of the Polish words for ‘true’ (‘p’) and ‘false’ (‘f’). All stimuli were counterbalanced and randomly distributed throughout the experiment.

2.1 Results

In this study, the sentence was presented before the picture was displayed, hence the true-false conditions were not included in the analysis. ANOVA with type of quantifier (2 levels: numerical, proportional), and monotonicity (2 levels: upward, downward) as the two within-subject factors was used to examine the differences in mean reaction times of sentence mean reading time (see Table 1 for means and standard deviations). There were no significant effects or interactions between the factors.

Table 1: Means (M) and standard deviations (SD) of the reading time in milliseconds of each quantifier

Quantifier	M	SD
More than seven	4054	1992
Fewer than eight	4345	2913
More than half	4459	2907
Fewer than half	4742	2863

Further, the times needed to verify the quantifiers were compared. ANOVA with type of quantifier (2 levels: numerical, proportional), monotonicity (2 levels: upward, downward), and the statement’s truth-value (2 levels: true, false) as three within-subject factors was used to examine differences in mean reaction times of sentence-picture verification (see Table 2).

The analysis indicated significant main effects of quantifier ($F(1, 68) = 146.73, p < 0.001, \eta^2=0.68$) and monotonicity ($F(1, 68) = 6.73, p = 0.012, \eta^2=0.09$), as well as the following interactions: quantifier \times monotonicity ($F(1, 68) = 13.32, p < 0.001, \eta^2=0.16$), quantifier \times truth ($F(1, 68) = 11.58, p = 0.005, \eta^2=0.15$), monotonicity \times truth ($F(1, 68) = 7.93, p = 0.006, \eta^2=0.10$), and quantifier \times monotonicity \times truth ($F(1, 68) = 12.64, p < 0.001, \eta^2=0.16$) (see Figure 1).

Analyzing the latter interaction effect, we compared the differences within each type of quantifier. We performed ANOVA with monotonicity and truth-value as two within-subject factors for the numerical and proportional quantifiers separately.

In the case of the numerical quantifiers, we found significant effects of monotonicity ($F(1, 68) = 43.61, p < 0.001, \eta^2=0.40$), truth ($F(1, 68) = 13.01, p = 0.001, \eta^2=0.16$) and the interaction monotonicity \times truth ($F(1, 68) = 37.54, p < 0.001, \eta^2=0.36$). Pairwise comparisons among means

Table 2: Means (M) and standard deviations (SD) of the verification time in milliseconds of each quantifier with respect to monotonicity and truth-value.

Quantifier		M	SD
More than seven	True	3793	1241
	False	3360	1218
	Overall	3577	964
Fewer than eight	True	3626	1296
	False	5029	1833
	Overall	4327	1307
More than half	True	6511	2454
	False	6475	2195
	Overall	6493	1959
Fewer than half	True	6509	2378
	False	6084	2464
	Overall	6296	2090

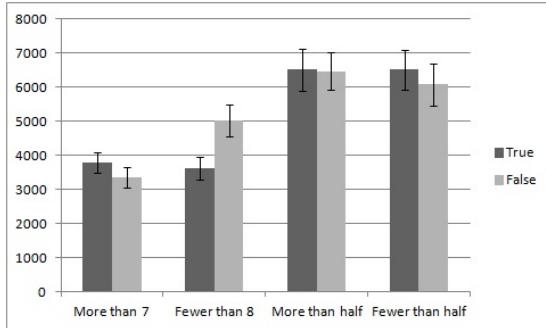


Figure 1: Average reaction time in milliseconds of each experimental condition. Note Error bars are for 95% intervals.

(with LSD test) revealed that the false ‘fewer than eight’ were processed longer than any other quantifier, while false ‘more than seven’ were performed the fastest. Moreover, there was no significant difference between both true conditions ($p > 0.05$).

There were no significant differences within proportional quantifiers ($p > 0.05$).

We also analyzed the main effect of the quantifier type. The analysis revealed that proportional quantifiers ($M=6395$, $SD=1832$) were processed longer than numerical ($M=3952$, $SD=1069$).

The accuracy analysis revealed only one significant effect of quantifier type ($F(1, 68) = 20.23$, $p < 0.001$, $p\eta^2=0.23$): proportional quantifiers were more difficult ($M=.85$, $SD=.02$) than numerical

quantifiers ($M=.94$, $SD=.01$).

Summing up, we can conclude that there are no differences within proportional quantifiers as far as the monotonicity and truth is concerned. With regards to the numerical sentences, their difficulty increased as follows: false ‘more than seven’, both true ‘more than seven’ and ‘fewer than eight’ (equal), and false ‘fewer than eight’. Finally, the proportional statements were generally more difficult than the numerical.

3 Discussion

We have examined the time needed to process sentences containing quantifiers. The quantifiers we have studied differed in their computational complexity, monotonicity, and true-false conditions. The results have confirmed the complexity hypothesis derived from the computational model. As predicted, sentences with the quantifier ‘more than seven’ were processed faster when they were false. In the case of ‘fewer than eight’ true sentences were easier to process than false sentences. In other words, we found an interaction effect between monotonicity and truth-value that reflects the effect of counting in the case of numerical quantifiers. Moreover, our data indicated that there was no significant monotonicity effect within the proportional quantifiers. Again, this is in agreement with the computational theory, according to which the mental strategies for the verification of proportional quantifiers resemble the push-down automata algorithm [2].

The running of the procedure does not differ between the upward monotone case ‘more than half’ and the corresponding downward monotone quantifier ‘less than half’. Furthermore, the complexity of the computation is similar between true and false instances of the proportional sentences. In both cases, one needs to compute and compare the cardinalities of two sets that cover the whole universe. These two facts explain why we found no effect of monotonicity or truth-value in the case of proportional quantifiers.

The average difference in reaction time was additionally consistent with the hypothesis that quantifiers trigger the corresponding minimal computation (counting up to seven or eight). Therefore, our research contributes another argument in favor of the cognitive plausibility of the automata-theoretic model of quantifier verification [2]. Furthermore, we observed that, in general, proportional quantifiers are more difficult than numerical quantifiers, which is again consistent with the theory that predicts a complexity difference between numerical and proportional quantifiers.

As we have already extensively justified throughout the introduction, our computational perspective brings a refinement to the theory proposed in [1]. Namely, we predicted and experimentally found an interaction effect between monotonicity and truth value in the case of numerical quantifiers. This effect follows from the corresponding differences in their computational complexity. Moreover, neither our theory nor the experiments indicate an involvement of monotonicity in the difficulty of proportional quantifier verification. These observations allow suggesting that monotonicity has only a relative effect on the difficulty of verification. Together with other aspects of the situation, monotonicity may influence the complexity of the verification.

In a syllogistic reasoning experiment, [4] has found that if the monotonicity profiles of two quanti-

fying expressions are the same, then they should be equally hard to process. They studied sentences like:

- (1) Some of the sopranos sang with more than three of the tenors.
- (2) None of the sopranos sang with fewer than three of the tenors.
- (3) Some of the sopranos sang with fewer than three of the tenors.

The results suggested that sentences with two upward monotone quantifiers, like (1), are easier to reason with than sentences with two downward monotone quantifiers, e.g., (2). While sentences with two quantifiers of different monotonicity, for instance sentence (3), are the hardest. According to our experiments, in the verification tasks, quantifiers with the same monotonicity profiles can differ with respect to their difficulty depending on the truth-value. This suggest that there might be crucial differences between reasoning with quantifiers and quantifier verification and that future studies devoted to monotonicity should control the true-false conditions and include them in the analyses.

Moreover, our results complement the proposal put forward in [5] that the counting stage in processing is affected by the number mentioned in the quantifier, rather than the critical number of objects needed to verify the statement. They reported a real time study of verification procedures for numerical quantifiers, like ‘more than n ’ and ‘fewer than n ’, using self-paced counting. The used methodology is an analogue of well-known self-paced reading experiments. Subjects hear a sentence and are asked to determine as fast and as reliably as possible its truth-value relative to an array of dots. The arrays are presented as three scattered rows of hexagonal plates. As participants press the space bar, the dots are uncovered in groups of 1, 2, or 3, while previously seen dots are recovered and masked. Participants may answer once they have enough information. The setting allows looking into the verification process by timing how the participants uncover the dots. Using this paradigm, [5] was able to show that reaching the number heard in the quantifier causes a slow down in the processing. This observation is consistent with our data emphasizing the interaction with the truth-value, as changes in the truth-value are necessarily bound with reaching the number n . This also suggests that next to the reading and verification stages, one should also take into account the *decision stage* in quantifier verification.¹ One would predict that the interactions among all three processing stages: reading, verification, and decision, may, for instance, play a crucial role in explaining the differences between comparative quantifiers, e.g., ‘more than 3’ and the equivalent superlative quantifiers, e.g., ‘at least 4’. We know that superlative quantifiers are harder to verify than the corresponding comparative quantifiers but there are no differences in reading times [3]. Our computational approach predicts no differences in the verification times as the counting processes for equivalent quantifiers are identical, i.e., the computations for ‘more than 3’ and ‘at least 4’ do not differ in complexity. Therefore, we would predict that the difference between comparative and superlative quantifiers is due to the decision stage. This prediction falls outside the scope of the current work but should be tested in the future experiments.

¹Interestingly, [5] chose the quantifiers in a way that no matter whether the sentence was true or false, the subjects always needed to count only up to seven, i.e., the number heard, n , varied across true and false items, e.g., ‘more than six’ (true) but ‘more than seven’ (false) and ‘fewer than eight’ (true) but ‘fewer than seven’ (true). As a result they found out that monotone increasing quantifiers are quicker to verify than falsify and monotone decreasing quantifiers are quicker to falsify than verify.

References

- [1] Jon Barwise and Robin Cooper. Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4:159–219, 1981.
- [2] Johan van Benthem. *Essays in Logical Semantics*. D. Reidel, Dordrecht, 1986.
- [3] Bart Geurts, Napoleon Katsos, Chris Cummins, Jonas Moens, and Leo Noordman. Scalar quantifiers: Logic, acquisition, and processing. *Language and Cognitive Processes*, 25(1):244–253, 2009.
- [4] Bart Geurts and Frans van der Silk. Monotonicity and processing load. *Journal of Semantics*, 22:97–117, 2005.
- [5] Jorie Koster-Moeller, Jason Varvoutis, and Martin Hackl. Verification procedures for modified numeral quantifiers. In N. Abner and J. Bishop, editors, *Proceedings of the 27th West Coast Conference on Formal Linguistics*, volume 1986, pages 310–317, Somerville, 2008. Cascadilla Proceedings Project.
- [6] Corey T. McMillan, Robin Clark, Peachie Moore, Christian Devita, and Murray Grossman. Neural basis for generalized quantifier comprehension. *Neuropsychologia*, 43:1729–1737, 2005.
- [7] Corey T. McMillan, Robin Clark, Peachie Moore, and Murray Grossman. Quantifiers comprehension in corticobasal degeneration. *Brain and Cognition*, 65:250–260, 2006.
- [8] Jakub Szymanik. *Quantifiers in TIME and SPACE. Computational Complexity of Generalized Quantifiers in Natural Language*. PhD thesis, University of Amsterdam, Amsterdam, 2009.
- [9] Jakub Szymanik and Marcin Zajenkowski. Comprehension of simple quantifiers. Empirical evaluation of a computational model. *Cognitive Science: A Multidisciplinary Journal*, 34(3):521–532, 2010.
- [10] Jakub Szymanik and Marcin Zajenkowski. Quantifiers and working memory. In M. Aloni and K. Schulz, editors, *Amsterdam Colloquium 2009, Lecture Notes In Artificial Intelligence 6042*, pages 456–464. Springer, 2010.
- [11] Jakub Szymanik and Marcin Zajenkowski. Contribution of working memory in parity and proportional judgments. *Belgian Journal of Linguistics*, 25(1):176–194, 2011.
- [12] Marcin Zajenkowski, Rafał Styła, and Jakub Szymanik. A computational approach to quantifiers as an explanation for some language impairments in schizophrenia. *Journal of Communication Disorders*, 44(6):595 – 600, 2011.
- [13] Marcin Zajenkowski and Jakub Szymanik. Most intelligent people are accurate and some fast people are intelligent: Intelligence, working memory, and semantic processing of quantifiers from a computational perspective. *Intelligence*, 41(5):456 – 466, 2013.

Focus association in superlatives and the semantics of *-est*

Barbara Tomaszewicz*

University of Southern California

barbara.tomaszewicz@usc.edu

Abstract

I provide new evidence from Polish that two lexical entries for the superlative morpheme *-est* are needed in the grammar of a single language: 3-place *-est* on which individuals are compared and 2-place *-est* comparing sets of degrees, (Heim, 1999) (Romero, 2010). I argue that 2-place *-est* does not associate with focus, but is used in cases of explicit comparison between degrees, e.g. in the presence of a degree relative clause in Polish. I show that focus association with 3-place *-est* accounts for the range of superlative readings found cross-linguistically (Pancheva & Tomaszewicz, 2012).

1 Superlative ambiguities and focus

The semantics of the superlative morpheme *-est* remains a question of debate, as does the role of focus in superlatives. First, the presence of a superlative expression typically allows for different interpretations of the sentence, some of which are sensitive to the presence of focus (Ross (1964), Heim (1985), Szabolcsi (1986), Gawron (1995) a.o.). For instance, the sentences (1a-b) are ambiguous. They each allow a reading where cakes are compared without regard as to who bought them for whom, as in (1i), the so-called ‘absolute reading’. Additionally, each sentence also has a so-called ‘relative’ reading, (1ii-iii). Focus on *John*, (1a), expressed as intonational prominence, clearly biases interpretation towards the relative reading in (1ii). Focus on *Mary*, (1b), reverses the effect on relative readings in favor of (1iii). The absolute reading is available irrespective of focus, thus the absolute/relative ambiguity cannot be due solely to the effect of focus.

- (1) a. [John]_{Focus} bought Mary [DP the most expensive cake]. Readings: (i), (ii), but not (iii)
b. John bought [Mary]_{Focus} [DP the most expensive cake]. Readings: (i), (iii), but not (ii)
(i) ‘John bought Mary the cake that was more expensive than any other cake.’ *Absolute*
(ii) ‘John bought Mary a more expensive cake than anyone else bought her.’ *Relative*
(iii) ‘John bought Mary a more expensive cake than he bought for anyone else.’ *Relative*

Second, the different kinds of comparison on each reading can be modeled in two ways, as comparison between individuals or as comparison between degrees (Heim, 1999). Theoretically, the two issues, focus sensitivity and the two “modes of comparison”, to use Kennedy’s (1999) term, can be treated independently. I will argue, however, that the empirical observations in Pancheva & Tomaszewicz (2012) reveal not only that association with focus is necessary to account for cross-linguistic differences in the interpretation of superlatives, but also that the semantics of the superlative

* I would like to thank Roumi Pancheva and Yael Sharvit for discussion and feedback that greatly improved the paper. Thanks are also due to Maria Luisa Zubizarreta, Barry Schein, Martin Hackl, Sigrid Beck, Mojmir Docekal and colloquium participants at Tübingen University and at Masaryk University. The errors are all mine.

involves comparison of individuals and comparison of degrees in different, mutually exclusive contexts.

Pancheva & Tomaszewicz (2012) show that there is cross-linguistic variation in the availability of certain relative readings. In English, comparison on relative readings can only be established with respect to a focused nominal that is external to the superlative DP ('John' or 'Mary' in (1)), but not with respect to a DP-internal focus, e.g. 'cake' in (1). Such DP-internal relative readings are found in the Slavic languages, however. The Polish sentence in (2a) can have the DP-internal relative reading (2b).

- (2) a. Jan kupił Marii naj-droższe ciastko. (Polish)
Jan bought for-Mary *est*-expensive cake Readings: (1i), (1ii), (1iii), (2b).
b. 'John bought Mary a more expensive cake than anything else he bought her.'

DP-Internal Relative

I add here an empirical argument that the reading (2b) requires focus on 'cake'. The sentence in (2), where the superlative adjective is in its base position, is four-way ambiguous and prosody disambiguates between different readings. In (3), however, where the adjective 'most expensive' is fronted (via Left Branch Extraction), only one reading obtains, the DP-internal relative reading (2b). The superlative adjective in (3) is topicalized (optionally, a topic marker '*to*' may be present), while the noun 'cake' is focused in conformity with intonational phrasing in Polish (cf. Féry et al. (2007)).

- (3) Context: 'Jan poszedł kupić różne drogie słodycze dla Marii. ...' (Polish)
'Jan went to buy various expensive sweets for Maria. ...'
[[NajDROższe₁]_{Topic} (to)]_i [Jan kupił Marii _{t₁} [CIAstko]_{Focus}]_i
LH* HL*
est-expensive (Topic) Jan bought for-Mary cake
Readings: (2b), *(1i), *(1ii), *(1iii)

The way focus determines the relative readings of *-est* resembles focus effects with the exclusive particle *only*, whose truth-conditional contribution relies on association with focus (Rooth (1985), (1992), Beaver & Clark (2008)). *Only* in the 'split construction' in Polish behaves in exactly the same way as the superlative. In (4a) *only* must associate with 'cake', while in (4b) association with 'cake' is impossible.

- (4) Context: 'Jan poszedł kupić różne słodycze dla Marii. ...' (Polish)
'Jan went to buy various sweets for Maria. ...'
a. Drogie_i *pro* kupił jej tylko _{t_i} ciastko.
expensive bought her only cake
'He bought her sweets of which only the cake was expensive.'
b. *pro* Kupił jej tylko drogie ciastko.
bought her only expensive cake
'He bought her only an expensive cake and nothing else.'

Crucially, the split superlative construction in (3) precludes both the absolute, (1i), and the DP-external relative readings, (1ii)-(1iii). This indicates that superlative ambiguities are grammatical and not just due to contextual vagueness. I follow Pancheva & Tomaszewicz (2012) in arguing that focus association is obligatory for relative readings (1ii-iii), (2b), and that the impossibility of moving *-est* outside the superlative DP in the presence of the definite article is what blocks the DP-internal relative reading (2b) for the sentence in (1). I show that only the lexical entry on which individuals are

compared (together with focus association) correctly derives the cross-linguistic differences in the availability of the DP-internal relative reading (section 3). I further show that the lexical entry on which degrees are compared can be used in contexts where focus association does not take place, but where comparison of degrees is required by the presence of a degree relative clause (section 4). In section 2, I introduce the syntactic and semantic assumptions for the derivation of superlative ambiguities.

2 The semantics of *-est*

Heim (1999) introduces two lexical entries for the *-est* morpheme with truth-conditionally equivalent meaning. The semantics in (5) involves a comparison between individuals (the comparison class C is of type $\langle e, t \rangle$)¹, while (6) calls for a comparison between sets of degrees (C is of type $\langle dt, t \rangle$).²

$$(5) \llbracket [-est_{3\text{-place}}] \rrbracket = \lambda C_{\langle e, t \rangle}. \lambda P_{\langle d, et \rangle}. \lambda x_e. \exists d [P(d)(x) \wedge \forall y \in C [y \neq x \rightarrow \neg(P(d)(y))]]$$

Presuppositions: $x \in C, \forall y [y \in C \rightarrow \exists d [P(d)(y)]]$

$$(6) \llbracket [-est_{2\text{-place}}] \rrbracket = \lambda C_{\langle dt, t \rangle}. \lambda P_{\langle d, t \rangle}. \exists d [P(d) \wedge \forall Q \in C [Q \neq P \rightarrow \neg(Q(d))]]$$

Presuppositions: $P \in C, \forall Q [[Q \in C \wedge Q \neq P] \rightarrow \exists d [Q(d)]]$

Both entries, (5)-(6), can derive the absolute and relative readings. Assuming that *-est* can take scope inside or outside of the superlative DP (Szabolcsi (1986), (2012), Heim (1999)), comparison of individuals (using *-est*_{3-place}, (5)) gives us comparison between cakes on the absolute reading, (7a), and between people who bought cakes for Mary on one of the relative readings, (7b). Degree comparison (using *-est*_{2-place}, (6)) involves comparison between different sets of degrees, depending on the scope of *-est*, (8a-b). We cannot judge by introspection whether cakes are compared, cake buyers, or prices of cakes (Heim (1999), Farkas & Kiss (2000), Sharvit & Stateva (2002) a.o.).

- | | | |
|--|-----------------------|----------------------|
| (7) a. Jan bought Mary [_{DP} the $\llbracket [-est C_{\langle e, t \rangle}] \lambda d. d\text{-expensive cake} \rrbracket$]
$C_{\langle e, t \rangle} = \{x: \exists d [x \text{ is a } d\text{-expensive cake}\}]$ | <i>Absolute</i> (1i) | |
| b. Jan $\llbracket [-est C_{\langle e, t \rangle}] \lambda d. \lambda x. x \text{ bought Mary } [\substack{\text{DP } d\text{-expensive cake}]} \rrbracket$
$C_{\langle e, t \rangle} = \{x: \exists d [x \text{ bought Mary a } d\text{-expensive cake}\}]$ | <i>Relative</i> (1ii) | |
| (8) a. Jan bought Mary [_{DP} the $\llbracket [-est C_{\langle dt, t \rangle}] \lambda d. d\text{-expensive cake} \rrbracket$]
$C_{\langle dt, t \rangle} = \{D: \exists x [D = \lambda d. x \text{ is a } d\text{-expensive cake}\}]$ | | <i>Absolute</i> (1i) |
| b. $\llbracket [-est C_{\langle dt, t \rangle}] \lambda d. \lambda x. x \text{ bought Mary } [\substack{\text{DP } d\text{-expensive cake}]} \rrbracket$
$C_{\langle dt, t \rangle} = \{D: \exists x [D = \lambda d. x \text{ bought Mary a } d\text{-expensive cake}\}]$ | <i>Relative</i> (1ii) | |

Unlike Heim (1999) and Szabolcsi (1986), (2012), Farkas & Kiss (2000) and Sharvit & Stateva (2002) argue that *-est* can only take DP-internal scope. On neither approach, however, is focus taken to be necessary for relative readings.

Heim (1999) introduced the 2-place semantics for *-est* specifically to allow focus to determine what enters the comparison class via the mechanism of Rooth's (1985), (1992) theory of focus interpretation. The crucial ingredient of this account is Rooth's (1992) ~ ('squiggle'), a focus operator which comes with its own restrictor variable, S . The ~ introduces the presupposition that S is a subset

¹ Technically, the variable C in (5) is the characteristic function of a set of individuals. I will refer to it as a set, for convenience. The same goes for C in (6).

² The lexical entries in (5)-(6) require the assumption that gradable predicates are *downward monotonic*, (i):

(i) A relation R between objects and degrees is downward monotonic iff: $\forall x \forall d \forall d' [R(x, d) = 1 \wedge d' < d \rightarrow R(x, d')]$

of the focus-value of the constituent to which $[\sim S]$ attaches, typically the clause. When a quantificational element associates with focus, its (covert) restrictor variable is set to be the subset of S . For instance, the truth-conditional contribution of the exclusive particle *only* is determined by association with focus, (9)-(10), (Rooth (1985), (1992), von Fintel (1994), Beaver & Clark (2008)).

- (9) a. John *only* gave [Mary]_F a cheap gift.
- b. LF for (9a): $[\text{Only } C] [[\sim S] [\text{TP}_1 \text{John gave [Mary]}_F \text{ a cheap gift}]]$
- c. $\llbracket (9b) \rrbracket = \lambda w. \forall p [(p \in C \wedge p \neq \llbracket \text{John gave Mary a cheap gift} \rrbracket) \rightarrow \neg p(w)]$
- d. $C \subseteq S, C = \{P: P = \exists x [\text{John gave } x \text{ a cheap gift}]\}$
- (10) a. John *only* gave Mary a [cheap]_F gift.
- b. LF for (10a): $[\text{Only } C] [[\sim S] [\text{TP}_1 \text{John gave Mary a [cheap]}_F \text{ gift}]]$
- c. $\llbracket (10b) \rrbracket = \lambda w. \forall p [(p \in C \wedge p \neq \llbracket \text{John gave Mary a cheap gift} \rrbracket) \rightarrow \neg p(w)]$
- d. $C \subseteq S, C = \{P: P = \exists \exists x [\text{John gave Mary a gift} \wedge \text{gift}(x) \wedge f(x)]\}$

Heim (1999) suggests that the truth-conditional effects of focus on the relative readings, (1), should be modeled as a contextual effect on the restrictor C , entirely parallel to the focus effects with *only*. Adding the focus association condition on the relation between C and S is straightforwardly achieved with $\text{-est}_{2\text{-place}}$, (11)-(12). Like *only*, $\text{est}_{2\text{-place}}$ takes scope over the entire sentence and in (11a) the focus value of the sister of $[\sim S]$, TP_2 , is a set of sets of degrees, hence S is of type $\langle dt, t \rangle$ – the right type for C . Focus is interpreted in-situ, (11c), a focus feature on ‘John’ specifies C as in (11e), while focus on ‘Mary’ gives C the specification in (12b).

- (11) a. LF for (1ii): $[\text{-est } C_{\langle dt, t \rangle}] [\text{TP}_3 [[\sim S_{\langle dt, t \rangle}] [\text{TP}_2 [\text{John}]_F \text{ bought Mary a } d\text{-expensive cake}]]]$
- b. $\llbracket (11a) \rrbracket = 1$ iff
 $\exists d \exists x [\text{cake}(x) \wedge \text{expensive}(x, d) \wedge \text{bought}(j, x, m)] \wedge \forall Q \in C [Q \neq \llbracket \text{TP}_3 \rrbracket \rightarrow \neg (Q(d))]$
- c. $C \subseteq S, S \subseteq \llbracket \text{TP}_2 \rrbracket^f$ (focus association)
- d. $\llbracket \text{TP}_2 \rrbracket^f = \{D: \exists x [D = \lambda d. x \text{ bought Mary a } d\text{-expensive cake}]\}$
- e. $C_{\langle dt, t \rangle} = \{D: \exists x [D = \lambda d. x \text{ bought Mary a } d\text{-expensive cake}]\}$
- (12) a. LF for (1iii): $[\text{-est } C_{\langle dt, t \rangle}] [\text{TP}_3 [[\sim S_{\langle dt, t \rangle}] [\text{TP}_2 \text{ John bought [Mary]}_F \text{ a } d\text{-expensive cake}]]]$
- b. $C_{\langle dt, t \rangle} = \{D: \exists x [D = \lambda d. \text{John bought } x \text{ a } d\text{-expensive cake}]\}$

Focus association with $\text{est}_{3\text{-place}}$ taking DP-external scope on relative readings, e.g. (7b), requires that either $[\sim S]$ attaches to the focus marked element that QRs to saturate the individual argument of $[\text{-est } C]$, (Heim, 1999), or that $[\sim S]$ attaches at the clausal level and the trace of the moved third argument is marked with the focus feature, (13) (Pancheva & Tomasiewicz, 2012).

- (13) a. LF for (1ii), (Pancheva & Tomasiewicz, 2012):
 $[\text{TP}_1 \text{John} [\text{TP}_2 [\text{-est } C] [\text{TP}_3 [[\sim S] [\text{TP}_4 [x]_F \text{ bought Mary a } d\text{-expensive cake}]]]]]$
- b. $S \subseteq \llbracket \text{TP}_4 \rrbracket^f$
- c. $\llbracket \text{TP}_4 \rrbracket^f = \{D: \exists d [D = \lambda x [x \text{ bought Mary a } d\text{-expensive cake}]]\}$
- d. $C = \cup S$ (focus association) (von Fintel, 1994)
- e. $C = \{x: \exists d [x \text{ bought Mary a } d\text{-expensive cake}]\}$ (focus association)
- f. $C = \{x: \exists d [x \text{ bought Mary a } d\text{-expensive cake}]\}$ (presupposition of -est , (5))

Since focus marking is realized prosodically as intonational prominence, ideally \sim should scope over both focus marked and unmarked portions of the sentence, as it does with $\text{est}_{2\text{-place}}$. However, focus effects in the interpretation of superlatives cannot be accounted for using the 2-place semantics for -est , as I show below.

3 Restrictions on DP-internal relative readings

3.1 Wrong predictions of *-est*_{2-place} for DP-internal relative readings

The 2-place semantics for *-est* makes the prediction that DP-internal focus, (14), triggers the relative reading in (14b) with the comparison set (14e). As shown in section 1, this reading is available in Polish, (2b), but not in English, even though the same DP-internal focus does have a truth-conditional effect with *only*, (15).

- (14) a. John bought Mary [_{DP} the most expensive [cake]_F].
b. ‘John bought Mary a more expensive cake than anything else he bought her.’
DP-Internal Relative
c. LF for (14a): [-est C_(dt,t)][TP₃ [~ S_(dt,t)] [TP₂ John bought Mary a *d*-expensive [cake]_F]]
d. $\llbracket \text{TP}_2 \rrbracket^f = \{D: \exists f_{(e,t)} \exists x [D = \lambda d. f(x) \wedge x \text{ is } d\text{-expensive} \wedge \text{John bought } x \text{ for Mary}\}\}$
e. $C_{(dt,t)} = \{D: \exists f_{(e,t)} \exists x [D = \lambda d. f(x) \wedge x \text{ is } d\text{-expensive} \wedge \text{John bought } x \text{ for Mary}\}\}$
- (15) John *only* gave Mary [_{DP} a cheap [cake]_F].
‘John gave Mary nothing else that was cheap except the cake.’

The contrast between Polish and English with respect to the reading in (14b)/(2b) does not follow from the presence vs. absence of *-est*_{2-place} in the language. Bulgarian allows DP-internal relative readings, so *-est*_{2-place} could be available in this language. I show, however, that using *-est*_{2-place} and focus association we cannot account for the fact observed by Pancheva & Tomaszewicz (2012) that in the presence of the definite determiner, (16), only DP-external relative readings are available, just like for the English (1). When the definite is absent, (17), both DP-internal and DP-external relative readings obtain, with suitable prosody, whereas the absolute reading is unavailable.³

- (16) Ivan kупи [_{DP} *naj-skupa-ta* torta] за Meri. (Bulgarian)
Ivan bought *est-expensive-the* cake for Mary Readings: *(14b), (1i), (1ii), (1iii)
(17) Ivan kупи [_{DP} *naj-skupa* torta] за Meri.
Ivan bought *est-expensive* cake for Mary Readings: (14b), *(1i), (1ii), (1iii)

Assuming that the same lexical entry is used to derive relative readings in both (16) and (17), the goal is to account for the blocking effect of the definite determiner on association with DP-internal focus. Pancheva & Tomaszewicz (2012) achieve this with the 3-place semantics for *-est*, as reviewed in section 3.2. I show that an account of the cross-linguistic differences is not possible with the 2-place semantics, in section 3.3. I further offer empirical evidence that *-est*_{2-place} does not associate with focus, whether DP-internal or external, on relative readings.

3.2 Constraining DP-internal relative readings with *-est*_{3-place}

Pancheva & Tomaszewicz (2012) argue on the basis of Bulgarian and other Slavic languages, that the DP-internal relative reading is derived by obligatory focus association, and that DP-external scope for *-est* (as in (13)) is possible only in the absence of the definite determiner. In the presence of *the*, *-est* is trapped inside the DP. The evidence that *-est* cannot be interpreted DP-internally on the DP-internal

³ Since in Slavic the availability/blocking of the DP-internal relative reading is strictly correlated with the definiteness of the superlative DP, Pancheva & Tomaszewicz (2012) propose that definite island effects on degree movement are universal. Their additional support for the claim that the definite DPs are degree islands comes from the fact that also QR of the comparative *-er* is blocked out of definite DPs, in English, and in the Slavic languages that have a definite article.

relative reading comes from sentences such as (3) and (17) where the absolute reading is not available, i.e. the comparison class determined on the basis of the sister of [-est C] can be neither $C_{(e,t)} = \{x: \exists d. x \text{ is a } d\text{-expensive cake}\}$ nor $C_{(dt,t)} \subseteq \{D: \exists x [D = \lambda d. x \text{ is a } d\text{-expensive cake}\}\}$. This shows that the absolute/relative ambiguity is not simply a matter of the same LF and different context as has been suggested (Farkas & Kiss (2000), Sharvit & Stateva (2002)).

Mandating DP-internal scope for $-est_{3\text{-place}}$ in the presence of the definite article, Pancheva & Tomaszewicz (2012) derive the impossibility of DP-internal relative readings. A definite superlative DP can QR and -est can associate with a DP-external focus. In (18a) where the focus is on ‘John’, -est can associate with it, since the condition on focus association, (18c), and the presupposition of -est, (18d) match.

- (18) a. LF for (1ii) in the presence of the definite (vs. (13)), (Pancheva & Tomaszewicz, 2012):
 $[\text{TP}_3 [\text{DP} \text{ the } [-est C_{(e,t)}] [\text{NP } d\text{-expensive cake}]]_1 [\text{TP}_2 [\sim S] [\text{TP}_1 [\text{John}]_F \text{ bought Mary } t_1]]]$
b. $S \subseteq [\text{TP}_2]^f, [\text{TP}_2]^f = \{P: \exists y [P = \lambda x [y \text{ bought Mary } x]\}$
c. $C = \cup S = \{x: \exists y [y \text{ bought Mary } x]\}$ (focus association)
d. $C = \{x: \exists d [x \text{ is a } d\text{-expensive cake}\}\}$ (presupposition of -est (5))

If focus were on ‘Mary’ in (18), the requirements of focus association and of the presupposition would also be compatible, but with focus on ‘cake’ they would clash, preventing the derivation of the DP-internal reading. Pancheva & Tomaszewicz (2012) demonstrate that when -est stays inside the definite superlative DP, whatever configuration with respect to the $[\sim S]$ complex is attempted, the DP-internal relative reading is never derived because the DP-internal -est fails to associate with focus, i.e. the condition on focus association and the presuppositions of -est, (5), are in conflict. In the next section I explore whether the blocking effect of the definite determiner on the association with DP-internal focus can be accounted for on the 2-place semantics for -est.

3.3 Attempting to constrain focus association of $est_{2\text{-place}}$

There are three possible configurations where $est_{2\text{-place}}$ takes scope within the DP (its QR being blocked by the presence of the definite article) and attempts to associate with the focus on ‘cake’ for the DP-internal relative reading. It turns out that the blocking effect of the definite in Bulgarian, (16), and in English, (14a-b), cannot be modeled using $est_{2\text{-place}}$. Forcing DP-internal scope for $est_{2\text{-place}}$ never results in an LF where only association with DP-external focus is possible. And allowing $est_{2\text{-place}}$ to take sentential scope cannot handle the cross-linguistic facts either.

The first option is for $[-est_{2\text{-place}} C]$ to simply raise below the definite article, (19), but then its second argument NP₂ is not of the right type. NP₂ has the type $\langle d, \text{et} \rangle$, and not $\langle dt \rangle$ as required.

- (19) John bought $[\text{DP} \text{ the } [-est C_{(dt,t)}]]_{(dt,t)} [\text{NP}_2 [\sim S] [\text{NP}_1 d\text{-expensive [cake]}_F]]_{\langle d, \text{et} \rangle}$

One way to avoid the type clash in (19) is to move $[-est_{2\text{-place}} C]$ to the edge of the DP, (20), and assume the determiner to contribute existential quantification. This second option, however, does not block the DP-internal reading. It results in an interpretation unattested in the presence of the definite, namely, that John bought a cake that was more expensive than anything relevant in the context, (20b). If $[\sim S]$ has sentential scope, we get the DP-internal relative reading in (14).

- (20) a. John bought $[\text{DP}_3 [-est C_{(dt,t)}]]_{(dt,t)} [\text{DP}_2 [\sim S] [\text{DP}_1 \text{ the/a } d\text{-expensive [cake]}_F]]_{\langle d, \text{et} \rangle}$
b. $[\text{DP}_3] = \exists d. \exists x. [\text{cake}(x) \wedge \text{expensive}(x,d) \wedge \forall Q \in C [Q \neq [\text{DP}_3] \rightarrow \neg(Q(d))]$
c. $C \subseteq S, S \subseteq [\text{DP}_1]^f$ (focus association)
d. $[\text{DP}_1]^f = \{D: \exists f \exists x [D = \lambda d. f(x) \wedge \text{expensive}(x,d)]\}$

$$e. \quad C_{(dt,t)} = \{ D : \exists f \exists x [D = \lambda d. f(x) \wedge \text{expensive}(x,d)] \}$$

The third option is to remove the superlative DP from the scope of the focus operator \sim , (21). This way we mandate DP-external focus, but we face further problems. We again need to avoid the type clash, e.g. by allowing the definite determiner to move leaving a variable of type e (as proposed by Romero (2010) for the derivation of the absolute reading with $-est_{2\text{-place}}$). More seriously, the focus association condition $C \subseteq S, S \subseteq [\![TP_2]\!]^f$, cannot be satisfied given that DP_2 , which by presupposition in (6) is a member of C , (21b), contains an unbound variable, (21c).

- (21) a. $[DP_2 \text{ the}_2 [-est C_{(dt,t)}]_{(d,t)} [DP_1 t_2 d\text{-expensive cake}]_{(d,t)}]_1 [TP_2 [\sim S] [TP_1 [\text{John}]_F \text{ bought } t_1]$
b. $[\![DP_2]\!] \in C$ (presupposition of $-est$, (6))
c. $[\![DP_2]\!]^g = \lambda d. [\text{cake}(g(2)) \wedge \text{expensive}((g2),d)]$

No configuration where the scope of $-est_{2\text{-place}}$ is constrained by the presence of the definite article derives DP-external relative readings while simultaneously blocking the DP-internal relative reading. This suggests that only $-est_{3\text{-place}}$ together with focus and independent facts about the structure of the DP (definiteness) can determine compositionally which superlative readings are (un)available in English and Bulgarian. But could $-est_{2\text{-place}}$ be assumed for relative readings in Polish since this language lacks the definite article, which has the blocking effect in Bulgarian and English?

4 Degree comparison ($-est_{2\text{-place}}$) and degree relative clauses

Polish has a dedicated degree *wh*-operator, so the compatibility of degree relative clauses and superlatives provides a diagnostic for the type of comparison: between individuals or degrees. Modification by a degree relative clause, (22), does not allow for the relative reading on which individuals are compared, (22b). Consequently, prosodic focus in (23) results in the unacceptability of the degree relative, in contrast to (24) where the superlative quantifier is focused.

- (22) Jan kupił najwięcej ciastek, ile było dozwolone.
Jan bought most cakes how-much was allowed
a. ‘Jan bought the largest amount of cakes that was allowed.’
b. *‘Jan bought a larger allowed amount of cakes than anyone else did.’
- (23) [JAN]_F kupił najwięcej ciastek, (*ile ktokolwiek kupił).
Jan bought most cakes how-much anyone bought
‘Jan bought the most cakes that anyone bought.’
- (24) Jan kupił [najWIĘcej]_F ciastek, ile ktokolwiek widział.
Jan bought most cakes how-much anyone saw
‘Jan bought the most cakes that anyone saw.’

Since prosodic focus on ‘Jan’ in (23) triggers the relative reading, but that reading is incompatible with the degree relative, I conclude that the relative interpretation is derived with $-est_{3\text{-place}}$. With $-est_{2\text{-place}}$ the comparison would be between different amounts of cakes, just like in (22), and the degree relative clause should be able to further specify this set of sets of degrees. Because (22) and (24) show that degree relative clauses can modify superlatives, we have evidence that $-est_{2\text{-place}}$ can be used in Polish, since the degree relative can be taken to express the comparison set $C_{(dt,t)}$, as proposed for English by Howard (2013) and for English modal superlatives by Romero (2010).

5 Conclusion

I have provided new evidence from Polish that two lexical entries for the superlative morpheme can be found in the grammar of a single language: the 3-place *-est* on which individuals are compared, (5), and the 2-place *-est* comparing sets of degrees, (6). The two morphemes are in complementary distribution in Polish and cross-linguistically. Unlike what is suggested in Heim (1999), *-est_{2-place}* does not associate with focus to derive the focus affected relative readings. Focus association with *-est_{2-place}* makes wrong predictions for the range of superlative interpretations available cross-linguistically as identified in Pancheva & Tomaszewicz (2012). However, *-est_{2-place}* is used in cases of explicit comparison between degrees, e.g. in the presence of a degree relative clause. I concluded that the grammar uses *-est_{3-place}* for the relative readings derived by focus association. This is in line with Szabolcsi's (2012) suggestion that different ways for building superlatives "may coexist in (varieties of) the same language".

References

- Beaver, D., & Clark, B. (2008). *Sense and Sensitivity: How Focus Determines Meaning*. Oxford: Wiley-Blackwell.
- Farkas, D., & Kiss, K. É. (2000). Comparative and absolute readings of superlatives. *Natural Language and Linguistic Theory*, 18, 417-455. doi: 10.1023/A:1006431429816.
- Féry, C., Paslawska, A., & Fanselow, G. (2007). Nominal split constructions in Ukrainian. *Journal of Slavic Linguistics*, 15 (1), 3-48.
- Gawron, M. (1995). Comparatives, superlatives, and resolution. *Linguistics and Philosophy*, 8, 333–380. doi: 10.1007/BF00984929.
- Heim, I. (1985). *Notes on comparatives and related matters*. University of Texas-Austin: Unpublished Manuscript.
- Heim, I. (1999). *Notes on superlatives*. Retrieved from Semanticsarchive.net: <http://semanticsarchive.net/Archive/TI1MTlhZ/Superlative.pdf>
- Howard, E. (2013). *Superlative degree clauses: Evidence from NPI licensing*. Semantics and Linguistic Theory (SALT) 23 May 3-5 2013 handout.
- Kennedy, C. (1999). *Projecting the Adjective: The Syntax and Semantics of Gradability and Comparison*. New York: Garland.
- Pancheva, R., & Tomaszewicz, B. (2012). Cross-linguistic Differences in Superlative Movement out of Nominal Phrases. In N. Arnett, & R. Bennett (Ed.), *Proceedings of the 30th West Coast Conference on Formal Linguistics* (pp. 292-302). Cascadilla Proceedings.
- Romero, M. (2010). Modal superlatives and 3-place vs. 2-place *-est*. *The Baltic International Yearbook Of Cognition, Logic And Communication*, 6, doi:10.4148/biyclc.v6i0.1577.
- Rooth, M. (1985). *A Theory of Focus Interpretation*. PhD dissertation, UMass, Amherst.
- Rooth, M. (1992). A Theory of Focus Interpretation. *Natural Language Semantics*, 1, 75–116.
- Ross, J. R. (1964). *A Partial Grammar of English Superlatives*. University of Pennsylvania, MA Thesis.
- Sharvit, Y., & Stateva, P. (2002). Superlative expressions, context, and focus. *Linguistics and Philosophy*, 25, 453-504. doi: 10.1023/A:1020875809794.
- Szabolcsi, A. (1986). Comparative superlatives. *MIT Working Papers in Linguistics*, 8, 245–266.
- Szabolcsi, A. (2012). Compositionality without word boundaries: (the) more and (the) most. *Proceedings of SALT 22* (pp. 1-25). eLanguage.
- von Fintel, K. (1994). *Restrictions on Quantifier Domains*. PhD dissertation, UMass, Amherst.

The type of adjectives

Galit Weidman Sassoon

Bar Ilan University

galitadar@gmail.com

Abstract

A compositional analysis of a range of readings of comparison constructions, as well as the positive form is proposed, which, unlike previous accounts, is compatible with multidimensional adjectives and has the power to explain differences between them and nouns.

To this end, adjectives are represented as properties of dimensional quantifiers, namely of sets of gradable properties; e.g., healthy $\Leftrightarrow \lambda GQ<\text{et}, t>$. n-many($\lambda F.F$ is a health dimension, $\lambda F.GQ(F)$), where *many* denotes a cardinality function and *n* sets up a standard. Comparison morphemes either set the standard of *many* or of the dimensions. Consequences are discussed for our understanding of the adjective-noun distinction and for the analysis of gradable morphology.

1. The challenge

Scholars generally agree that natural languages provide evidence for a taxonomy of predicates consisting of word classes such as nouns, adjectives and verbs (Baker 2003: 1-16). Some semantic analyses distinguish between verbs and other predicate types, analyzing verbs as denoting event types (Landman 2000), but they do not say what distinguishes adjectives from nouns. Nouns tend to occur in argument position, where their main function is to refer to objects, whereas adjectives typically occur in predicate- or modifier-position, as in (1a-b) and (2a), respectively. However, nouns can freely occur in predicate and modifier positions, too, as in (1c) and (2b) (Baker 2003: chap. 4).

- 1) a. John is healthy.
b. John is sick.
c. Tweety is a duck.
- 2) a. The healthy boys came.
b. This is an elephant turtle.

The compatibility of adjectives with so-called degree morphemes, as in, for instance, (3a)-(4b), calls for a more complicated type than that of a mere predicate or predicate modifier. A dominant response to this challenge is a degree-function analysis.

- 3) a. John is healthier than Mary (is).
b. John is the healthiest.
- 4) a. The table is longer than the sofa is wide.
b. The sofas are more similar than dissimilar.
c. John is more a linguist than a philosopher.

Let a model M be a quadruple $\langle I, D_t, D_e, D_d \rangle$ consisting of an interpretation function I and domains of truth values, entities and degrees, respectively, and let G be a set of variable assignments g . On a degree analysis, in every M and G , adjectives denote degree functions, $f: D_e \rightarrow D_d$, or corresponding relations, $R_f = \lambda x \in D_e \lambda d \in D_d. f(x) \geq d$ (von Stechow 1984; Kennedy 1999). Hence, the positive form, e.g., ‘John is healthy’, is translated to the form $F(\text{John})(c)$ (‘John is c F’), which involves a null element c . This element sets up a categorization threshold such that for $g(c) = d$, $R_f(j)(d)$ is true iff the degree $f(j)$ is equal to or bigger than d .

Alternative analyses to gradability and comparison exist, which similarly postulate a null parameter c for the assignment of truth value to the positive form, but do not postulate degrees. In particular, on a comparison-class analysis, dimensions, $F_{\langle et, et \rangle}$, are functions from an entity set called comparison class, c , to a subset, the positive extension of F in c . Thus, assuming a comparison class variable c in the logical form (where $g(c_{et}) \subseteq D_e$), $F(\text{John})(c)$ is true iff John is in the positive extension of F in $g(c)$.

A traditional analysis of comparison, compatible with both approaches, translates (3a) to $\exists c, F(\text{John})(c) \& \neg F(\text{Mary})(c)$; i.e., John is healthier than Mary is true iff some degree exists, which John's health exceeds, but Mary's health does not (Schwarzchild 2008), or alternatively, some comparison class exists, relative to which John falls in the positive extension of *healthy*, but Mary does not (Klein 1980). Both approaches to gradability explain the noun-adjective distinction by postulating that nouns denote entity sets. This explains their incompatibility with most degree morphemes, illustrated by the oddness of, for example, #birder; #birdest; #very bird', and #too bird. However, a problem with this view of the noun-adjective distinction is that nouns can freely occur in between-predicate comparisons, as in (4c) above.

Existing analyses of between-predicate comparisons (see Morzycki 2011 and references therein) yield wrong predictions regarding other gradable constructions. The problem is that postulating either semantic gradability, or even only ad-hoc, contextual, meta-linguistic, last resort gradable interpretations for nouns to capture the meaning of between-noun comparisons, such as (4c), results in wrong predictions for, e.g., within-noun comparisons. For example, infelicitous structures such as #'This bird is more a duck than that one' or #'This bird is the most duck of all' are then predicted to be equally felicitous.

An alternative approach to the adjective-noun distinction considers the nature of the concepts they denote. Intuitively, people suppose that nouns like 'bird' denote object categories, while adjectives like 'red' denote properties. Psycholinguists employ this division, but do not explicate what exactly it amounts to. A related common view is that categorization under adjectives is a matter of a single dimension, such as height for 'tall', whereas nouns are multidimensional, e.g., categorization under 'bird' depends on dimensions such as has a bird-genotype, bird descendant, can interbreed with birds, winged, feathered and small. However, categorization under many adjectives depends on multiple dimensions, too. For example, 'healthy' and 'sick' consider dimensions such as flu, chickenpox, cholesterol, and sugar intake. A person may be 'healthy' in some respects, but not in others (Kamp 1975).

As we saw above, formal semanticists typically ignore the complexity of multiple dimensions, modeling all adjectives after 'tall' and 'long', i.e., in terms of a unique scale and degree relation. To illustrate, the degree analysis assigns to adjectives a type of a degree function, $\langle e, d \rangle$, or, as in derivation (5), a corresponding relation, $\langle e, \langle d, t \rangle \rangle$. These types are obviously too thin to encompass multiple degree functions or relations which simultaneously affect the interpretation of an adjective.^{1,2}

$$5) \quad \begin{array}{ccc} & \text{Tall}(j)(c_{\text{tall}}) & \Leftrightarrow \\ \text{John} & \diagup \quad \diagdown & \\ & \text{Is} & \text{Tall} \quad \lambda x. \lambda d. \text{Tall}(i)(c) \end{array}$$

In response to these challenges, this paper proposes a new approach to the semantic composition of sentences with adjectives, which captures multidimensional adjectives and the way they differ from nouns (consequences for modifier position fall outside the scope of this paper). The interpretation of statements with multidimensional adjectives, on the new approach, **necessitates quantification over dimensions**.

2. The proposed solution: Adjectives in the positive form

Let us assume that degree relations constitute dimensions and call their type, $\langle e, dt \rangle$, type f. Multiple degree relations in the interpretation of an adjective can be accommodated into the representation by analyzing multidimensional adjectives as dimensional generalized quantifiers, type $\langle ft, t \rangle$ (sets of sets of degree relations). Like entity quantifiers such as 'every boy' ($\langle et, t \rangle$), they have to move from their surface position to resolve type mismatch. They leave a trace of type f. Its value, F_F , combines with an entity and a degree argument, $F(j)(c_F)$, and is abstracted over at the clause level. The resulting clause interpretation is a set of relations. Thus, e.g., 'John is c_F F' denotes the set of degree relations F whose contextual norm John exceeds, $\lambda F. F(j)(c_F)$. For instance, this set includes the dimension cholesterol iff John's cholesterol level is sufficiently close to the ideal level in the context of evaluation. This clause interpretation, $\lambda F. F(j)(c_F)$, is of type ($\langle f, t \rangle$), which is precisely the right one to combine with the raised adjective, which is of type $\langle ft, t \rangle$.

The interpretation of a multidimensional adjective such as *healthy* as a dimensional quantifier makes use of a contextually determined **set of dimensions**. Thus, let $\lambda F. \text{Dim}(\text{healthy}, F)$ symbolize the set of contextually

¹ But see Kamp (1975) and Klein (1980) for discussions and proposed degree-less analyses.

² Analyses of positive forms 'X is F' often postulate a null morpheme to mediate the interpretation. The representation in (5) simplifies this aspect, assuming only a null variable c_F that can be read as either a comparison class or a degree parameter. When the variable remains free, the assignment function sets its value to the default comparison class or membership norm stored for F in the lexicon or calculated for it within context.

relevant health measurements or respects (cholesterol, fever, ...). Dimensional adjectives such as *long* can also be associated with a dimension set, $\lambda F.\text{Dim}(\text{long}, F)$, the set of length measurements, assuming in each context this set is a singleton.

The interpretation of an adjective type $\langle ft, t \rangle$, then, involves a relation between the set of contextually relevant dimensions, e.g., $\lambda F.\text{Dim}(\text{healthy}, F)$, and the set of dimensions which norms the entity argument exceeds, $\lambda F.F(\text{John})(c_F)$. For example, intuitively, the positive forms of multidimensional adjectives such as ‘*healthy*’ and ‘*sick*’ involve quantification over dimensions; e.g., (1a) conveys that John is healthy in all relevant respects, and (1b) conveys that he is sick in some respect. By contrast, (1c) does not naturally lend itself to an interpretation equivalent to quantificational paraphrases such as *Tweety is a duck in some/most/all respects* (Wittgenstein 1953).

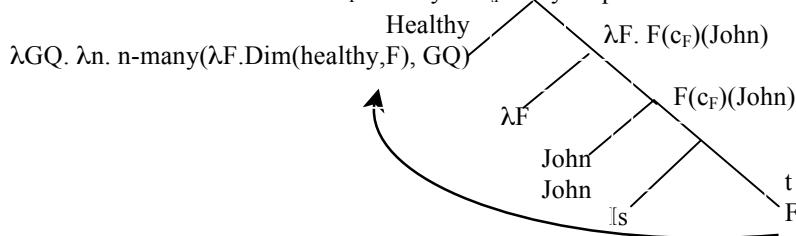
The above descriptions of the positive forms in (1) are motivated by empirical and experimental findings whereby grammatical operations can freely access the dimensions of adjectives and operate on them, whereas access to the dimensions of nouns is more restricted. Surveys of felicity judgments suggest that explicit dimensional quantifiers more naturally combine with adjectives than with nouns, as in *healthy/ safe/ clean in every respect* and *sick/ dangerous/ dirty in some respect(s)*, vs. the odd combinations *#duck/tree in every/some respect*.

Moreover, the felicity of sentences of the form ‘X is P’ with adjectives, but not nouns, correlates with that of their quantified equivalents ‘X is P in every/most/some respect(s)’ (Sassoon 2014a). Furthermore, surveys and corpus studies suggest that exception phrases, which distribution is restricted to universally quantified propositions, combine with adjectives (as in ‘*healthy except for cholesterol*’) and their negated antonyms (as in ‘not sick except for cholesterol’), but they combine neither with nouns, nor negated nouns (#’not a duck except for size/the beak’; Sassoon 2012, 2013a,b). These findings suggest that the interpretation of adjectives, but not nouns, appears to be mediated by dimensional quantifiers.

Thus, on the emerging view, (1a) translates to ‘John is healthy in every respect’, or more formally, $Q(\lambda F.\text{Dim}(\text{healthy}, F), \lambda F.F(\text{John})(c_F))$; e.g., if Q is a universal quantifier, this reads as ‘for all the dimensions F of *healthy*, John is c_F F’. Recall, however, that ‘every’ itself, and determiners more generally, can also be represented in terms of a relation between the cardinality of a set and a membership norm n, as they do in generalized quantifier theory (Barwise & Cooper 1981). This leads us to a representation as in $n\text{-many}(\lambda F.\text{Dim}(\text{healthy}, F), \lambda F.F(\text{John})(c_F))$, where n represents the number of adjectival dimensions whose norm John has to exceed for the sentence to count as true. Thus, *Healthy* translates to $\lambda GQ_{\ll e, dt \gg, t} \cdot \lambda n. n\text{-many}(\text{Dim}(\text{healthy}), GQ)$. This reduces to $\lambda GQ_{\ll e, dt \gg, t} \cdot \lambda n. |\text{Dim}(\text{healthy}) \cap GQ| \geq n$. The type of adjectives is, therefore, $\langle ft \langle d, t \rangle \rangle$.

The value of n may range from the total cardinality of the dimension set (yielding universal force) to 1 (existential force). In analogy with c_F , a free variable c_n (representing the number of dimensions which norm the entity argument is required to exceed) surfaces in the positive form. The assignment function sets the value of this variable to a default value for n stored in the semantics of the adjective (e.g., *healthy* is universal; *sick* is existential), or else, to a contextually selected value (as with *clever*; for empirical support see Sassoon 2012–2013a,b).

$$6) \quad [\lambda n. n\text{-many}(\lambda F.\text{Dim}(\text{healthy}, F), \lambda F.F(c_F)(\text{John}))](c_n) \\ \Leftrightarrow \text{John is } c_F \text{ healthy in } c_n \text{ many respects } F.$$



To wrap up, positive forms such as (1a) break down to the following two parts, illustrated in (6): (i) a clause with a dimensional trace applied to an entity argument (e.g., *John*) and abstracted over to denote the set of the entity’s gradable properties, e.g., $\lambda F.F(\text{John})(c_F)$, and (ii) an adjective that denotes a function from such a set to truth, or more precisely, a relation between such a set and a norm n. The generalized quantifier interpretation $\lambda F.F(\text{John})(c_F)$ saturates the generalized quantifier argument of the adjective to yield a truth value.³

³ Adjective movement within nominal structures has been postulated in the past (e.g., Kayne 1994 and Alexiadou & Wilder 1998), but no explanation was provided for why the AP cannot remain in its base position. We provide motivation: Movement leaves a trace of a dimension variable, thus creating the right type of argument for adjectives. But this means that movement must be postulated across the board, even in predicate position. An alternative way to approach the problem without movement is via type shifting of the entity argument

3. Adjectives and gradability

The relevance of multiple dimensions to the interpretation of gradability morphemes can be seen in the fact that the interpretation of comparisons of various types can be given paraphrases relating to dimensions. Examples (7a-b) include paraphrases for within-adjective comparisons with a dimensional and a multidimensional adjective, respectively. Examples (8a-b) include paraphrases for between-predicate comparisons with multidimensional adjectives and nouns, respectively.

- 7) a. The sofa is 2 centimeters longer than the table (is) \Leftrightarrow The difference between the degree of the sofa and the table in the dimension underlying entity classification as long vs. not long, f_{length} , equals twice the degree of a centimeter.
- b. (Generally) John is healthier than Bill \Leftrightarrow (Generally), the difference between the degrees of John and Bill in the dimensions underlying entity classification as healthy vs. not healthy in the context, $f_{\text{blood pressure}}, f_{\text{cholesterol}}, f_{\text{chickenpox}}, \dots$, exceeds zero.
- 8) a. These sofas are more similar than different \Leftrightarrow The number of dimensions along which the two sofas classify as similar exceeds the number of dimensions along which they classify as different.
- b. John is more a linguist than a philosopher \Leftrightarrow The percentage of dimensions of a linguist along which John classifies positively exceeds the percentage of dimensions of a philosopher along which John classifies positively.

These paraphrases suggest that comparison morphemes involve quantification over or counting of dimensions; e.g., it follows from (7a,b) that a degree difference in at least SOME dimensions of an adjective should be present for within adjective comparisons to hold true. In positive multidimensional adjectives such as ‘healthy’, the requirement might be stronger such that a degree difference should generally be present in ALL or MOST of the dimensions. Furthermore, it follows from (8b) that a larger PERCENTAGE of dimensions of a noun in comparison to another noun should be observed for the between-noun comparison to hold true of an entity (see Sassoon 2014b for discussion).

The proposed account extends naturally to gradable constructions; e.g., (3a) breaks down to an adjective and the rest of the sentence, which in this case amounts to $\lambda F. \exists c, F(\text{John})(c) \& \neg F(\text{Mary})(c)$ (Klein 1980; Schwarzschild 2008), we get that (3a) translates to $[n\text{-many}(\lambda F. \text{Dim}(\text{healthy}, F), \lambda F. \exists c, F(\text{John})(c) \& \neg F(\text{Mary})(c))] (c_n)$, i.e., $|\lambda F. \text{Dim}(\text{healthy}, F) \cap \lambda F. \exists c, F(\text{John})(c) \& \neg F(\text{Mary})(c)| \geq c_n$. In words, for n many (all) dimensions F of *healthy*, John is more F than Mary.

Preliminary corpus evidence suggests that the default force of quantifier over dimensions, modeled here through the value of c_n , is inherited from adjectives like *healthy* to their comparative form *healthier*. Evidence for such readings is formed by frequencies of occurrence of explicit quantifiers over comparative dimensions, as in, *John is healthier {in every/some respects; except w.r.t. cholesterol}* (Sassoon 2013a).

In terms of semantic composition, the comparison morpheme can either take two entity type arguments, or entity arguments type shifted to a set of relational properties. In the latter case, illustrated in (9a), the null parameter c of these properties is bound by the comparative morpheme.

But this is not the only reading derived. In (3a), *more* can either operate on each dimension F separately (as in (9a)) or on *many*, yielding that ‘John is healthy in more respects than Mary’, $\exists n$, such that $n\text{-Many}(\lambda F. \text{Dim}(\text{healthy}, F), \lambda F. F(\text{John})(c_F)) \& \neg n\text{-Many}(\lambda F. \text{Dim}(\text{healthy}, F), \lambda F. F(\text{Mary})(c_F))$. In words, there is a number n, such that John is c_F healthy in n many respects F, but Mary does not. As illustrated in (8b), this reading involves an unsaturated standard variable n in the interpretation of the adjective.

In (4b), illustrated in (10), *more* analyzed as the comparative of *many* yields that for some n, the sofas are similar in n-many respects, but not dissimilar in n-many respects, meaning that they are more similar than dissimilar.

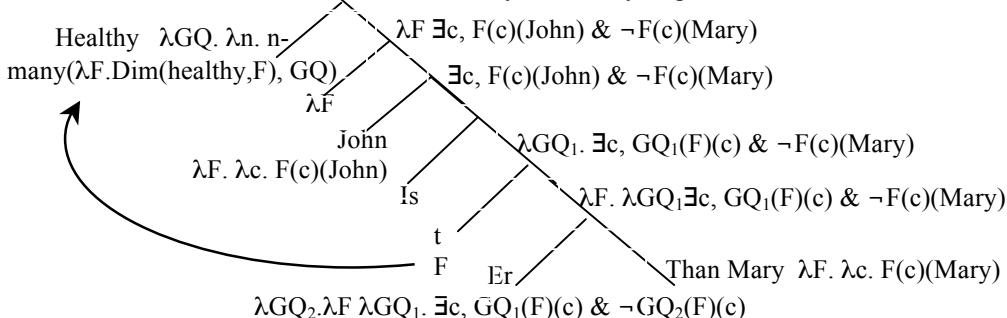
On the emerging view, *more* is a cross categorical morpheme, just like negation and conjunction. It combines with any set of arguments of any types that can combine to the form $\exists a, R_1(a, b_1) \& \neg R_2(a, b_2)$ and be interpreted consistently. To this end, the arguments can comprise of one relation R and two entities b_1 and b_2 , or to one argument b and two relations R_1 and R_2 . R can be an adjective type $\langle d, \langle f, t \rangle \rangle$ whose degree argument would be bound existentially and then it would be fed by two dimensional quantifier type $\langle f, t \rangle$, as in (9b). Alternatively, R can be of a dimension type f whose degree argument is bound and then it applies to two entity arguments, or two entity quantifiers apply to it, as in (9a). R can also be a relation over noun meanings as in ‘more boys than girls arrived’, which can be similarly represented as $\exists n, n\text{-many}(b, \text{arrived}) \& \neg n\text{-many}(g, \text{arrived})$.

(e.g., *John*) to a set of dimensions, $\lambda F. F(\text{John})(c_F)$, which can then combine directly with an adjective interpretation insitu. Grammar may, therefore, allow both options.

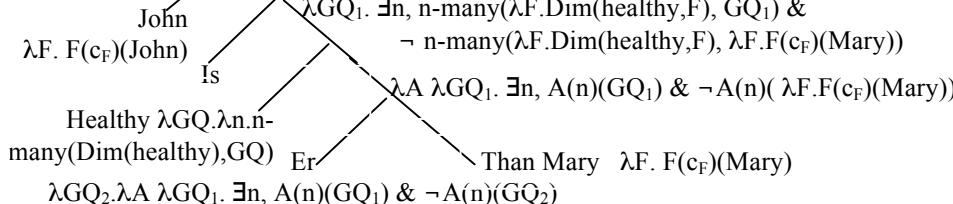
Replacing *longer* with *more long* in (4a) yields that the sofa is c_F long in more respects F than the table is c_F wide, i.e., $\exists n, n\text{-many}(\lambda F.\text{Dim(long},F), \lambda F. F(c_F)(\text{the sofa})) \& \neg n\text{-many}(\lambda F.\text{Dim(wide},F), \lambda F. F(c_F)(\text{the table}))$. Since the two adjectives in this example are one dimensional, this reduces to the requirement that the sofa be c_{long} long, and the table not be c_{wide} wide. But precisely this message is conveyed by the simpler positive forms: *The table is long and the sofa is not wide*. Thus, decrease in felicity is correctly predicted.⁴

The ban on comparison of number of dimensions which norm an entity exceeds given two dimensional adjectives (as in, e.g., *more long than wide*) can also be formulated as a more general restriction on the distribution of ‘more’ to scales comprising more than two degrees (Frank Veltman, p.c.). The scale from which the value of n is drawn in the case of a dimensional adjective comprises of only two values, 0 and 1, unlike the case of a multidimensional adjective, in which it normally comprises of many more values. This ban explains unavailability of the first reading of *more* (as in 9a) with predicates whose dimensions are not gradable (*prime*, *triangle*).

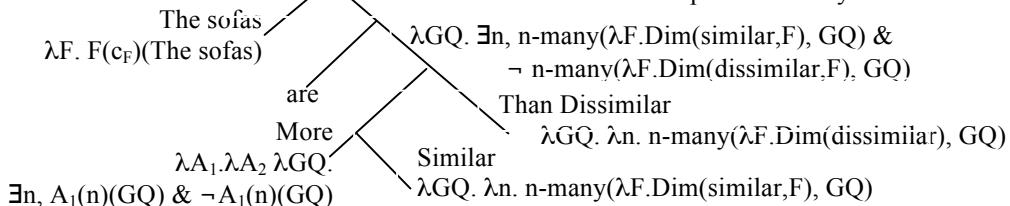
- 9) a. $[\lambda n. n\text{-many}(\lambda F.\text{Dim(healthy},F), \lambda F. \exists c, F(c)(\text{John}) \& \neg F(c)(\text{Mary}))](c_n)$
 $\Leftrightarrow \text{John is healthier than Mary in } c_n \text{ many respects F.}$



- b. $\exists n, n\text{-many}(\lambda F.\text{Dim(healthy},F), \lambda F. F(c_F)(\text{John})) \&$
 $\neg n\text{-many}(\lambda F.\text{Dim(healthy},F), \lambda F. F(c_F)(\text{Mary}))$
 $\Leftrightarrow \text{John is healthy in more respects F than Mary is.}$



- 10) $\exists n, n\text{-many}(\lambda F.\text{Dim(similar},F), \lambda F. F(c_F)(\text{the sofas})) \&$
 $\neg n\text{-many}(\lambda F.\text{Dim(dissimilar},F), \lambda F. F(c_F)(\text{the sofas})) \Leftrightarrow$
 $\text{The sofas are similar in more respects than they are dissimilar.}$



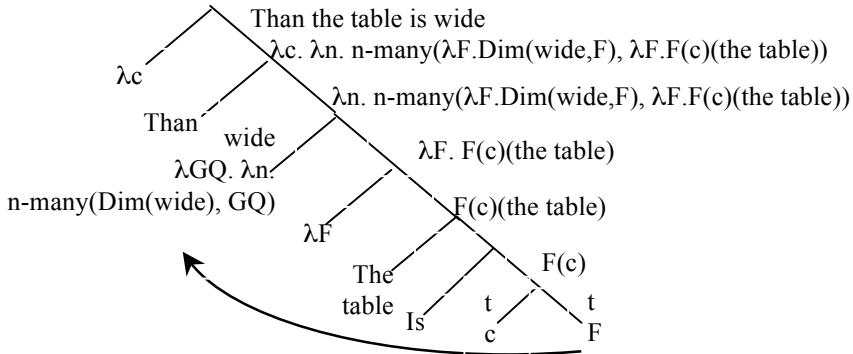
The comparison morphemes in the examples we considered so far either bind the standard of the dimension counter *many* (like *more* in (9b-10) or the dimensional standards c_F (like *-er* in (9a)). However, direct-comparison morphemes like *-er* in (4a), which surface a measure phrase argument, appear to bind both standard variables. An extension of the traditional account of comparison is required to capture their meaning. Comparison can be viewed as involving a set $\lambda a, R_1(a,b_1) \& \neg R_2(a,b_2)$, which is not necessarily bound existentially (Schwarzschild 2008).

⁴ To discern the interpretation from that of the conjunction of positive forms, it is inferred that the latter is false. To accommodate this, a degree lower than c_{long} must be used as the value of the contextual standard, for the sofa to only be close to being long, yet longer than the table is wide.

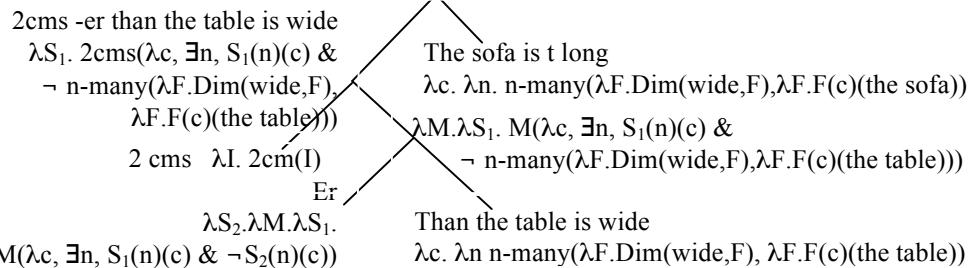
For example, (4a) may translate to $2\text{CM}(\lambda c, \exists F \in \text{DIM}(\text{long}), F(\text{the sofa})(c) \ \& \ \neg \exists F \in \{\text{DIM}(\text{wide}), F(\text{the table})(c)\})$, i.e., the property of being 2 centimeters long holds of the interval comprising of the set of degrees c such that for some dimension F of ‘long’ the sofa is c long, but for no dimension F of wide the sofa is c wide. Thus, we get for (4a) that the set of degrees c —such that the sofa is c long, but the table is not c wide—stretches along an interval of 2 centimeters, as desired (Kennedy 1999).

The compositional semantics for the than-clause in (11a) and matrix clause in (11b) is compatible with a standard analyses of *er* as a determiner over degree predicates, whereby a silent WH operator moves up within the than-clause leaving a degree trace to be bound by a lambda operator as in (11a). The matrix clause is analyzed as a degree predicate based on the assumption that the whole *er*-phrase moves at LF, leaving a degree trace to be bound by a lambda operator.

11) a.



b. $2\text{cms}(\lambda c, \exists n, n\text{-many}(\lambda F.\text{Dim}(\text{wide},F), \lambda F.F(c)(\text{the sofa})) \ \& \ \neg n\text{-many}(\lambda F.\text{Dim}(\text{wide},F), \lambda F.F(c)(\text{the table})))$



4. Nouns and gradability

Nouns denote sets of entities. Yet, the role dimensions play in categorization under nouns is a main predictor of their (in)felicity in gradable constructions, as follows.

On psychological similarity analyses, entities classify under nouns iff their values on multiple dimensions sufficiently match the ideal values for the noun. The degree of an entity in a given noun is built by addition or multiplication of its degrees in multiple dimensions. The resulting weighted sum or product should exceed a membership standard. Interestingly, additive classification characterizes mostly social concepts (artifacts and human traits), whose dimensions are relatively independent (Hampton et al. 2009). For instance, typically, a ‘linguist’ works in linguistics departments, investigates languages, and reads Chomsky’s work, but a person violating some of these features may still count as a linguist. By contrast, multiplicative classification characterizes natural kind concepts (plants and animals), where the dimensional values always go together (Hampton et al. 2009); e.g., an entity halfway between a zebra and a donkey classifies as neither. Only small deviations from the ideal are allowed.

To appreciate this distinction, consider a set of binary dimensions with equal weights. For entities that match the ideal in all of them, 1 ... 1, except for one mismatch, 0, addition gives $0 + 1 + \dots + 1 \gg 0$, while multiplication gives $0 \times 1 \times \dots \times 1 = 0$. Thus, with binary dimensions of equal weights, additive, but not multiplicative classification is rendered equivalent to quantification (mere counting of dimensions). Entities are required to have sufficiently many (all/ most/ some) of the dimensions.

Since adjectives favor quantification, additive nouns are predicted to be judged more felicitous than multiplicative ones in adjective-selecting linguistic constructions. This prediction is borne out (Sassoon 2014a). Findings suggest that domain (additive vs. multiplicative) is a main predictor of noun felicity in various

constructions, including within and between predicate comparisons. Hence, grammar is sensitive to conceptual structure. In particular, comparison morphemes are acceptable with a noun to the extent to which its interpretation can approximate that of an adjective, namely be modeled by means of quantification over dimensions.

The general moral with regard to the noun-adjective distinction is that the interpretation of the former, but not the latter involves quantification over dimensions. Thus, most gradability morphemes, whose interpretations involve quantification over dimensions as well, freely select adjectives, but not nouns.⁵

References

- Alexiadou, A., & Chris W. (1998). Adjectival modification and multiple determiners. In Artemis Alexiadou and Chris Wilder (Eds.) *Possessors, predicates and movement in the determiner phrase*: 303–332. Amsterdam: John Benjamins.
- Barwise, J. & Cooper, R. (1981). Generalized quantifiers and natural language, *Linguistics and Philosophy* 4: 159–219.
- Hampton, A. James, Gert Storms, Claire L. Simmons and Daniel Heussen (2009). Feature integration in natural language concepts, *Memory and cognition* 37(8): 1150-1163.
- Heim, I. & Kratzer, A. (1998). *Semantics in Generative Grammar*. Blackwell, Oxford.
- Kamp, H. (1975). Two theories about adjectives. In Keenan, E (ed.) *Formal Semantics of Natural Language*. Cambridge Uni. Press, England;
- Kayne, R. (1994). The antisymmetry of syntax. Cambridge, MA: MIT Press.
- Kennedy, C. (1999). *Projecting the adjective: the syntax and semantics of gradability and comparison*. Garland, NY;
- Klein, E. (1980). A semantics for positive and comparative adjectives. *Linguistics & philosophy* 4(1);
- Landman, F. (2000). Events and Plurality. Dordrecht: Kluwer, 143-164.
- Morzycki, M. (2011). ‘Metalinguistic comparison in an alternative semantics for imprecision’. *Natural Language Semantics* 19(1), 39–86.
- Sassoon, G. W. (2012). The double nature of negative antonymy. *Proceedings of Sinn und Bedeutung* 16: 543-556, Utrecht, Netherlands.
- Sassoon, G. W. (2013a). A typology of multidimensional adjectives. *Journal of Semantics* 30: 335–380.
- Sassoon G. W. (2013b). *Vagueness, Gradability, and Typicality, The Interpretation of Adjectives and Nouns*, CRISPI, Brill: Leiden.
- Sassoon, G. W. (2014a). Dimension accessibility as a predictor of morphological gradability. Paper in progress based on a talk for the workshop on concept composition 1-2/09/2013, Utrecht. Bar Ilan University.
- Sassoon, G. W. (2014b). Between noun comparisons. To appear in *Selected papers from the 10th TbiLLC conference*, Lecture Notes in Artificial Intelligence, Springer.
- Schwarzschild, R. (2008). The semantics of comparatives and other degree constructions. *Language and Linguistics Compass* 2.2: 308–331.
- von Stechow, A. (1984). Comparing semantic theories of comparison. *Journal of Semantics* 3, 1–77.
- Wittgenstein, L. (1953). *Philosophical Investigations*. Translated by G. E. M. Anscombe. Blackwell. Oxford.

⁵ Quantifying modifiers, such as *perfectly*, *slightly*, and *in {every, some} respects*, may bind n or c as well; e.g., x may be *perfectly healthy* if x is c_F F in every respect F, or if in c_n respects F, x is F to every degree c. This gives additional motivation for the richer adjectival structure proposed, to be studied in the future.

Distributivity and agreement: new evidence for groups as sets

Hanna de Vries

Utrecht University, Utrecht, The Netherlands
h.devries1@uu.nl

1 Introduction

It is a well-known observation that morphologically singular group NPs in British English can occur with either a singular or a plural VP (example (1) is from (Jespersen, 1914)):

1. a. Mine is an old family.
b. My family are early risers.
2. a. The hotel staff is friendly.
b. The hotel staff are friendly.

This phenomenon has attracted some attention from formal semanticists, who have mainly investigated the way semantics influences agreement choice by attempting to relate the preference for singular or plural to various lexical properties of the predicate (e.g. Bennett, 1974; Barker, 1992; Pollard and Sag, 1994; Schwarzschild, 1996). However, the reverse - whether agreement choice influences semantics - has rarely been investigated.

In this paper, I argue that there are systematic differences in interpretation between British English group noun sentences with a singular VP and their plural-VP counterparts, and that these differences tell us something about the set-theoretic structure of group noun denotations. First, I show how the availability of certain distributive interpretations can be used as a diagnostic for the semantic plurality of an NP. Then, I will show that these distributive interpretations are available with morphologically singular group NPs, but only if the VP is plural. This central observation is exemplified by contrasts like the following:¹

3. The semanticists are singing or dancing.
4. a. The group (of semanticists) is singing or dancing.
b. The group (of semanticists) are singing or dancing.

Sentence (3) is compatible with a ‘split’ situation in which part of the semanticists are singing and the rest are dancing, which is the interpretation we get when we apply the disjunction to each individual member of the plurality *the semanticists*. In contrast, sentence (4a) lacks such a distributive interpretation: it is only compatible with a situation in which the entire team is doing the same thing. This contrast can easily be explained under the common assumption (e.g Barker, 1992; Schwarzschild, 1996) that group NPs denote atomic entities, whose subparts are inaccessible to the compositional semantics. Without access to individual group members, the disjunction can only apply to the group as a whole.

¹The contrast in this particular example, which is responsible for inspiring the line of research developed in this paper, was originally pointed out to me by Hazel Pearson and Michelle Sheehan.

However, this analysis is complicated by the fact that the split interpretation is available in British English when the VP is plural, as in (4b). According to our reasoning above, this would suggest that *the group* in (4b) does not denote an atomic entity, but a set, like *the semanticists* in (3). The way we solve this puzzle has consequences for semantic theory in general: do we allow for language-specific assumptions to deal with this cross-linguistic variation (strategies like this have been proposed for independent reasons: Schwarzschild (1996) suggests that group nouns might be systematically ambiguous in British but not in American English, and Sauerland and Elbourne (2002) propose that British English group nouns are semantically plural while their American counterparts are semantically singular), or can we analyse the above facts while maintaining a uniform semantic analysis of group nouns across languages?

In this paper, I propose to solve the puzzle in the following way. We will assume that group nouns basically range over sets in all varieties of English; then, we derive the atomic interpretation by means of a typeshift that is triggered by a type mismatch between the NP (which denotes a set) and a singular VP (which ranges over atoms). Very few of these assumptions are new. Set-based analyses of group nouns are unusual (Bennett, 1974; Pearson, 2011, are two rare examples), but the required typeshifting operation (which maps a set into an ‘impure’ atom) has been present in the literature since Link (1984); Landman (1989); similarly, the idea that morphologically plural VPs range over sets while morphologically singular NPs range over atoms is quite common and goes back to at least Bennett (1974).²

There are two main advantages to this approach. First, there is no need for any language-specific semantic machinery: group nouns are uniformly analysed as sets, and any interpretational differences between languages follow from their morphosyntactic properties. Second, it allows us to analyse the interpretation of group nouns using only general semantic mechanisms; there is no need for an additional typeshifting operation that ‘breaks up’ groups, as in e.g. Landman (1989); Barker (1992); Schwarzschild (1996); Sauerland (2004).

2 Background: quantificational distributivity

2.1 Two kinds of distributivity

I follow researchers like Roberts (1978); Hoeksema (1988); Winter (1997); Champollion (2010) in assuming that distributive interpretations are sometimes derived lexically and sometimes structurally, by means of a (covert) operator. The sentence in (5a) is an example of the first kind of distributivity, which I will call *P-distributivity* (following Winter 1997). The meaning of (5a) can be derived by simple predication over a plural individual, without the assumption of further covert semantic machinery: information about the way individual members of *the children* participate in the laughing event is not formally specified, but inferred on the basis of the lexical semantics of the predicate. (We know that in order to be able to laugh one needs lungs and a vocal apparatus, and we know that individual children have this but groups as a whole do not.)

- 5. a. The children laughed.
- b. The children are hiding somewhere.

On the other hand, the distributive interpretation of (5b) - according to which each of the children is hiding in a potentially different place - cannot be derived in such a way. *Somewhere*

²In this paper, I approach plurality in set-theoretical terms, but everything should carry over straightforwardly to a Link-style lattice-theoretical framework.

is a quantifier, which in the absence of other quantificational elements takes scope over the entire sentence; this derives only the collective interpretation according to which the children are all hiding in the same spot. To derive the distributive interpretation, we need to resort to some kind of covert quantificational mechanism comparable to overt *each*. I will call this kind of distributivity *Q-distributivity* (again following Winter).

2.2 Q-distributivity as a diagnostic for semantic number

The distinction between P- and Q-distributivity is relevant because the first, but not the second, is available with singular group noun subjects (when the VP is also singular). Compare the possible interpretations of (6a-b) with those in (5a-b):

6. a. The cricket team laughed.
- b. The cricket team is hiding somewhere.

While we can easily interpret (6a) distributively, the Q-distributive interpretation for (6b) - according to which each of the cricket players has their own hiding spot - is unavailable. The only interpretation for (6b) is the one according to which everyone is hiding in the same place.

We find the same contrast between plural definites and group NPs with other examples of Q-distributivity:

7. a. The members of the Jones family are blond or red-haired.
- b. The Jones family is blond or red-haired.
8. a. The children are drawing and sleeping.
- b. # The class is drawing and sleeping.

Sentence (7a) is compatible with a situation in which some of the Jones are blond while the others are redheads; for sentence (7b) to be true, however, all the Joneses need to have the same hair colour. Similarly, sentence (8a) but not (8b) is true in a situation where half of the children are drawing and the other half are sleeping (sentence (8b) makes the weird claim that all children are simultaneously doing both).³

As already noted in the introduction, these data are not surprising if we assume that group NPs denote atomic entities (following e.g. Landman, 1989; Barker, 1992; Schwarzschild, 1996). Q-distributivity is a quantificational mechanism, so it only works if it has a set to quantify over; if plural definites denote sets but group NPs do not, the observed contrast follows straightforwardly. This means that we can use the availability of Q-distributive interpretations as a diagnostic for the semantic plurality of an NP: if such an interpretation is available, the NP in question should be analysed as a set, and if not, it should be analysed as an atom.

With this idea in mind, it is time to take another look at British English.

3 Q-distributivity with group nouns in British English

3.1 Group nouns and agreement in British English

There does not seem to be a straightforward rule that governs when it is more appropriate to use a singular or a plural VP with a particular group subject - speakers' preferences for singular

³I am not sure how to analyse the semantics of sentences like (8a) that allow these non-Boolean split interpretations. I classify them as Q-distributive here because they pattern with the other Q-distributivity tests in terms of the plural/group contrast.

or plural vary depending on both the group noun and the predicate (Corbett, 2000). Jespersen (1914) notes that the choice of agreement sometimes reflects whether we conceive of a particular group as a single entity (*family* in (1a)) or as a collection of individuals (*family* in (1b)), but this is not necessarily the case: a simple Google search confirms Jespersen's observation that both collective and distributive predicates may take either singular or plural form when they occur with a group subject.

However, the choice of singular or plural is not always arbitrary. Pollard and Sag (1994) mention that predicates that express properties of the groups themselves never trigger plural agreement, as in (9) (from Pollard & Sag 1994):

9. *A new committee have been constituted.

Being constituted is a property of the committee as an institution, completely independent from any properties of its members; predicates like these obligatorily appear in the singular. A related observation is made by Barker (1992): unlike (10a), which can mean that the committee is an old institution, (10b) (where the group noun appears with a plural VP) can only mean that the members of the committee are old.⁴

10. a. The committee is old.
- b. The committee are old.

The number of the VP also has more structural semantic consequences, that are not related to the lexical meaning of the predicate itself. As observed by Sauerland and Elbourne (2002), whether certain British English sentences with a group noun subject display scope ambiguity depends on the number of the VP. Sentence (11a) - with a singular VP - has both the surface scope and the inverse scope reading, but sentence (11b) - with a plural VP - only has the former.

11. a. A northern team is likely to be in the final. $(\exists > \text{likely}, \text{likely} > \exists)$
- b. A northern team are likely to be in the final. $(\exists > \text{likely}, * \text{likely} > \exists)$

These interesting data suggest that there is a connection between the semantic structure of a group NP and the number of the VP it agrees with: an NP that agrees in the singular apparently has different scope properties than an NP that agrees in the plural. As we will see, the Q-distributivity facts for British English are in line with this.

3.2 Q-distributivity in British English

In this section, I present this paper's main data, which show that the number of the VP influences the availability of Q-distributivity for British English sentences with a group subject. Since the judgements are sometimes subtle, I will support my claims with quantitative evidence obtained from a small group of native BE speakers.⁵

⁴Barker also claims that (10a) does not have the 'old members' interpretation, but this appears to be too strong a claim: according to my informants, sentence (10a) can have both interpretations, just as it does in other varieties of English.

⁵Judgements were obtained from 6 native speakers from England and Wales by means of a pen-and-paper truth value judgement task. Each test item consisted of a sentence and a picture of a situation that was compatible with a distributive interpretation of that sentence, but incompatible with a collective one; subjects were asked to judge whether the sentence could be true in the depicted situation. The test items (18 in total) were balanced out by fillers. A few days after filling out the first questionnaire, the informants were asked to fill out a second version, with the order of the items reversed. In total, 23% of the singular-VP group-subject sentences was judged true, against 61% of their plural-VP counterparts and 83% of the sentences with a plural subject.

In section 2.2, I have given three different examples of Q-distributive sentences: one involving disjunction, one conjunction, and one involving another quantificational element. All three cases showed the same contrast between plural definites and group NPs: while sentences with the former could all receive a Q-distributive interpretation, that interpretation was absent when the subject was a group NP (with a singular VP). Below, I give the same (or very similar) examples, but this time I also compare group noun sentences with a singular VP with their plural-VP counterparts. The percentages indicate the number of ‘true’ judgements by my informants, who were asked to evaluate the truth of these sentences in a ‘distributive’ situation (for example, the picture that accompanied the sentences in (12) showed several different hiding places with a child hiding in each of them).

- | | | |
|-----|---|---------|
| 12. | a. The children are hiding somewhere. | 100% |
| | b. The class is/are hiding somewhere. | 40%/83% |
| 13. | a. The Joneses are very short or very tall. | 100% |
| | b. The Jones family is/are very short or very tall. | 17%/40% |
| 14. | a. The children are drawing and sleeping. | 75% |
| | b. The class is/are drawing and sleeping. | 40%/83% |

The variation between informants and between structurally indistinguishable sentences suggests that people’s judgements are influenced by many factors that are unrelated to the formal semantics of these sentences - for example, in both the disjunction and the conjunction cases, the group noun sentences were judged ‘false’ much more often when the coordinated predicates were adjectives (as in 13b) than when they were verbs (as in 14b). Even so, a clear pattern emerges from this small quantitative study: while British English group noun sentences with a singular VP behave as they would in American English, group noun sentences with a plural VP behave similar to plural-subject sentences, and may receive a Q-distributive interpretation. Combined with our earlier observation that the availability of Q-distributivity serves as a diagnostic for semantic number, this leads us to conclude that singular group NPs receive an atomic denotation when they agree with a singular VP, but a set denotation when they agree with a plural VP.

3.3 A related observation: reciprocity

The observed pattern is also found with reciprocal predication in British English, which should not come as a surprise given that reciprocity, like Q-distributivity, is usually analysed as a form of quantification and hence requires a semantically plural argument. It has been noted in the literature that reciprocal expressions are fine with plurals, but at best marginal with group nouns (Schwarzschild, 1996; Lønning, 2011):

15. a. The cricket players are friends / usually coach each other.
- b. *The cricket team is friends / usually coaches each other.

However, Schwarzschild hypothesises that reciprocal predication over a group subject may be grammatical with British English when the VP is plural, a suggestion supported by the data in (16) (from Pearson, 2011). Many examples of these sentences ‘in the wild’ can be found with Google (17a-d):

16. The family can’t stand each other.

- a. The Team are friends on track as well as off track, and are as much family as we are friends.
- b. Can a scientific program really change the way the Diaz family love each other?
- c. It is puzzling when medical staff disagree with each other.
- d. Remember that your group are neighbours who have to get along outside the group as well as within it.

Like the Q-distributivity data, these reciprocity data are in line with the idea that the semantic number of a morphosyntactically singular group NP depends on the number of the VP it agrees with.

4 Analysis: group NPs as sets

In principle, there are a couple of ways to account for the observed ‘mixed’ behaviour of group NPs. We could say that British English group nouns are ambiguous between set and atom predicates (cf. Schwarzschild, 1996). Alternatively, we could take one of the denotations as basic and derive the other one by a typeshift of some kind (cf. Landman, 1989; Barker, 1992; Schwarzschild, 1996; Sauerland, 2004). I will propose an analysis along the second lines; however, in contrast to previous accounts, I will take the set denotation to be basic and the atom denotation to be derived. (We will see an empirical advantage of this in section 4.1.)

Apart from the assumption that group nouns basically range over sets, we will adopt the following two common assumptions from the literature. First, that syntactically plural predicates range over sets (cf. Bennett, 1974; Link, 1983; Winter, 2002, and many others); I will follow Link in assuming that these semantically plural predicates are derived from their singular, *et*-type counterparts by a pluralisation operation $*$ (where $*P$ is defined as $\wp(P) - \emptyset$). Second, that set-denoting referential NPs can be mapped onto a corresponding ‘impure’ atom (cf. Link, 1984; Landman, 1989; Winter, 2002) by a type-shifting operation that we will write as \uparrow (following Landman 1989).⁶ Together, these assumptions present a problem that every language will have to solve in some way: with ‘correct’ agreement (a singular VP to match the singular group NP), it is impossible to say anything meaningful about groups, as group NPs denote sets but singular predicates range over atoms; in other words, we always end up with a type mismatch, as exemplified in (17).

17. $\llbracket \text{The group is tall} \rrbracket = \mathbf{tall}_{et}(\mathbf{the_group}_{et})$

There are two ways to resolve this type mismatch. The semantic way, which I propose is always available, is to shift the group NP into its corresponding impure atom (18a); the morphosyntactic way, which is available in British but not in American English, is to allow the group NP to occur with a plural VP (18b):

18. a. $\llbracket \text{The group is tall} \rrbracket = \mathbf{tall}_{et}(\uparrow(\mathbf{the_group})_e)$
b. $\llbracket \text{The group are tall} \rrbracket = * \mathbf{tall}_{et,t}(\mathbf{the_group}_{et})$

From this system, it follows that group NPs behave like sets when they occur with a plural VP, but like atoms when they occur with a singular VP. To further illustrate how this accounts for the observed pattern with Q-distributivity, let’s have a look at an example:

⁶Even though I adopt Landman’s notation, I am not committed to his or any other particular formalisation of impure atom formation.

19. a. $\llbracket \text{the group is singing or dancing} \rrbracket = 1$ iff $\uparrow(\text{the_group}) \in \text{sing} \cup \text{dance}$
 $\Leftrightarrow \uparrow(\text{the_group}) \in \text{sing}$ or $\uparrow(\text{the_group}) \in \text{dance}$
b. $\llbracket \text{the group are singing or dancing} \rrbracket = 1$ iff $\text{the_group} \in *(\text{sing} \cup \text{dance})$
 \Leftrightarrow for all $x \in \text{the_group}$: $x \in \text{sing}$ or $x \in \text{dance}$

In (19a), where *the group* receives an atomic interpretation, the sentence is truth-conditionally equivalent to ‘the group is singing or the group is dancing’, which is incompatible with a split situation. In (19b), *the group* is interpreted as a set, and the pluralised predicate *singing or dancing* denotes the set of all sets whose members are either singing or dancing; because of the way $*$ is defined, the set **the_group** is in the extension of this predicate just in case each of its members is either singing or dancing. This is the distributive interpretation, compatible with a split situation in which part of the group is singing and the rest is dancing.

Along similar lines, we can now explain our earlier data from Barker (1992) and Pollard and Sag (1994) ((9) and (10) from section 3.1), who observed that group-level predicates (like *be constituted* and *be old* when the age of the group is intended) are incompatible with plural morphology. Under our assumptions, the semantics of *The committee are old* is analysed as $*\text{old}(\text{the_committee})$. For this to be true, each individual member of **the_committee** needs to be in the extension of (unstarred) **old**; in other words, *The committee are old* is true just in case the committee’s members are old. Similarly, *The committee were constituted in 2001* requires, nonsensically, that each individual member of the committee was constituted in 2001. On the other hand, *The committee is old* and *The committee was constituted in 2001* require that the impure atom corresponding to **the_committee** be in the extension of **old** or **constituted_in_2001**. In the first case, we may still draw a P-distributive inference about the individual members of the group, which is in line with our observation that *The committee is old* can have both an old-members and an old-committee interpretation.

4.1 Additional support: *there-sentences in colloquial English*

One of the main distinctions between the present analysis and its two main alternatives (ambiguity and a typeshift the other way around) is that the present analysis has general applications beyond the domain of group NPs. Whether this is an advantage or not depends on whether the analysis makes the right predictions in those cases. There seems to be at least one phenomenon in English that suggests that it does. In colloquial English, many speakers accept *there-sentences* with mismatched agreement, as in (20b):

20. a. There are two semanticists singing or dancing in my garden.
b. There’s two semanticists singing or dancing in my garden.

According to most of my informants, these sentences show a contrast similar to the British English Q-distributivity cases: while (20a) is compatible with a split situation, (20b) is not. It seems that our present analysis can account for this in a straightforward way: in (20b), the type mismatch between the singular VP and the plural NP coerces the latter to shift into an impure atom, which rules out a Q-distributive interpretation. Since these cases involve ‘ordinary’ plurals rather than group nouns, neither of the alternative analyses can be extended in a way that covers them.

5 Conclusions

I have shown that morphologically singular group NPs in British English behave like atoms when they occur with a singular VP, but like sets when they occur with a plural VP. If we assume

that group NPs are basically set-denoting, their behaviour with different kinds of agreement follows from common assumptions about the semantics of number morphology. The advantage of this proposal is that it reduces crosslinguistic variations in interpretation to morphosyntactic differences between languages, allowing us to maintain a uniform semantics for group nouns. It also has the empirical advantage of being applicable to a much wider range of semantic phenomena involving number mismatches - our next step should be to look beyond English and see if the analysis holds up under further typological scrutiny.

References

- Barker, Chris. 1992. Group terms in English: Representing groups as atoms. *Journal of Semantics* 9:69–93.
- Bennett, Michael. 1974. Some extensions of a Montague fragment of english. Doctoral Dissertation, UCLA.
- Champollion, Lucas. 2010. Parts of a whole: Distributivity as a bridge between aspect and measurement. Doctoral Dissertation, UPenn.
- Corbett, Greville G. 2000. *Number*. New York: Cambridge University Press.
- Hoeksema, Jack. 1988. The semantics of non-Boolean ‘and’. *Journal of Semantics* 6:19–40.
- Jespersen, Otto. 1914. *A modern English grammar on historical principles, part II, volume 1*.
- Landman, Fred. 1989. Groups. *Linguistics and Philosophy* 12:559–605, 723–744.
- Link, Godehard. 1983. The logical analysis of plurals and mass terms: a lattice-theoretical approach. In *Meaning, use and interpretation of language*, ed. Rainer Bauerle, Christoph Schwarze, and Arnim von Stechow. Berlin: De Gruyter.
- Link, Godehard. 1984. Hydras. On the logic of relative constructions with multiple heads. In *Varieties of formal semantics*, ed. Fred Landman and Frank Veltman. Dordrecht: Foris.
- Lønning, Jan Tore. 2011. Plurals and collectives. In *Handbook of logic and language (second edition)*, ed. Johan van Benthem and Alice ter Meulen. Elsevier.
- Pearson, Hazel. 2011. A new semantics for group nouns. In *Proceedings of WECOL 28*.
- Pollard, Carl, and Ivan Sag. 1994. *Head-driven phrase structure grammar*.
- Roberts, Craige. 1978. Modal subordination, anaphora, and distributivity. Doctoral Dissertation, UMass.
- Sauerland, Uli. 2004. ‘a team’, definitely. *Snippets* 9. URL <http://ledonline.it/snippets/>.
- Sauerland, Uli, and Paul Elbourne. 2002. Total reconstruction, PF movement, and derivational order. *Linguistic Inquiry* 33:283–319.
- Schwarzchild, Roger. 1996. *Pluralities*. Dordrecht: Kluwer.
- Winter, Yoad. 1997. Choice functions and the scopal semantics of indefinites. *Linguistics and Philosophy* 20:399–467.
- Winter, Yoad. 2002. Atoms and sets: a characterization of semantic number. *Linguistic Inquiry* 33:493–505.

Indicative Scorekeeping

Malte Willer

University of Chicago
willer@uchicago.edu

Abstract

Folklore has it that counterfactual Sobel sequences favor a variably strict analysis of conditionals over its plainly strict alternative. Recent discussions of the lore have focussed on the question whether data about reverse counterfactual Sobel sequences actually speak in favor of a dynamic revival of the strict analysis. This paper takes the discussion into a new direction by looking at straight indicative Sobel sequences. The observation is that a variably strict analysis fails to predict the felicity of these sequences given minimal semantic and pragmatic assumptions. A properly elaborated dynamic analysis of indicatives, in contrast, handles the data with grace.

1 Indicative Sobel Sequences

Lewis [6] famously argues that the felicity of Sobel sequences speaks against a strict analysis of counterfactuals and instead supports a *variably* strict interpretation on which conditional consequents are evaluated at the *closest* possible worlds that verify the antecedent. Recent discussions of Lewis's argument have focussed on Heim's observation that *reverse* counterfactual Sobel sequences tend to be infelicitous: von Fintel [2] and Gillies [4] argue that this observation favors a *dynamic* strict interpretation of counterfactuals over a (static) variably strict analysis, but Moss [7] counters that the data can be handled by a pragmatic supplement to Lewis's account. This paper takes the discussion into a new direction by arguing that straight *indicative* Sobel sequences favor a dynamic strict interpretation of conditionals.

Indicative Sobel sequences are just as felicitous as their counterfactual cousins:

- (1) (a) If Alice comes to the party, it will be fun. (b) But if Alice and Bert come, it will not be fun. (c) But if Charles comes as well, it will be fun...

A variably strict analysis of indicatives à la Stalnaker [8] predicts that (1) is consistent, the simple observation being that the closest possible worlds at which Alice comes to the party need not be worlds at which Bert comes too and that the closest possible worlds at which both show up may be such that Charles stays at home.

Variably Strict Analysis. Take a connected and transitive relation \leqslant that is provided by context and keeps track of relative similarity or closeness between worlds:

1. $\min_{\leqslant, w}(\phi) = \{w' : w' \in \llbracket \phi \rrbracket \text{ and for all } w'' : \text{if } w'' \in \llbracket \phi \rrbracket, \text{ then } w' \leqslant_w w''\}$
2. $\llbracket (\text{if } \phi) (\psi) \rrbracket^w = 1 \text{ iff } \min_{\leqslant, w}(\phi) \subseteq \llbracket \psi \rrbracket$

This is fine as far as it goes, but it is not sufficient to predict that the sequence in (1) is felicitous since indicative conditionals also impose distinct constraints on the discourse context to be assertible. Compare:

- (2) (a) # Mary is not in New York. If she is in New York, she will meet Alex.
- (b) ✓ Mary is not in New York. If she were in New York, she would meet Alex.

The textbook explanation for the observation that (2a) is marked goes as follows. The first member of the sequence expresses the proposition that Mary is not in New York and eliminates from the context set—the set of possible worlds compatible with what is common

ground—all possible worlds at which that proposition is false (see [10]). But for the subsequent indicative to be felicitous there must be at least one such world in the context set because indicative conditionals—unlike their counterfactual cousins—presuppose that their antecedents are compatible with the common ground (see [9]).

The textbook explanation just given relies on two pragmatic assumptions that are widely accepted in the literature. Given some context c , let s_c be the context set of c :

1. If ϕ expresses a proposition $\llbracket\phi\rrbracket$ in context c , then the result of asserting ϕ in c , $c + \phi$, is such that $s_{c+\phi} = s_c \cap \llbracket\phi\rrbracket$.
2. An utterance of an indicative conditional of the form '(if ϕ) (ψ) ' in context c presupposes that $s_{c+\phi} \neq \emptyset$.

The problem for the variably strict analysis is this: while it correctly predicts that indicative Sobel sequences are consistent, it wrongly predicts that such sequences are infelicitous given minimal constraints on the semantics of conditionals. I assume—as Lewis and Stalnaker do—that modus ponens is valid and thus that the members of (1) are bounded from below by the material conditional (\supset) in the following sense:

3. (if ϕ) (ψ) $\vDash \phi \supset \psi$

Accordingly, on a variably strict analysis the similarity relation for the conditionals in (1) must be *weakly centered*: for all w and w' , $w \leq_w w'$. (Linguists are less concerned with modus ponens than philosophers but there are no attested counterexamples to this rule for the kind of indicative conditionals that figure in simple Sobel sequences.) But then the result of uttering '(if A) (F) ' and '(if $A \wedge B$) $(\neg F)$ ' in any context cannot contain an $(A \wedge B)$ -world since this would have to be an $(F \wedge \neg F)$ -world, and so the antecedent of (1c) is incompatible with what is common ground once its predecessors have been asserted. So if a variably strict semantics were right, we would expect the utterance of (1c) in the Sobel sequence to be marked due to presupposition failure, which is simply not the case.

One might suggest that the presupposition carried by (1c) is accommodated via expanding the context set but then we would expect this expansion to take place in (2a) as well since in both cases the conditional's antecedent is incompatible with the context set. One might also suggest that the presupposition violation is real but overlooked/ignored in discourse, but then Sobel sequences would have no bite against a plain strict analysis in the first place since it is hard to see why *its* predictions should not be overlooked/ignored in discourse as well. A variably strict interpretation thus has no good explanation for why indicative Sobel sequences should be felicitous.

The purpose of the next sections is to demonstrate that a proper *dynamic* strict analysis of indicative conditionals avoids the problem of its static variably strict alternative while preserving the pragmatic and semantic constraints articulated earlier. Assuming that conditionals are subject to a uniform semantic analysis, this result also suggests that a dynamic analysis of counterfactuals is on the right track.

2 Dynamics

The dynamic story I am about to tell is in the spirit of Gillies's [3] analysis of conditionals as tests on an information carrier and thus owes a lot of inspiration to Veltman's [12] update semantics. But it crucially departs from the classical conception of an input context as a plain set of possible worlds in order to distinguish between possibilities that are merely compatible with what is common ground and those that are 'live' in the sense that they are explicitly treated as relevant. The guiding intuition here is that indicative conditionals have the potential to highlight certain hitherto ignored possibilities in virtue of their presupposed content and then, in virtue of their asserted content, establish the corresponding material conditional as necessary in light of the possibilities treated as live. These two features

conspire to account for the felicity of indicative Sobel sequences, as we will see momentarily. But the first point to notice is that, at least for current purposes, it will not do to think of contexts as plain sets of possible worlds since in these models all possibilities are created equal. Instead I will model a context as a *set of sets of possible worlds* and define the notions of a (live) possibility in a supervaluationist fashion (see [13]).

I will first define the language under consideration and then make the notion of a context state more precise.

Definition 1 (Language). \mathcal{L} is the smallest set that contains a set of sentential atoms $\mathcal{A} = \{p, q, \dots\}$ and is closed under negation (\neg), conjunction (\wedge), the epistemic modals *might* (\diamond) and *must* (\Box), and the indicative conditional (if \cdot) (\cdot) . $\mathcal{L}^+ = \mathcal{L} \cup \{\partial\phi : \phi \in \mathcal{L}\}$. \mathcal{L}_0 is defined as the non-modal fragment of \mathcal{L} . Disjunction (\vee) and the material conditional (\supset) are defined in the usual way.

Definition 2 (Possible Worlds, Context States). w is a *possible world* iff $w : \mathcal{A} \rightarrow \{0, 1\}$. W is the set of such w 's, $\mathcal{P}(W)$ is the powerset of W . Σ is a *context state* iff $\Sigma \subseteq (\mathcal{P}(W) \setminus \emptyset)$, that is, a context state is a (possibly empty) set of nonempty sets of possible worlds. I is the set of such Σ 's. The set of *minimal* elements of Σ is defined as $\text{MIN}(\Sigma) = \{\sigma \in \Sigma : \neg \exists \sigma' \in \Sigma. \sigma' \subset \sigma\}$.

A context state, intuitively, is the set of sets of possible worlds satisfying everything that is common ground, and we distinguish between p being a possibility in the common ground and it being common ground that p is a possibility in the common ground. If p is a possibility in the common ground, $\neg p$ fails to be common ground and so there is at least one set of possible worlds satisfying everything that is common ground and that contains a p -world. So it makes sense to say that if Σ is our model of the common ground, then p is a possibility in Σ just in case there is *some* $\sigma \in \Sigma$ that contains a p -world. And whenever p is a live possibility in Σ —whenever it is common ground that p is a possibility in discourse—then *every* $\sigma \in \Sigma$ contains a p -world. Relatedly, notice that the set of minimal elements of Σ corresponds to the set of possible worlds just weak enough to accommodate all the possibilities that are live in discourse, and so it makes sense to say that ϕ is a necessity in light of the live possibilities in Σ —what I call a ‘live necessity’ in Σ —just in case ϕ is true throughout the minimal elements of Σ .

Definition 3 (Propositions, Possibilities, Necessities). Consider any $\Sigma \in I$ and $\phi \in \mathcal{L}_0$, and let the function $\llbracket \cdot \rrbracket$ assign to each $\phi \in \mathcal{L}_0$ a *proposition*, understood as a subset of W , in the familiar fashion. Define:

1. ϕ is a *possibility* in Σ iff $\exists \sigma \in \Sigma \exists w \in \sigma : w \in \llbracket \phi \rrbracket$
2. ϕ is a *live* possibility in Σ iff $\forall \sigma \in \text{MIN}(\Sigma) \exists w \in \sigma : w \in \llbracket \phi \rrbracket$
3. ϕ is a *necessity* in Σ iff $\forall \sigma \in \Sigma \forall w \in \sigma : w \in \llbracket \phi \rrbracket$
4. ϕ is a *live* necessity in Σ iff $\forall \sigma \in \text{MIN}(\Sigma) \forall w \in \sigma : w \in \llbracket \phi \rrbracket$

Notice here that live necessities may be *defeated* as hitherto ignored possibilities come into view: raising a possibility to a live possibility may *expand* the minimal elements of the common ground and thus remove a proposition from the set of its live necessities.

Context states are updated by updating each of their elements and so we first define an update operation on elements of such states.

Definition 4 (Updates on Elements of Context States). Consider any $\Sigma \in I$ and $p \in \mathcal{A}$, $\phi, \psi \in \mathcal{L}^+$. The operation $\uparrow : I \mapsto (\mathcal{L}^+ \mapsto \mathcal{P}(W)^{\mathcal{P}(W)})$ is defined as follows:

1. $\sigma \uparrow_\Sigma p = \{w \in \sigma : w(p) = 1\}$

2. $\sigma \uparrow_{\Sigma} \neg\phi = \sigma \setminus (\sigma \uparrow_{\Sigma} \phi)$
3. $\sigma \uparrow_{\Sigma} (\phi \wedge \psi) = (\sigma \uparrow_{\Sigma} \phi) \uparrow_{\Sigma'} \psi$, where $\Sigma' = \{\sigma \uparrow_{\Sigma} \phi : \sigma \in \Sigma\}$
4. $\sigma \uparrow_{\Sigma} \Diamond\phi = \{w \in \sigma : \exists \sigma' \in \Sigma. \sigma' \subseteq \sigma \text{ & } \sigma' \uparrow_{\Sigma} \phi \neq \emptyset\}$
5. $\sigma \uparrow_{\Sigma} \Box\phi = \{w \in \sigma : \exists \sigma' \in \Sigma. \sigma' \subseteq \sigma \text{ & } \sigma' \uparrow_{\Sigma} \phi = \sigma'\}$
6. $\sigma \uparrow_{\Sigma} \partial\phi = \sigma$ iff $\sigma \uparrow_{\Sigma} \phi = \sigma$
7. $\sigma \uparrow_{\Sigma} (\text{if } \phi)(\psi) = \sigma \uparrow_{\Sigma} (\partial\Diamond\phi \wedge \Box(\phi \supset \psi))$

Updating is defined relative to some context state Σ since the update rules for modals do not simply appeal to the information provided by the input set of possible worlds, as we will see momentarily. The rules for atomic sentences, negation, and conjunction are straightforward. An update of σ with an atomic sentence eliminates all possible worlds from σ at which the atomic sentence is false. To update with a negation, eliminate all possible worlds in the result of updating with what is negated. Updating σ with a conjunction basically comes down to an update with the first conjunct followed by an update with the second conjunct, but to arrive at an ‘internally dynamic’ conception of conjunction we stipulate that the second update takes place in light of a derived context state whose elements have been updated with the first conjunct.

To understand the update rules for *might* and *must*, remember that a context state is a subset of $\mathcal{P}(W)$ and thus corresponds to a set of sets of possible worlds ordered by the inclusion relation \subseteq , and that we have identified the live possibilities and necessities in the common ground in terms of its minimal elements. The simple idea then is that epistemic *might* and *must* update the common ground so that its prejacent becomes a live possibility and live necessity, respectively. Precisely, epistemic *might* trims each set of sets of possible worlds ordered by \subseteq so that its minimal element can be consistently updated with the prejacent. Epistemic *must* trims each set of sets of possible worlds ordered by \subseteq so that its minimal element satisfies the prejacent. Whenever the prejacent ϕ is a propositional formula, this just means that such a set of sets of possible worlds is eliminated from the common ground just in case ϕ fails to be true throughout its minimal sphere.

The rule for the presupposition operator repeats Beaver’s [1] analysis, which in turn articulates Heim’s [5] conception of presuppositions as definedness conditions: updating σ with $\partial\phi$ returns σ in case updating σ with ϕ idles, and is undefined otherwise.

The final clause articulates the hypothesis that indicative conditionals are strict over a dynamically evolving domain of quantification: they can be analyzed as articulating the presupposition that their antecedent might be the case coupled with the assertion that the corresponding material conditional is a live necessity in the common ground.

Define the following relations between a context state Σ and some $\phi \in \mathcal{L}^+$:

Definition 5 (Satisfaction, Support, Admission). Consider arbitrary $\sigma \subseteq W$, $\Sigma \in I$, and $\phi \in \mathcal{L}^+$:

1. σ satisfies ϕ with respect to Σ , $\sigma \Vdash_{\Sigma} \phi$, iff $\sigma \uparrow_{\Sigma} \phi = \sigma$
2. Σ supports ϕ , $\Sigma \Vdash \phi$, iff for all $\sigma \in \text{Min}(\Sigma) : \sigma \Vdash_{\Sigma} \phi$
3. Σ admits ϕ , $\Sigma \triangleright \phi$, iff for some $\sigma \in \text{Min}(\Sigma) : \sigma \uparrow_{\Sigma} \phi \neq \emptyset$

An element σ of a context state Σ satisfies ϕ just in case updating σ with ϕ in light of Σ idles. Σ supports ϕ just in case ϕ is satisfied by the minimal elements of Σ , that is, just in case ϕ is a weak necessity in Σ . Finally, Σ admits ϕ just in case some of its minimal elements may be consistently updated with ϕ , that is, just in case ‘ $\neg\phi$ ’ fails to be a weak necessity in Σ . This is all we need to say how exactly context states are updated:

Definition 6 (Updates on Context States). Consider arbitrary $\Sigma \in I$ and $\phi \in \mathcal{L}^+$. The update operation $[\phi] : I \mapsto I$ is defined as follows:

$$\Sigma[\phi] = \{\sigma : \Sigma \triangleright \phi \& \sigma \neq \emptyset \& \exists \sigma' \in \Sigma. \sigma' \uparrow_{\Sigma} = \sigma\}$$

An update of Σ with ϕ in effect tests whether Σ already accepts the negation of ϕ . If it does, the update returns the absurd state; otherwise, we update each element of Σ with ϕ and collect the results, leaving out the empty set. This yields the output state.

The notions of entailment and consistency are defined in the obvious fashion.

Definition 7 (Entailment, Consistency). Consider arbitrary $\phi_1, \dots, \phi_n, \psi \in \mathcal{L}^+$:

1. ϕ_1, \dots, ϕ_n entails ψ , $\phi_1, \dots, \phi_n \models \psi$, iff for all $\Sigma \in I$: $\Sigma[\phi_1] \dots [\phi_n] \Vdash \psi$
2. ϕ_1, \dots, ϕ_n is consistent iff for some $\Sigma \in I$: $\Sigma[\phi_1] \dots [\phi_n] \neq \emptyset$

An argument is valid just in case its conclusion is supported by every state once updated with its premises. A sequence is consistent just in case there is some context state that can be updated with that sequence without resulting in the absurd state. Notice that ϕ_1, \dots, ϕ_n entails ψ just in case $\phi_1, \dots, \phi_n, \neg\psi$ is inconsistent.

So much for the semantics. As far as the pragmatics is concerned, all we need to say here is that the result of uttering ϕ in some context c , $c + \phi$, is such that $\Sigma_{c+\phi} = \Sigma_c[\phi]$, where Σ_c and $\Sigma_{c+\phi}$ are the context states of c and $c + \phi$, respectively. Notice here that $\bigcup \Sigma_c = s_c$, which just means that context states allow us to say everything about discourse contexts that context sets allow us to say, and—as we saw earlier—a bit more.

3 Output

Let me now make a few observations about the framework developed here, the first one being that we preserve the constraints that figured prominently in the textbook explanation of why the sequence in (2a) is marked. Start with the following fact:

Fact 1. Consider arbitrary $\phi \in \mathcal{L}_0$ and $\Sigma \in I$: $\bigcup \Sigma[\phi] = \bigcup \Sigma \cap \llbracket \phi \rrbracket$

It follows that an utterance of an element from the propositional fragment of \mathcal{L} in some context adds the proposition expressed to the context set. This is just the familiar picture, but as in Veltman’s update semantics we leave room for context change that is not mediated by propositional content.

The next fact concerns the presuppositions of indicative conditionals. Start by defining what it takes for a sentence to presuppose another one (see [1] for inspiration):

Definition 8 (Presupposition). ϕ presupposes ψ , $\phi \gg \psi$, iff for all $\sigma \subseteq W$ and $\Sigma \in I$: if $\sigma \uparrow_{\Sigma} \phi$ is defined, then σ satisfies ψ with respect to Σ .

It then follows immediately that indicative conditionals presuppose that their antecedent might be the case:

Fact 2. $(\text{if } \phi)(\psi) \gg \diamond \phi$

Given the semantics for epistemic *might*, it is then easy to verify that $\Sigma[(\text{if } \phi)(\psi)] \neq \emptyset$ only if $\exists \sigma \in \Sigma : \sigma \uparrow_{\Sigma} \phi \neq \emptyset$, that is, only if $\bigcup \Sigma[\phi] \neq \emptyset$. Notice, furthermore, that the *might*-content carried by an indicative conditional *projects* just in the way we expect in case it is presupposed. For instance, we predict that ‘(if ϕ)(ψ)’ as well as its negation presuppose that the antecedent ϕ might be the case.

The previous two observations, together with the general pragmatic assumption that assertions affect a context by updating the corresponding context state, show that the framework developed here preserves the key aspects of the textbook explanation for why a sequence such as (2a) is marked. In addition, modus ponens is valid:

Fact 3. $(\text{if } \phi)(\psi) \models \phi \supset \psi$

This is because an update with ' $(\text{if } \phi)(\psi)$ ' involves an update with ' $\Box(\phi \supset \psi)$ ' and thus establishes the material conditional as a live necessity in the common ground.

All of that, and we can *still* make sense of the observation that indicative Sobel sequences are felicitous. To show that this is so, it is sufficient to demonstrate that indicative Sobel sequences are consistent, since updating with a presupposition that cannot be accommodated is guaranteed to result in the absurd state. Here is the first thing to notice:

Fact 4. $(\text{if } A)(F), (\text{if } A \wedge B)(\neg F)$ is consistent

For consider $\Sigma' = \Sigma[(\text{if } A)(F)]$: then $\Sigma' \Vdash \Diamond A$ and $\Sigma' \Vdash A \supset F$. But $\Diamond A \not\models \Diamond(A \wedge B)$ and so it may very well be that $\Sigma' \not\models \Diamond(A \wedge B)$. In that case, accommodating the presupposition carried by $(\text{if } A \wedge B)(F)$ in Σ' will induce a non-trivial change and, in particular, may defeat $A \supset F$ as a live necessity, leaving room for consistent update with $\Box((A \wedge B) \supset \neg F)$.

What caused the trouble for a variably strict semantics was that (1a) and (1b) are consistent but entail that Alex and Bert will not both come to the party. This result is avoided in the current framework:

Fact 5. $(\text{if } A)(F), (\text{if } A \wedge B)(\neg F) \not\models \neg(A \wedge B)$

True enough, $(\text{if } A)(F) \models A \supset F$ and $(\text{if } A \wedge B)(\neg F) \models (A \wedge B) \supset \neg F$, and also $A \supset F, (A \wedge B) \supset \neg F \models \neg(A \wedge B)$. But crucially $(\text{if } A)(F), (\text{if } A \wedge B)(\neg F) \not\models A \supset F$ since live necessities may be defeated as open but hitherto ignored possibilities come into view. So just as it is possible to update $\Sigma[(\text{if } A)(F)]$ with $(\text{if } A \wedge B)(\neg F)$ since $\Diamond A \not\models \Diamond(A \wedge B)$, it is possible to update $\Sigma[(\text{if } A)(F)][(\text{if } A \wedge B)(\neg F)]$ with $(\text{if } A \wedge B \wedge C)(F)$ since $\Diamond(A \wedge B) \not\models \Diamond(A \wedge B \wedge C)$.

A very simple example of a context that may be consistently updated with the sequence in (1) is the following. Consider $\Sigma = \{\{w_1\}, \{w_1, w_2\}, \{w_1, w_2, w_3\}\}$ and suppose that $w_1 \in \llbracket A \wedge \neg B \wedge \neg C \wedge F \rrbracket$, $w_2 \in \llbracket A \wedge B \wedge \neg C \wedge \neg F \rrbracket$, and $w_3 \in \llbracket A \wedge B \wedge C \wedge F \rrbracket$. Then $\Sigma[(\text{if } A)(F)] = \Sigma$ since $\Sigma[\partial \Diamond A] = \Sigma$ and $\text{MIN}(\Sigma) \Vdash A \supset F$ (simply observe here that $\text{MIN}(\Sigma) = \{\{w_1\}\}$). But $\Sigma' = \Sigma[\partial \Diamond(A \wedge B)] = \{\{w_1, w_2\}, \{w_1, w_2, w_3\}\}$ and so $\text{MIN}(\Sigma') = \{\{w_1, w_2\}\}$. Accordingly, even though it holds that $\Sigma[\Box((A \wedge B) \supset \neg F)] = \emptyset$, $\Sigma[\partial \Diamond(A \wedge B)] \Vdash \Box((A \wedge B) \supset \neg F)$. Finally, $\Sigma'' = \Sigma'[\partial \Diamond(A \wedge B \wedge C)] = \{\{w_1, w_2, w_3\}\}$ and clearly $\Sigma'' \Vdash \Box((A \wedge B \wedge C) \supset F)$. Accordingly, Σ can be consistently updated with the sequence in (1) without resulting in the empty set, which is sufficient to show that indicative Sobel sequences are predicted to be felicitous.

What explains the consistency of indicative Sobel sequences is that the dynamic logical consequence relation is *nonmonotonic*, and in particular we have:

Fact 6. $(\text{if } A)(F) \models (\text{if } A)(F)$ but $(\text{if } A)(F), (\text{if } A \wedge B)(\neg F) \not\models (\text{if } A)(F)$

On the view developed here, then, indicative Sobel sequences are consistent because each member of the sequence *defeats* its predecessor. The reason is familiar: the predecessor establishes its corresponding material conditional as a live necessity, but live necessities may be defeated as hitherto ignored possibilities come into view—which is just what the subsequent conditional does in virtue of its presupposed content. This dynamic approach crucially differs from the static variably strict analysis on which Sobel sequences are classically consistent since their consequents are evaluated with respect to disjoint modal domains. If what I have said is right here, a nonmonotonic perspective on indicative Sobel sequences is inevitable if we take the validity of modus ponens seriously and subscribe to the textbook explanation of why (2a) is marked.

4 Previous Dynamic Analyses

Let me briefly outline how the framework developed here differs from previous dynamic analyses of conditionals. The proposals by Gillies [3] and Starr [11] also analyze indicative conditionals as strict material conditionals with an additional presuppositional flavor but model their semantics as tests on contexts understood as plain sets of possible worlds. This is good enough for their purposes, but it is easy to verify that it will not be good enough to handle the data that interest us here. For suppose that input contexts were just singletons of the form $\Sigma = \{\sigma\}$. Then $\Sigma' = \Sigma[(\text{if } \phi)(\chi)]$ does not admit '(if $\phi \wedge \psi$) ($\neg\chi$)' for either $\Sigma' \not\models \diamond(\phi \wedge \psi)$ and thus $\Sigma'[\partial\diamond(\phi \wedge \psi)] = \emptyset$ or $\Sigma \models \diamond(\phi \wedge \psi)$ and then $\Sigma[\square((\phi \wedge \psi) \supset \neg\chi)] = \emptyset$. A more complex conception of a context and, accordingly, of how assertions operate on contexts is needed to articulate how the elements of a dynamic analysis conspire to leave room for felicitous indicative Sobel sequences.

The proposals by von Fintel [2] and Gillies [4], while different in detail, effectively assign to counterfactuals strict truth-conditions with respect to a possible world w and a single, dynamically evolving system of spheres initially centered on w . They do not address indicatives but it makes sense to ask why we cannot do something similar to handle the case at hand. To see the issue more clearly, define the *admissible domains* around w as S_w and say that $\sigma \in S_w$ iff $\exists v \in W$ such that $\sigma = \{u : u \leq_w v\}$. The suggestion then is that indicatives are assigned truth-values relative to a possible world w and a *hyperdomain* π at w , which is nothing but a subset of S_w ordered by \subseteq . Using the framework developed earlier, we may then state what it takes for an indicative to be true as follows:

$$\llbracket (\text{if } \phi)(\psi) \rrbracket^{w,\pi} = 1 \text{ iff } \pi[\partial\diamond\phi] \models \square(\phi \supset \psi)$$

If we now assume that the hyperdomain dynamically evolves as discourse proceeds, the resulting proposal looks close to the framework developed earlier. But there are very good reasons for telling the story just the way I did. Let me explain.

First, the truth-conditional proposal under consideration does not avoid the main problem of its classical alternative since it predicts that (if A)(F) and (if $A \wedge B$) ($\neg F$) entail $\neg(A \wedge B)$. For suppose that $w \in \llbracket A \wedge B \rrbracket$ and let π be an arbitrary hyperdomain at w : since $\{w\}$ is a subset of $\text{MIN}(\pi)$ by design, the *might*-presuppositions carried by (if A)(F) and (if $A \wedge B$) ($\neg F$) cannot trigger an expansion of the minimal sphere in π , and so there would be no way for the two conditionals to be true at w and π . Accordingly, it once again becomes hard to see why a context that has been updated with (1a) and (1b) should admit an update with (1c). For no such context—assuming that contexts inherit the semantic commitments that come with their updates—will admit the presupposition that Alice and Bert and Charles might come to the party, and so we once again do not expect that (1c) may be uttered without fuss once (1a) and (1b) have been asserted.

The second reason is that the framework developed here offers distinct advantages over the dynamic truth-conditional proposal when it comes to *might*-conditionals. Conditionals, I have said, transform open but hitherto ignored possibilities into view in virtue of their presupposed content. But that is not the only source of shiftiness since, intuitively, *might*-conditionals may do so as well in virtue of their asserted content. Consider:

- (3) (a) If Alice comes to the party, it will be fun. (b) But if Alice comes, Bert might come as well, and then it will not be fun...

The fact that (3) seems felicitous strongly suggests that (3b) highlights the possibility of Alice and Bert's coming to the party, and this is just what the framework developed here predicts. For remember that an indicative establishes the corresponding material conditional as a live necessity in virtue of its asserted content and that ϕ is a live necessity in Σ just in case each of its minimal elements satisfies ϕ with respect to that context state. But σ satisfies $A \supset \diamond(A \wedge B)$ just in case σ either contains no A -world or includes at least some $(A \wedge B)$ -world. But every element of Σ will contain an A -world once the

presupposition of (3b) has been accommodated, and so updating Σ with its asserted content establishes Alice and Bert's coming to the party as a live possibility.

On the truth-conditional alternative under consideration, in contrast, a conditional's asserted content does *not* have the potential to modify a hyperdomain but is simply evaluated for truth or falsity with respect to some hyperdomain. Gillies recognizes this limitation and suggests that conditionals do not only presuppose that their antecedent is a possibility, but also that their consequent is a possibility. While this move predicts that (3b) highlights the possibility of Alice and Bert's coming to the party, it seems fair to say that it relies on a plain stipulation with little independent support. No such additional stipulation is needed in the dynamic framework suggested here.

5 Conclusion

The felicity of straight indicative Sobel sequences favors a dynamic strict analysis of conditionals over its static variably strict alternative. I have not discussed how the story told here may be expanded to cover counterfactuals and what a dynamic story has to say about the complex data surrounding reverse Sobel sequences. While doing so is not an entirely trivial affair, I submit that the data discussed here strongly suggest that a dynamic semantic analysis of conditionals is on the right track.

Acknowledgments An earlier version of this paper was presented at the University of Texas at Austin. Special thanks to the audience for comments and discussion, as well as to Fabrizio Cariani and Itamar Francez.

References

- [1] David I. Beaver. *Presupposition and Assertion in Dynamic Semantics*. CSLI Publications, Stanford, 2001.
- [2] Kai von Fintel. Counterfactuals in a dynamic context. In Michael Kenstowicz, editor, *Ken Hale: A Life in Language*, pages 123–152. MIT Press, Cambridge, MA, 2001.
- [3] Anthony S. Gillies. Epistemic conditionals and conditional epistemics. *Noûs*, 38(4):585–616, 2004.
- [4] Anthony S. Gillies. Counterfactual scorekeeping. *Linguistics and Philosophy*, 30(3):329–360, 2007.
- [5] Irene Heim. *The Semantics of Definite and Indefinite Noun Phrases*. PhD thesis, University of Massachusetts, Amherst, 1982.
- [6] David K. Lewis. *Counterfactuals*. Harvard University Press, Cambridge, MA, 1973.
- [7] Sarah Moss. On the pragmatics of counterfactuals. *Noûs*, 46(3):561–586, 2012.
- [8] Robert C. Stalnaker. A theory of conditionals. In Nicholas Rescher, editor, *Studies in Logical Theory*, pages 98–112. Blackwell, Oxford, 1968.
- [9] Robert C. Stalnaker. Indicative conditionals. *Philosophia*, 5(3):269–286, 1975.
- [10] Robert C. Stalnaker. Assertion. In Peter Cole, editor, *Syntax and Semantics 9*, pages 315–332. New York Academic Press, New York, 1978.
- [11] William B. Starr. A uniform theory of conditionals. *Journal of Philosophical Logic*, Forthcoming.
- [12] Frank Veltman. Defaults in update semantics. *Journal of Philosophical Logic*, 25(3):221–261, 1996.
- [13] Malte Willer. Dynamics of epistemic modality. *Philosophical Review*, 122(1):45–92, 2013.

Deontic scope restrictions beyond polarity *

Igor Yanovich

Universität Tübingen, igor.yanovich@uni-tuebingen.de

Abstract

Deontic modals often have restricted scope with respect to clausemate negation. Recently attempts have been made to attribute all such restrictions to the polarity status of the respective modals ([Iatridou and Zeijlstra, 2013], [Homer, 2013]). I argue that once we examine the full range of cross-linguistic data, such a purely-syntactic account is hardly tenable. I suggest to complement polarity restrictions on deontics with semantic filters on scope. Conventions regarding such filters arise as speakers generalize from the range of actual uses that a pre-deontic construction used to have.

The aim of this paper is to show that the current syntactic accounts that view the scope restrictions of deontic modals as polarity restrictions cannot capture the whole range of empirical data, and need to be complemented with semantic-convention filters on scope configurations. After examining a wider range of data on deontic scope than it has been yet done in the formal literature, I argue that the observed diversity makes a purely syntactic account highly implausible. I therefore propose to complement syntactic polarity mechanisms with a semantic mechanism which may restrict the available range of scope construals for a given tense-aspect-mood form of the modal.

1 Deontics and clausemate negation: the state of the art

Deontic modals often have restricted scope with respect to clausemate negation. Possibility deontics (that is, permission modals) seem to universally scope under negation, see [van der Auwera, 2001, Sec. 5.6, 5.7], but necessity deontics (obligation modals) show a range of different behaviors. I have nothing to say about the scope restrictions of permission deontics, and set them aside for the purposes of this paper. From this point on, only obligation deontics will be discussed. [Iatridou and Zeijlstra, 2013] distinguish three types of such modals, given in 1:

- (1) I&Z's three polarity types of deontics:
 1. **PPI modals**: as other PPIs, need to be licensed by being in a non-downward-entailing (DE) context. **Examples**: *must*, *should*, Dutch *moeten*, Greek *prepi*
 2. **NPI modals**: as other NPIs, need to be licensed by being in a DE context
Examples: *need*, German *brauchen*, Dutch *hoeven*
 3. **Neutral modals**: are OK in upward-entailing contexts, but in the presence of negation scope under it. **Examples**: *have to*, German *müssen*

*The research reported here has benefitted from presentations at Ottawa University, UConn, MIT, Stanford and UCLA, and closely corresponds to a part of the Chapter 5 of my dissertation, [Yanovich, 2013]. The input of my advisers Kai von Fintel, Irene Heim and Sabine Iatridou has helped me to enormously improve the work reported, and is hereby most gratefully acknowledged. I also benefitted a lot from the discussions with Enoch Aboh, Jonathan Bobaljik, Cleo Condoravdi, Nathalie Dion, Olga Fischer, Sabine Iatridou, Magdalena Kaufmann, Stefan Kaufmann, Paul Kiparsky, Natalia Korotkova, Sven Lauer, Yael Sharvit, Sali Tagliamonte and Yakov Testelets.

Iatridou and Zeijlstra employ two mechanisms to derive the empirically observed scope configurations from the surface structure. First, they argue that the modal appearing in the TP zone may reconstruct to a position within VP, and thus below negation. This is how I&Z derive the narrow scope for the NPI modal *need* in 2.

- (2) Mary needn't leave.
 = 'It is not that Mary needs to leave'
 ≠ 'Mary needs to not leave'

OK $\neg > \square$
 * $\square > \neg$

Second, when a modal that occurs below negation in the surface syntax needs to scope above it, as in 3, I&Z posit covert, QR-like movement of the modal over the negation.

- (3) O Yanis dhen prepi na figi.
 John NEG \square -DEONTIC leave
 ≠ 'It is not that John has to leave'
 = 'John has to not leave'

* $\neg > \square$
OK $\square > \neg$

Thus reconstruction takes care of the cases when the modal needs to scope lower than it stays in the surface syntax, and covert QR-like movement applies when the modal needs to scope higher than its surface position. In the case of NPI or PPI modals, I&Z's analysis straightforwardly derives the facts: 1) due to the polarity requirements, only one of the two possible scope configurations allows the polarity-item-modal to be licensed; 2) if the only licensed scope configuration can be read off the surface structure, so be it; and if not, then either reconstruction or covert QR-like movement derives it. But for the third type of modals, which I&Z call "neutral", more must be said. Those modals are happy in upward-entailing environments, unlike NPIs. Yet when they occur in the same clause with negation, they obligatorily scope under it. So on the one hand, they are not polarity items, but on the other, they have fixed scope with respect to negation. English *have to* and German *müssen* are two examples:

- (4) Mary doesn't have to leave.
 = 'It is not that Mary has to leave'
 ≠ 'Mary has to not leave'
- (5) Hans muss nicht abfahren.
 Hans \square -DEONTIC NEG leave
 = 'It is not that John has to leave'
 ≠ 'John has to not leave'

OK $\neg > \square$
 * $\square > \neg$

OK $\neg > \square$
 * $\square > \neg$

In the surface structure, *have to* appears below negation, while *müssen* appears above it. I&Z take modals like *have to* to pose no problems in their system: "we need say nothing further for neutral modals that surface to the right of negation. They are simply interpreted where they appear in the overt syntax; that is, their syntactic and semantic scopes are identical", I&Z's p. 547. Now, that is not quite correct: as I&Z themselves discuss, modals have to have the option of undergoing covert QR-like movement for scope purposes; that option, other things being equal, should be available to *have to* in 4, and produce wide scope for the modal. But this is not what happens.

I&Z's account for non-NPI modals like German *müssen* or Spanish *tener que* which appear above negation in the surface structure is also problematic. For such modals, I&Z have to explain why their scope cannot be just read off the surface structure. To do that, I&Z introduce a principle forcing such modals to have narrow scope:

- (6) [Hatridou and Zeijlstra, 2013, (69b)]:

Head-movement reconstructs unless reconstruction would result in a grammatical violation.

I&Z argue that given the principle in 6, the scopal properties of 5 follow: 1) as *müssen* is a non-polarity modal according to I&Z, there is no reason that would prevent it from scoping under the negation; 2) from 6, we derive that *müssen* obligatorily reconstructs.

There are two problems with this explanation. First, the principle in 6 only rules out one class of derivations that may lead to the $\square > \neg$ construal. There are at least two other types of derivations that need to be independently blocked: A) after reconstructing due to 6, a modal like *müssen* may undergo QR-like movement upwards across the negation; B) the modal may undergo QR-like movement from its surface position, ending up in a position above the negation; for modals like *müssen*, that would block head-movement reconstruction, rendering the principle in 6 irrelevant.

I do not claim that one absolutely cannot introduce constraints that would rule out all the derivations that lead to illicit scope construals. But for *have to*-type and *müssen*-type modals, we would have to introduce very different constraints that result in the same interpretational restrictions. For *have to* we need to prohibit QR-like movement from the base position of the modal. For *müssen* we need to prohibit QR-like movement from a raised position at T. At the same time, we cannot prohibit QR-like movement for modals in general, as it is needed to derive the observed scope configurations for other modals. There seems to be no principled way to carve out the relevant constraints on covert movement which would apply to *have to* and *müssen* without directly mentioning those modals—rather than some structural configuration—in the definition. In other words, the narrow scope of *have to* and *müssen* seems to be an idiosyncratic property of those modals, and not the consequence of how general syntactic mechanisms work.

The second problem with I&Z's explanation only reinforces that conclusion. In addition to non-polarity modals with fixed scope, there exist freely-scoping deontics. For example, I&Z themselves discuss French *devoir* which has free scope freely in simple present sentences like 7:

- (7) Il ne doit pas partir.
 He EXPL.NEG \square -DEONTIC NEG leave
 = 'It is not that John has to leave'
 = 'John has to not leave'

$OK \neg > \square$
 $OK \square > \neg$

Modals like *devoir* do not fit into I&Z's classification in 1, but I&Z suggest that perhaps the principle 6 forcing obligatory reconstruction of head movement is language-specific, and does not exist in French. However, saying that French does not obey 6 is still not enough to account for the behavior of *devoir*, because French does not permit free scoping for *devoir* in all cases. As I&Z themselves note, in indicative perfectives, *devoir* is required to scope below the negation, 8.

- (8) Jean n' a pas du prendre l'autobus.
 Jean EXPL.NEG have NEG \square -DEONTIC take the bus
 = 'It is not that John had to take the bus'
 \neq 'John had to not take the bus'

$OK \neg > \square$
 $* \square > \neg$

While in 7, *devoir* has moved over the interpretable negation *pas* (with higher *ne* being an expletive, omittable negative particle), in 8 the T position above *pas* is occupied by the auxiliary *avoir* that blocks head movement by *devoir*. I&Z hypothesize that the presence of the auxiliary that blocks head movement of the modal to T may be the reason why scope of *devoir* is restricted in 8,

and call for future research on the issue. But the solution suggested by I&Z seems untenable in view of the data from irrealis perfectives, unknown to them. In examples like 9, the scope of *devoir* is fixed, but the modal scopes *above* negation. Thus in both 8 and 9, there is an auxiliary that would have blocked head movement over negation by *devoir*. Yet the scope is fixed differently in the two constructions.¹

- (9) Jean n' aurait pas du prendre l'autobus.
 Jean EXPL.NEG would.have NEG □-DEONTIC take the bus
 ≠ 'It is not that John should have taken the bus'
 = 'John should not have taken the bus'
- * $\neg > \square$
OK $\square > \neg$

The problems with *devoir* add to the problems with *have to* and *müssen*: there are plenty of restrictions on the scope of non-polarity-item modals, but they do not appear to be caused by general syntactic principles. Rather it seems that individual lexical items, or even the pair of a lexical item and a particular tense-aspect-mood combination, may have associated scope constraints.

Summing up the discussion so far, we can conclude the following:

- (10) Positing that modals by default reconstruct to VP-internal positions cannot by itself derive the narrow scope of I&Z's "neutral" modals with respect to negation.
 (from German *müssen* and English *have to*)
- (11) Constraints on the scoping of modals and negation may be specific to particular tense-aspect-mood combinations.
 (from French *devoir*)

In the next section, I will introduce data from Russian deontics that further illustrate the insufficiency of I&Z's system. While French *devoir* scopes freely in the present and the non-finite forms, its scope is at least sometimes fixed, but Russian features modals that are truly neutral in that they always permit both scope construals with respect to clausemate negation. In view of the existence of such true neutrality, the scope restrictions of modals such as German *müssen*, English *have to* and French *devoir* should be viewed as idiosyncratic to those particular words.

2 Russian deontics: true neutrality with respect to negation

Normally, if a scope-bearing expression is neither an NPI or a PPI, its scope with respect to negation is not fixed. For example, indefinites such as *two books* are polarity-neutral, and therefore may scope both above and below clausemate negation. In contrast to that, in the modal domain I&Z assign the label of polarity-neutral modals to *have to* and *müssen* which obligatorily scope below clausemate negation. This was a reasonable move given that I&Z did not find any necessity deontic that would be completely neutral with respect to negation. French *devoir* gets closest to that, but it still has restricted scope in some tense-aspect-mood forms. In this section, I provide data from Russian necessity deontics that are truly neutral: they scope freely with respect to their clausemate negation. From here on, I will reserve the term *neutral* to such truly polarity-neutral expressions. Thus I&Z's category of "neutral" modals in fact features modals subject to scope restrictions, even though those may stem from something different than polarity sensitivity.

¹Enoch Aboh (p.c.) notes that it is often the case that the syntactic structure of indicative and irrealis clauses is different. He conjectures that there could be a way to explain on syntax-internal grounds why an indicative auxiliary would block, but an irrealis auxiliary would not block the movement of the modal. In the present author's view, even if that is possible, it still remains hard to see why the narrow scope construal would be banned in 9.

Another important feature of the Russian system is that in addition to free-scope deontics, Russian also has both $\square > \neg$ and $\neg > \square$ fixed-scope necessity deontics. Free-scoping and restricted deontics have similar syntax in Russian, all occurring below negation in the surface structure. This further illustrates that scope restrictions are often specific to particular modals, rather than stem from the general properties of the syntactic system (contra I&Z, who propose the language-specific principle of obligatory reconstruction of head movement in order to derive the fixed scope of German *müssen* and Spanish *tener que*).

2.1 Russian free-scope deontics

Morphologically and syntactically, most deontics in Russian are predicative adjectives taking as arguments a Nominative or a Dative subject, and an infinitive clause. Predicative adjectives in Russian require the presence of copula *bytj* ‘be’. In the present, the copula’s form is \emptyset , so it is not visible on the surface.

Russian modals *dolžná* (that can have the deontic, teleological and epistemic modal flavors) and *núžno* (need/deontic/teleological) have free scope with respect to clausemate negation:

- (12) Ona ne dolžna upominatj o svojom znakomstve s Anej.
she NEG \square -DEONTIC mention about her acquaintance with Anya
 $\square > \neg$: ‘She mustn’t mention she’s acquainted with Anya.’
- (13) Maša objasnila, što Anja ne dolžna pisatj otčot.
Masha explained that Anya NEG \square -DEONTIC write report
 $\neg > \square$: ‘Masha explained that Anya does not have to write a report.’
- (14) Ej ne núžno segodnja prinositj svoj obed:
she.DAT NEG \square -DEONTIC today bring her lunch
'She {mustn't / doesn't have to} bring her lunch today: '
 - a. ... xolodiljnik slomalsja, i dekan poprosila poka ne prinositj svoju edu.
fridge broke, and chair asked yet NEG bring one's food
 $\square > \neg$: ‘... the fridge broke down, and the chair asked (everyone) to not bring their food until further notice.’
 - b. ... na fakultete budet furšet.
on department will.be catered.food
 $\neg > \square$: ‘... there will be catered food in the department.’

The same freedom of scoping is retained in the past tense:

- (15) Ona ne dolžna byla upominatj o svojom znakomstve s Anej.
she NEG \square -DEONTIC was mention about her acquaintance with Anya
 $\square > \neg$: ‘She had to keep silent about her acquaintance with Anya.’
 $\neg > \square$: ‘She didn’t have to keep silent about her acquaintance with Anya.’
- (16) Ej ne núžno bylo upominatj o svojom znakomstve s Anej.
she.DAT NEG \square -DEONTIC was mention about her acquaintance with Anya
 $\square > \neg$: ‘She had to keep silent about her acquaintance with Anya.’
 $\neg > \square$: ‘She didn’t have to keep silent about her acquaintance with Anya.’

Both *dolžna* and *nužno* appear below sentential negation in the surface syntax: negation in Russian always occupies a high position, cliticizing on the left to the highest finite element in its clause.² Thus the scope configuration $\neg > \square$ may be read off the surface structure, but the $\square > \neg$ interpretation has to be derived by covert QR-like movement of the modal, or its analogue.

2.2 Russian fixed-scope deontics

Not all Russian modals scope freely with respect to clausemate negation. In particular, *stoit* (an “advice” modal) always scopes over negation, while *objazana* (a deontic) always scopes below it:

- (17) Tebe ne stoit begatj po utram.
you.DAT NEG \square -ADVICE run on mornings
OK $\square > \neg$: “You should not run in the mornings.”
* $\neg > \square$: “It is not that you should run in the mornings.”
- (18) Ona ne objazana byla pisatj otčot o pomezke.
she NEG \square -DEONTIC was write report about trip
* $\square > \neg$: “She was required to not write a report about the trip.”
OK $\neg > \square$: “She did not have to write a report about the trip.”

Objazana belongs to the same morphosyntactic category as *dolžna*, so the comparison between the two shows that it really depends on the lexical item alone which scope construals are allowed. *Stoit* is an impersonal verb, not a predicative adjective, but its surface position is under negation, just as for predicative adjectives, so there is no reason to think the morphosyntactic differences between *stoit* and *nužno* or *dolžna* should have any consequence with respect to their interaction with clausemate negation.

²The role of that “highest finite element” in 15 and 16 is taken up by the predicative adjective: on the surface, it looks as if the adjective has head-moved across the copula, in a pattern similar to the well-known “long head movement” pattern in South and West Slavic. It is still an open question what exact underlying structure corresponds to the linear order (*Neg*) *Adj Aux* in different Slavic languages, the several analytical options being: (1) [Rivero, 1994] argues for a long head movement analysis, with the adjectival element (a participle or an actual adjective) appearing in C; (2) [Bošković, 1995] argues for incorporation analysis, in which the adjectival element merges into the copula head; (3) [Ackema and Čamđić, 2003] arguing for base-generating the adjectival element in the higher position. This does not exhaust the analytical options, and some authors, esp. [Embick and Izvorski, 1997], caution against lumping together all instances of the order *Adj Aux*, within the same language or across different Slavic languages (cf. also [Borsley and Rivero, 1994] and [Broekhuis and Migdalski, 2003]).

Russian data, previously not described in connection to “long head movement”/“Adj movement across the copula” in other Slavic, should bear on that debate, though much future research is needed. For example, the fact that the Russian sentence in (i) is grammatical shows that either the incorporation analysis of Bošković does not work for Russian, or it does not work for BCS. Bošković argues that if there is a need for the incorporated adjectival element to move further up, it necessarily excorporates, stranding the copula below. So when Adj has an independent reason to move higher from the complex *Neg-Adj-Aux*, it strands *Neg-Aux* — as can be observed in BCS. But in the Russian (i), negation is not stranded, but taken along by the moving adjective. Thus either Russian orders *Adj Aux* do not feature Bošković-style incorporation, or Bošković is wrong about the excorporation requirement, which would destroy his account of the BCS data.

- (i) Ne nužno Maše bylo tuda ezdijt.
NEG \square -DEONTIC Masha.DAT was there go
 $\square > \neg$: ‘Masha should not have gone there.’
 $\neg > \square$: ‘It is not that Masha should have gone there.’

New Russian data above add further evidence that scope restrictions of modals are so diverse that they are likely to be associated directly with individual modals, or even forms thereof, rather than stem from major principles regulating the work of the syntactic component.

3 Semantic-convention filters on modal scope construals

In view of the evidence we discussed above, I argue for the following:

- (19) Modals may be PPIs (like *must*) or NPIs (like *need*), or they may be not polarity-sensitive at all (like English *have to*, French *devoir*, Russian *nužno*).
- (20) The syntactic component allows a non-polarity-sensitive modal to have either scope with respect to clausemate negation.
(After [Iatridou and Zeijlstra, 2013], we can take the mechanisms responsible for this to be reconstruction and scope-changing covert movement.)
- (21) Not every modal uses all the possibilities made available by the syntax.
Namely, there may be further constraints in place that rule out a particular scope configuration for a given modal word (English *have to*, Russian *objazana*), or for a modal word within a particular environment (French *devoir* in indicative vs. irrealis perfectives).

Thus polarity-sensitivity is not the only mechanism that may constrain how a modal scopes with respect to clausemate negation. In particular, the case of *devoir* suggests that scope restrictions may be tied to particular “constructions” rather than attached to lexical items. It is hardly possible to derive such construction-specific constraints using general-purpose syntactic mechanisms. But do we have any way to account for scope constraints not using syntactic mechanisms?

I argue that we do, and that fixed-scope constraints may be imposed by the lexical and construction-specific semantics and pragmatics of the language. We know independently that certain meanings may be indexed to particular constructions rather than follow from the compositional semantics of the lexical items alone. For example, *Can you pass me the salt?* conventionally conveys a request, while *Are you capable of passing me the salt?* is not (cf. [Horn and Bayer, 1984]). Even the latter sentence may give rise to the implicature of a request, but only the first does so conventionally, without requiring much pragmatic reasoning. There is no *a priori* reason why restrictions on relative scope could not be conventionalized, too.

A semantic convention imposing fixed scope would be learned in the same way speakers learn the lexical meanings of words. After language learners hear a word used a large number of times in a similar way, they abstract from those occurrences a semantic representation for the word. The semantics of a word gets generalized from individual instances in such a way as to be capable of explaining each of those. Statistically significant absence of positive evidence works as negative evidence in the creation of such conventions: we know that *rabbit* cannot denote a frog because we never heard people use *rabbit* that way. Finally, meanings may be associated not with individual words, but with larger chunks of structure, as the existence of idioms shows.

The acquisition of a semantic filter on scope construals would proceed similarly. For example, the learners would hear surface string such as $\square \text{ Neg}$, and due to the existence of the constraint in the speech of competent speakers, that string would only be used in sentences conveying the $\text{Neg} > \square$ reading. With only a few examples of this sort encountered, the learners could have not noticed the pattern. But the more frequent the surface string $\square \text{ Neg}$ is, the more striking it becomes that it is only used to convey the $\text{Neg} > \square$ reading. As learners are sensitive to such statistical evidence,

interpreting it as a sign that something should be ruled out by the grammar, they acquire a scope constraint. If the constraint can be tied to the workings of a general syntactic mechanism (e.g., to the licensing of polarity items), then learners may acquire a syntactic constraint. But if the constraint seems to be idiosyncratically tied to a particular modal, or even to its combination with a particular tense-aspect-mood form, a semantic filter may become established in the grammars being acquired.

Once a semantic convention is established, it will perpetuate itself, other things being equal. The usage of all members of the linguistic community will be constrained by the convention, and new speakers will learn to conform to the same convention as they acquire language, unless there is pressure for language change. So the explanatory burden associated with positing a particular semantic convention restricting modal scope is to demonstrate how it got conventionalized: once it is established, the speakers will use the restriction until they have a good reason not to; it is the rise of the restriction that is not a trivial matter. For reasons of space I omit here case studies showing how we can trace a scope filter's rise; in [Yanovich, 2013, pp.203-226], I discuss the conventionalization of the wide scope of Russian *stoit* 'should' and the narrow scope of English *have to*.

References

- [Ackema and Čamđić, 2003] Ackema, P. and Čamđić, A. (2003). LF complex predicate formation: The case of participle fronting in Serbo-Croatian. In *UCL Working Papers in Linguistics*, volume 15, pages 131–175.
- [Borsley and Rivero, 1994] Borsley, R. D. and Rivero, M. L. (1994). Clitic auxiliaries and incorporation in Polish. *Natural Language and Linguistic Theory*, 12:373–422.
- [Bošković, 1995] Bošković, Željko. (1995). Participle movement and second position cliticization in Serbo-Croatian. *Lingua*, 96:245–266.
- [Broekhuis and Migdalski, 2003] Broekhuis, H. and Migdalski, K. (2003). Participle fronting in Bulgarian as XP-movement. *Linguistics in the Netherlands*, 20:1–12.
- [Embick and Izvorski, 1997] Embick, D. and Izvorski, R. (1997). Participle-auxiliary word orders in Slavic. In Brown, W., Kondrashova, N., and Zec, D., editors, *Formal Approaches to Slavic Linguistics: The Cornell meeting 1995*, pages 210–239, Ann Arbor, MI. Michigan Slavic Publications.
- [Homer, 2013] Homer, V. (2013). Neg-raising and positive polarity: The view from modals. To be published with *Semantics and Pragmatics*, pending minor revisions. Accessed at <http://tinyurl.com/d82c5cr> in January 2013.
- [Horn and Bayer, 1984] Horn, L. R. and Bayer, S. (1984). Short-circuited implicature: a negative contribution. *Linguistics and Philosophy*, 7:397–414.
- [Iatridou and Zeijlstra, 2013] Iatridou, S. and Zeijlstra, H. (2013). Negation, polarity and deontic modals. *Linguistic Inquiry*.
- [Rivero, 1994] Rivero, M. L. (1994). Clause structure and V-movement in the languages of the Balkans. *Natural Language and Linguistic Theory*, 12(1):63–120.
- [van der Auwera, 2001] van der Auwera, J. (2001). On the typology of negative modals. In Hoeksema, J., Rullmann, H., Sánchez-Valencia, V., and van der Wouden, T., editors, *Perspectives on negation and polarity items*, pages 23–48. John Benjamins, Amsterdam.
- [Yanovich, 2013] Yanovich, I. (2013). *Four pieces for modality, context and usage*. PhD thesis, MIT, Cambridge, MA. Directed by Kai von Fintel, Sabine Iatridou (committee chairs) and Irene Heim.

Questioning and Asserting at the Same Time: the L% Tone in A-not-A questions

Mengxi Yuan¹ and Yurie Hara²

¹ City University of Hong Kong
mxyuan2-c@my.cityu.edu.hk

² City University of Hong Kong
y.hara@cityu.edu.hk

1 Introduction

Mandarin A-not-A questions, hereafter referred to as ANAQs, have distinct syntactic and prosodic features. As shown in (1), ANAQs make both positive and negative answers explicit by conjoining the verb and its negative counterpart. Furthermore, ANAQs obligatorily end with a final low tone (see Section 2.2.2). These features distinguish ANAQs from *ma* questions (hereafter, MAQs), which lack a final low tone and only make the positive answer *p* syntactically explicit with the form *p-ma*, as shown in (2).

- (1) Lin xihuan bu xihuan Wu?
Lin like not like Wu?
'Does Lin like or not like Wu?'
- (2) Lin xihuan Wu ma?
Lin like Wu Q
'Does Lin like Wu?'

Another property of ANAQs is that they cannot be used when a possible answer to the question has been asserted. As in (3), B cannot use the ANAQ (1) when A has already provided the positive answer. Unlike the ANAQ, the MAQ (2) can be used by B in (3).

- (3) A: Lin xihuan Wu. 'Lin likes Wu.' B: #(1) ✓(2)

In this paper, we show how the syntactic and prosodic features of ANAQs derive the neutrality requirement that ANAQs can only be used in neutral contexts. ANAQs have the same question meaning as MAQs (i.e., an update in the Question Under Discussion (QUD; Roberts, 1996)) AND the meaning of an assertion of $p \vee \neg p$. The reduplication feature R of ANAQs introduces a set which contains the positive answer *p* and the negative answer $\neg p$. The sentence final particle *ne* in ANAQs adds this set onto QUD and creates a disjunction, $p \vee \neg p$. The low boundary tone L%, represented as the ASSERT morpheme, attaches to the disjunction to create an assertion of $p \vee \neg p$. The assertion of $p \vee \neg p$ indicates the speaker's ignorance about the issue *p* or $\neg p$ and thus requires a neutral context. When a possible answer is presented, the context is not neutral, therefore ANAQs cannot be used. Our analysis also predicts the similarities and differences between ANAQs and alternative questions (hereafter, ALTQs) with the form 'p or not p'. We will specify the semantics of ANAQs in Section 2, and then turn to the discussion of ANAQs and ALTQs in Section 3. Section 4 concludes this study.

2 Mandarin A-not-A questions (ANAQs)

Section 2.1 shows that ANAQs can only be used in neutral contexts. In Section 2.2, we explain this neutrality requirement by proposing that ANAQs update the QUD and make an assertion $p \vee \neg p$.

2.1 Neutrality Requirement

Our empirical observation regarding ANAQs is summarized in (4).

- (4) Unlike MAQs, ANAQs can only be used in neutral contexts, i.e., cannot be used in biased contexts.

The concept of ‘biased context’ in (4) is based on Gunlogson’s (2003) proposal, given in (5). If a proposition p is publicly asserted by one discourse participant, it is possible that both participants commit themselves to p in the end. However, there is no possibility that both participants are committed to $\neg p$. In other words, the context is biased towards p in the sense that it is possible that p is in the Common Ground (referred to as the CG, a set of propositions representing the common belief shared by all the discourse participants, Stalnaker, 1978) and it is impossible that $\neg p$ is in the CG. Therefore, the context is neutral with respect to an issue if no one publicly asserted a proposition about this issue before.

- (5) If a proposition p is publicly asserted and $\neg p$ is not publicly asserted, the context is biased towards p .
 (Modified from Gunlogson, 2003)

Let us motivate (4) with examples. In (6), no information about Lin’s feeling towards Wu has ever been mentioned before, and thus the context is neutral towards the issue whether Lin likes Wu. The speaker can use either an ANAQ or a MAQ to seek the information.

- (6) Context: Your friend arranged a blind date for Lin and Wu. After the date, you ask your friend:
 A: Lin xihuan bu xihuan Wu? (ANAQ) A': Lin xihuan Wu ma? (MAQ)

MAQs can also be used in biased contexts as seen in (3). In (3), one discourse participant A already asserted p ‘Lin likes Wu’ in the previous context. The context is biased towards p . The MAQ used by B indicates B’s doubt towards p and his request for more evidence for p . ANAQs are infelicitous in such biased contexts.

To summarize, an ANAQ can only be used in neutral contexts where no possible answers to the question have been asserted before.

2.2 Semantics of ANAQs: Deriving the Neutrality Requirement

This section derives the meaning of ANAQs compositionally from the meaning of a feature R, the sentence final particle *ne* and the low boundary tone L%.

2.2.1 The Feature R and the Particle *ne*

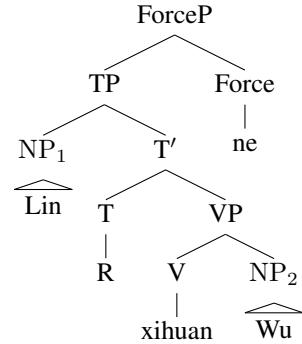
Our proposals regarding the compositional analysis of ANAQs are summarized as below:

- (7) a. The feature R, located between the subject NP and the VP in the deep structure of ANAQs, is realized by a reduplication rule.
 b. The feature R combines with the VP and the subject NP to create a set which contains a proposition p ‘NP VP’ and its negative counterpart.
 c. The sentence final particle *ne* adds this set onto QUD and creates a disjunction, $p \vee \neg p$.

Now, let us illustrate these proposals. (7-a) is based on Huang’s (1991) analysis of ANAQs. Following Huang (1991), we propose that the ANAQ in (1), repeated here as (8), is derived from the deep structure in (9). The feature R is realized by a reduplication rule, which copies a sequence following T and inserts *bu* ‘not’ between the original and its copy. Here, R copies the verb *xihuan*, leading to the structure in (8). The sentence final particle *ne* optionally occurs in ANAQs, ALTQs and *wh*-questions in Mandarin. This *ne* particle introduces the interrogative force, and thus occupies the head position of a Force Phrase (ForceP).

- (8) Lin xihuan bu xihuan Wu (ne)?
 Lin like not like Wu Q?
 ‘Does Lin like or not like Wu?’

(9)



(7-b) specifies the semantics of the feature R. We propose that the semantics of R is as in (10). According to (10), the feature R combines with the VP and the subject NP to create a set which contains a proposition and its negative counterpart, as shown in (11).

$$(10) \quad \llbracket R \rrbracket = \lambda P. \lambda x. \{P(x), \neg P(x)\}$$

$$(11) \quad \llbracket TP \rrbracket = \llbracket R(\text{like.Wu})(\text{Lin}) \rrbracket = \{ \text{‘Lin likes Wu’}, \text{‘Lin does not like Wu’} \}$$

(7-c) characterizes the semantics of the particle *ne*. The particle *ne* introduces an interrogative force head Q_1 . As in (12), the semantics of Q_1 consists of two formulae: 1) a primary formula $\lambda S. \lambda C. [\text{QUD}(C) + S]$, which takes in a set of propositions (Hamblin, 1973) and then adds this set onto the QUD à la Roberts. ‘+’ is an update function which adds a set of propositions to a discourse context (Heim, 1982). As a result, $\text{QUD}(C) + S$ is a context that resembles $\text{QUD}(C)$ except that $\text{QUD}(C) + S$ now contains S as the most immediate question under discussion. 2) a secondary formula $\lambda S. (r_1 \vee r_2 \vee \dots \vee r_{|S|})$, which takes in a set of propositions and yields the disjunction of all the propositions in the set.

$$(12) \quad \llbracket Q_1 \rrbracket = \lambda S. \lambda C. [\text{QUD}(C) + S] \times \lambda S. (r_1 \vee r_2 \vee \dots \vee r_{|S|}), r_i \in S \text{ for all } 1 < i \leq |S|$$

(Notation: If p and q are formulae, $p \times q$ is a formula, where p is the primary formula, q is the secondary formula; S is of type $\langle \langle s, t \rangle, t \rangle$.)

The interpretation of the ForceP in (9) is given in (13).

$$(13) \quad \llbracket \text{ForceP} \rrbracket = \llbracket Q_1 \rrbracket (\llbracket \text{TP} \rrbracket) = \lambda C. [\text{QUD}(C) + \{p, \neg p\}] \times (p \vee \neg p) \quad p = \text{‘Lin likes Wu’}$$

In summary, the feature R produces a set of propositions. The particle *ne* adds this set onto QUD, i.e., produces a question meaning, and then creates a disjunction of the propositions in the set.

2.2.2 The low boundary tone L% in ANAQs

An ANAQ is made up of a ForceP and the final low tone. The last section derived the meaning of the ForceP. Now we need to obtain the meaning of the final low tone. To see this, this section compares the intonation of ANAQs with the intonations of declaratives, ALTQs and MAQs.

Shen (1990) points out that MAQs end with a final rise tone, while ANAQs end with a final fall tone, which is the characteristic of declarative intonation. This distinction is depicted in Figure 1, which is a summary of Shen (1990) given by Schack (2000). Shen (1990) also concludes that ALTQs and Mandarin *wh*-questions share the same intonation pattern with ANAQs and end with a final low tone.

Shen’s (1990) conclusion is supported by our case study of a Mandarin speaker, who is the first author of this paper. We recorded four utterances: a declarative, a MAQ, an ANAQ and an ALTQ, and analyzed them in the Pan-Mandarin ToBI system (Peng et al., 2005). In the test sentence *Wulin na yinyu* ‘Wuli carries silver-fish’, each syllable is pronounced with tone 2 (the mid-rising tone, labeled as 35).

As can be seen from Figure 2 to Figure 5,¹ ANAQs and ALTQs end with the low boundary tone L%, just like the declarative. In contrast, MAQs end with the high boundary tone H%.

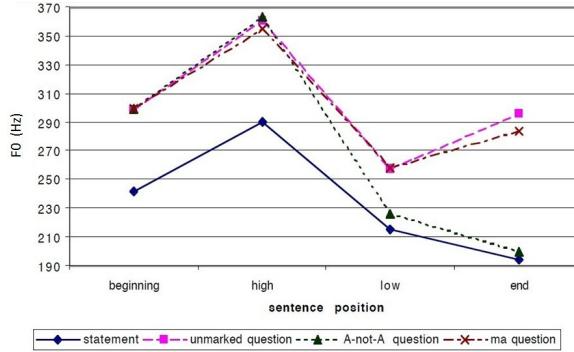


Figure 1: The average F0 in Shen's study

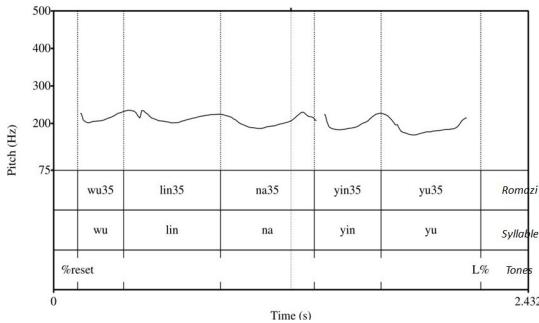


Figure 2: Declarative ‘Wulin carries silver-fish.’

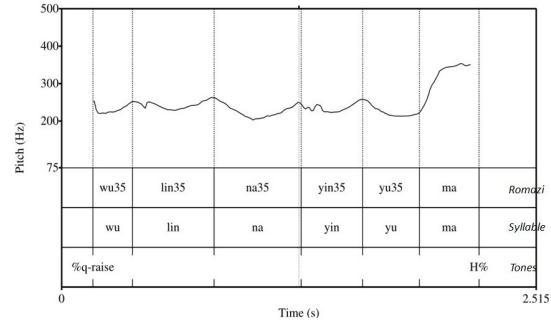


Figure 3: MAQ ‘Does Wulin carry silver-fish?’

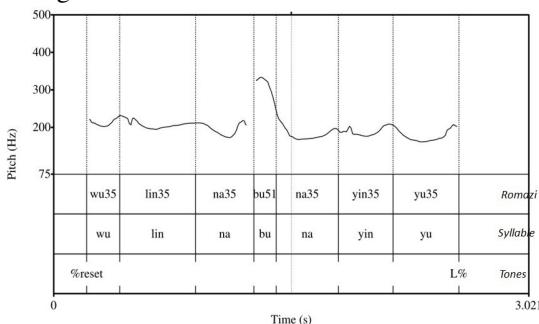


Figure 4: ANAQ ‘Does Wulin carry or not carry silver-fish?’

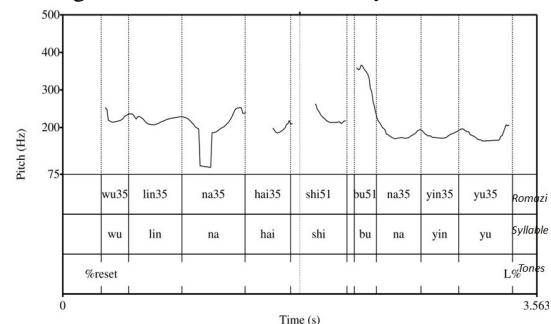


Figure 5: ALTQ ‘Does Wulin carry or not carry silver-fish?’

In a nutshell, declaratives, ANAQs and ALTQs are marked by the low boundary tone L%, whereas MAQs are marked by the high boundary tone H%.

2.2.3 L% as ASSERT and paratactic association

In this section, we propose that the low boundary tone L% provides an assertive force. We show how the paratactic association of L% to the ForceP create an assertion ‘ $p \vee \neg p$ ’, and derive the speaker’s ignorance towards the issue p or not p from this assertion.

¹Since we are only interested in boundary tones, we adopt only three tiers: Romanisation, Syllables and Tones, out of the seven tiers in the Pan-Mandarin ToBI annotation system.

Our proposals regarding the semantics of ANAQs are summarized as below:

- (14) a. L% in ANAQs and declaratives represents the abstract ASSERT morpheme.
b. ASSERT is paratactically attached to the secondary formula ' $p \vee \neg p$ ' to produce an assertion ' $p \vee \neg p$ '.
c. The assertion ' $p \vee \neg p$ ' indicates the speaker's ignorance and thus requires a neutral context.

First, let us illustrate (14-a) and (14-b). Our idea of associating the meaning of L% with the meaning of ForceP is inspired by Bartels (1997). Bartels (1997) proposes that English ALTQs end with the low phrasal tone L- (See also Pruitt & Roelofsen, to appear), and that L- tone represents the abstract ASSERT morpheme. ASSERT is paratactically associated with ALTQs and performs the dynamic assertive update.

Following Bartels (1997), we propose that the low boundary tone L% in Mandarin declaratives, ANAQs and ALTQs represents the ASSERT morpheme. ASSERT is attached to ANAQs by paratactic association, that is, ASSERT is not integrated with the sentence syntactically, but paratactically attached to either of the two formulae in (13). Since ASSERT is a force head of type $\langle\langle s, t \rangle, \langle c, c \rangle\rangle$, ASSERT should be attached to the secondary formula $p \vee \neg p$ (of $\langle s, t \rangle$ type) rather than the primary formula $\lambda C. [QUD(C)+\{p, \neg p\}]$ (of $\langle c, c \rangle$ type).

After the paratactic association of the low boundary tone, we obtain the semantics of (8) as in (15). The semantics of an ANAQ consists of two parts: 1) the primary meaning $\lambda C. [QUD(C)+\{p, \neg p\}]$, which updates the QUD with the set $\{p, \neg p\}$; 2) the secondary meaning, i.e., an assertion ' $p \vee \neg p$ '.

(15) $\lambda C. [QUD(C)+\{p, \neg p\}] \times \text{ASSERT}(p \vee \neg p)$ $p = \text{'Lin likes Wu'}$

Now, let us expand the proposal in (14-c), i.e., derive the speaker's ignorance from the assertion ' $p \vee \neg p$ '. The proposition ' $p \vee \neg p$ ' is a tautology, i.e., the informativeness of this assertion is zero since it is always true. When the speaker asserts a tautology such as (16), the speaker can have at least three possible states of mind: First, the speaker in fact knows that Lin likes Wu (or knows that Lin does not like Wu), but does not want to provide the addressee with this information.

(16) Lin likes Wu or Lin does not like Wu.

Second, the speaker does not care whether Lin likes Wu or not. In other words, the speaker is indifferent to the issue.² Third, the speaker has no idea if Lin likes Wu or not, i.e., the speaker is ignorant about this issue. In case of ANAQs, the first two possibilities should be ruled out due to the incompatibility with the question meaning of ANAQs; thus, the assertion ' $p \vee \neg p$ ' in ANAQs indicates the speaker's ignorance.

The first possibility that the speaker knows p but conceals this information is eliminated, because the speaker's knowledge about p would make the question act of ANAQs infelicitous. In Searle's (1969) theory of felicity conditions, an illocutionary act is felicitous only when it meets a set of conditions. For a question to be felicitous, it has to meet the following conditions:

- (17) a. Preparatory: $S [= \text{speaker}]$ does not know the answer to the question.
b. Sincerity: S wants the missing information. (Modified from Searle, 1969: 66)

²For example, in an unconditional like (i), the speaker indicates that whether Lin likes Wu or not does not matter to him.

(i) Whether Lin likes Wu or Lin does not like Wu, I recommend Wu to be our chairman.

Our analysis, together with Hamlin's (1973) pointwise application, ensures the right interpretation of A-not-A constructions in conditionals, as in (ii). The particle *ne* and L% do not appear in (ii). Only the feature R makes its contribution. Thus, *xia bu xiayu* indicates a set which contains the two alternatives: p 'it rains' and $\neg p$ 'It doesn't rain'. See also Rawlins (2008) for more discussions about conditionals.

(ii) Mingtian xia bu xiayu wo dou yao qu yuanzu.
tomorrow rain not rain I DOU will go hike
'Whether it rains or not, I will go hiking tomorrow.'

(iii) $[\![\text{xia bu xiayu}]\!] = \{p, \neg p\}$ $p = \text{'it rains'}$

Suppose that the speaker knew p. Then, the question act would be infelicitous, since both the preparatory condition and the sincerity condition are not met: the speaker knows the answer to the question, and the speaker does not want the information. Thus, the first possibility does not arise in case of ANAQs.

Now, let us see why the second possibility is eliminated. Suppose that the assertion ‘ $p \vee \neg p$ ’ represented the speaker’s indifference towards the issue p or not p. Then, the meaning of an ANAQ would be contradictory: according to the primary meaning $\lambda c. [QUD(c) + \{p, \neg p\}]$ of an ANAQ, the speaker adds the question onto QUD, indicating that the speaker is interested in the issue and thus seeks the answer. However, the assertion meaning says that the speaker is indifferent to this issue. Therefore, the indifferent reading of a tautology is incompatible with ANAQs.

Finally, the third interpretation, the speaker’s ignorance towards the issue, is compatible with the question meaning (i.e., the speaker is interested in the issue). Therefore, the assertion ‘ $p \vee \neg p$ ’ in ANAQs indicates that the speaker is ignorant about the issue whether p or $\neg p$. In other words, the speaker is claiming that the probabilities of p and $\neg p$ are the same in context c. Therefore, the assertion ‘ $p \vee \neg p$ ’ indicates $P_c(p) = P_c(\neg p)$, i.e., the context is neutral. Thus, we rewrite (15) as (18). (18) shows that the speaker is ignorant towards the issue p or not p but not indifferent.

- (18) The semantics of ANAQs: $\lambda c. [QUD(c) + \{p, \neg p\}] \times [P_c(p) = P_c(\neg p)]$
(P_c is a probability function that takes a proposition and returns its probability in context c.)

The assertion meaning of ANAQs, i.e., $P_c(p) = P_c(\neg p)$, explains why ANAQs cannot occur in biased contexts. In biased contexts where one possible answer to the question has already been asserted, the context is biased towards p, i.e., the probability of p is larger than the probability of $\neg p$ in context c ($P_c(p) > P_c(\neg p)$). This contradicts the assertion meaning of ANAQs ($P_c(p) = P_c(\neg p)$). Unlike ANAQs, MAQs simply indicate an update in the QUD.³ MAQs lack the meaning of the assertion ‘ $p \vee \neg p$ ’ and thus can occur when an answer has already been provided, as we have seen in Section 2.1.

In this section, we account for the neutrality requirement of ANAQs by proposing that ANAQs make an assertion $p \vee \neg p$ and that this assertion indicates the speaker’s ignorance. This assertion meaning is derived from the paratactic association of the low boundary tone L% with the ForceP.

3 ANAQs and alternative questions (ALTQs)

This section compares ANAQs with ALTQs in the form of ‘p or not p’. We briefly show that following Levinson’s (2000) M-principle, ALTQs with the form ‘p or not p’ can only be used in marked contexts.

A Mandarin ALTQ connecting two contradictory alternatives, as in (19), resembles an ANAQ except that *haishi* ‘or’ does not appear in the ANAQ. The disjunction *haishi* is used in questions and conditionals. The other disjunction *huo(zhe)* ‘or’ only occurs in declaratives, as in (20).

- (19) Lin xihuan haishi bu xihuan Wu (ne)? (20) Lin xihuan Wu huoze Zhang.
Lin like or not like Wu Q Lin like Wu or Zhang
'Does Lin like or not like Wu?' 'Lin likes Wu or Zhang.'

Like ANAQs, ALTQs cannot be used in biased contexts like (3). ALTQs with the form ‘p or not p’ are always used in forceful neutral contexts in which the speaker failed to get an answer before and is eager to get the answer this time. ANAQs are also felicitous in such contexts. As in (21), Li asked the question before but hasn’t got the answer due to Xiaoming’s contradictory statements. Now, Li can use either an ALTQ or an ANAQ to force the addressee to provide the answer immediately (See also Biezma, 2009).

³In Yuan and Hara (to appear), we propose that a MAQ introduces an interrogative force head Q_2 , construed as a function from a proposition (of $\langle s, t \rangle$ type) to a CCP (of $\langle c, c \rangle$ type): $\llbracket Q_2 \rrbracket = \lambda p. \lambda c. [QUD(c) + \{p, \neg p\}]$.

- (21) Li asks Xiaoming if he goes hiking tomorrow. Xiaoming says “Yes, I will go. Ah, I haven’t finished my homework”. Then, Li asks Xiaoming:
- | | |
|------------------------------|------------------------------|
| A: Ni qu haishi bu qu? | Ni qu bu qu? |
| you go or not go | you go not go |
| ‘Do you go or not go?’(ALTQ) | ‘Do you go or not go?’(ANAQ) |

The difference between these two questions is that ALTQs with the form ‘p or not p’ are infelicitous in examples like (22) whereas ANAQs are felicitous. (22) is a normal neutral context, but not a forceful context. It is impolite to use an ALTQ in (22), since the ALTQ with the form ‘p or not p’ indicates that the speaker is forcing the addressee to pick an answer from p and $\neg p$. The ANAQ does not have such a forceful connotation, and is felicitous in such a context.

- (22) Context: You tell your friend that you are going hiking this weekend. Then, you ask him:
 A: #Ni qu haishi bu qu?(ALTQ) ✓Ni qu bu qu?(ANAQ)

ALTQs with the form ‘p or not p’ cannot occur in biased contexts since they involve an assertion meaning, just like ANAQs. We propose that *haishi* collects two or more propositions and creates a set which contains these propositions,⁴ as in (23). In p-or-not-p ALTQs, *haishi* creates a set $\{p, \neg p\}$. The particle *ne* adds this set onto QUD and creates a disjunction $p \vee \neg p$. L% is paratactically associated to the sentence to produce an assertion ‘ $p \vee \neg p$ ’. Thus, the p-or-not-p ALTQs have the same semantics as ANAQs, as in (24).⁵ Both indicate the speaker’s ignorance and require a neutral context.

- (23) $\llbracket \text{[}haishi\text{]} \rrbracket = \lambda p_1 \dots \lambda p_n. \{p_1, \dots, p_n\}, n \geq 2$.
- (24) $\lambda C. [\text{QUD}(C) + \{p \vee \neg p\}] \times \text{ASSERT}(p \vee \neg p)$

ALTQs with the form ‘p or not p’ can only be used in forceful neutral contexts, because they are marked expression compared with ANAQs and are thus infelicitous in unmarked contexts such as the normal neutral context. According to Levinson’s M-principle (2000), the speaker cannot use a marked expression without reason. In other words, an abnormal and marked expression implicates a marked meaning and thus is used in a marked context. A speaker would violate the M-principle if he used a marked expression in an unmarked context. We speculate that ANAQs are unmarked expressions whereas ALTQs with the form ‘p or not p’ are marked.⁶ Asking a question forcefully is a marked context compared with asking a question normally. Therefore, marked expressions, i.e., ALTQs with the form ‘p or not p’, are appropriate for marked forceful contexts. A speaker would violate the M-principle if he used ALTQs with the form ‘p or not p’ in an unmarked context like the normal neutral context.

This section discusses the differences between ALTQs and ANAQs. To avoid violating the M-principle, ALTQs with the form ‘p or not p’ can only be used in forceful neutral contexts.

4 Conclusion

In this paper, we presented a compositional semantics for Mandarin ANAQs. In ANAQs, the reduplication feature R produces a set which contains a proposition p and $\neg p$. The question operator *ne* combines with this set to update the QUD and to yield a proposition $p \vee \neg p$. The low boundary tone L%, represented as the abstract ASSERT morpheme, attaches to $p \vee \neg p$ to produce the an assertion ‘ $p \vee \neg p$ ’. This assertion

⁴*Huozhe* collects two or more propositions and yields the disjunction of these propositions.

⁵See Erlewine (to appear) for a similar derivation of the semantics of ALTQs.

⁶ANAQs are more unmarked than ALTQs with the form ‘p or not p’, since: 1) the former takes a shorter form than the latter; 2) the former can be used in more contexts than the latter; 3) the former is used much more frequently than the latter. We found 61320 ANAQs and only 193 ALTQs with the form ‘p or not p’ in the Modern Chinese Corpus of Center for Chinese Linguistics.

indicates that the speaker is ignorant towards the issue p or not p and thus requires that the context be neutral. Therefore, ANAQs cannot be used in biased contexts in which a possible answer has been asserted. ANAQs express two speech acts at the same time, i.e., assertion act and question act.

Our proposal accounts for the similarities and differences between ANAQs, MAQs and ALTQs. MAQs do not have an assertion meaning and thus can be used in biased contexts. ALTQs with the form ‘ p or not p ’ are more marked than ANAQs. Following the M-principle, ALTQs with the form ‘ p or not p ’ can only be used in a marked context, i.e., the forceful neutral context.

References

- [1] Christine Bartels. *Towards a compositional interpretation of English statement and question intonation*. PhD thesis, University of Massachusetts, 1997.
- [2] Maria Biezma. Alternative vs. polar questions: The cornering effect. In *Proceedings of SALT 19*, 2009.
- [3] Maria Biezma and Kyle Rawlins. Responding to alternative and polar questions. *Linguist and Philos*, 35:361–406, 2012.
- [4] Michael Yoshitaka Erlewine. Alternative questions through focus alternatives in Mandarin Chinese. Proceedings of the 48th Annual Meeting of the Chicago Linguistic Society (CLS 48), To appear.
- [5] Christine Gunlogson. *True to form: Rising and falling declaratives as questions in English*. New York: Routledge., 2003.
- [6] Charles Hamblin. Questions in montague grammar. *Foundations of Language*, 10:41–53, 1973.
- [7] Irene Heim. *The semantics of definite and indefinite noun phrases*. PhD thesis, University of Massachusetts, Amherst., 1982.
- [8] C.-T. James Huang. Modularity and Chinese A-not-A questions. In Carol Georgopolous and Robert Ishihara, editors, *Interdisciplinary Approaches to Language*, pages 305–322. Dordrecht: Kluwer, 1991.
- [9] Stephen C. Levinson. *Presumptive meanings: the theory of generalized conversational implicature*. MIT Press, 2000.
- [10] S. Peng, M. Chan, C. Tseng, T. Huang, O. Lee, and M. Beckman. Towards a Pan-Mandarin system for prosodic transcription. In Sun-Ah Jun, editor, *Prosodic Typology: The Phonology of Intonation and Phrasing*, pages 230–270. Oxford University Press, 2005.
- [11] Kathryn Pruitt and Floris Roelofsen. Interpretation of prosody in disjunctive questions. *Linguistic Inquiry*, To appear.
- [12] Kyle Rawlins. *(Un)conditionals: An investigation in the syntax and semantics of conditional structures*. PhD thesis, University of California, Santa Cruz, 2008.
- [13] Craige Roberts. Information structure: Towards an integrated formal theory of pragmatics. In Jae Hak Yoon and Andreas Kathol, editors, *OSUWPL Volume 49: Papers in Semantics*, pages 305–322. Dordrecht: Kluwer, 1996.
- [14] Katrina Schack. Comparison of intonation patterns in Mandarin and English for a particular speaker. *University of Rochester Working Papers in the Language Sciences*, 1:24–55, 2000.
- [15] John Searle. *Speech acts: an essay in the philosophy of language*. Cambridge, England: Cambridge University, 1969.
- [16] Xiao-Nan Shen. *The prosody of Mandarin Chinese*. Berkeley, CA: University of California Press, 1999.
- [17] Robert Stalnaker. Assertion. *Syntax and Semantics*, 9:315–332, 1978.
- [18] Mengxi Yuan and Yurie Hara. The semantics of the two kinds of questions in Mandarin: a case study of discourse adverbs. Proceedings of The 44th Meeting of the North East Linguistic Society, To appear.

Universal Quantifier PPIs

Hedde Zeijlstra

Georg-August-University Göttingen, Germany
hzeijls@uni-goettingen.de

Abstract

Why have Positive Polarity Items (PPIs) that are universal quantifiers only been attested in the domain of modal auxiliaries (cf. Homer t.a., Iatridou & Zeijlstra 2010, 2013) and never in the domain of quantifiers over individuals? No PPI meaning *everybody* or *everything* has ever been reported. In this paper, I argue that universal quantifier PPIs actually do exist, both in the domain of quantifiers over individuals and in the domain of quantifiers over possible worlds, as, I argue, is predicted by the Kadmon & Landman (1993) – Krifka (1995) – Chierchia (2006, 2013) approach to NPI-hood. However, since the covert exhaustifier that according to Chierchia (2006, 2013) is induced by these PPIs (and responsible for their PPI-hood) can act as an intervener between the PPI and its anti-licenser, it is concluded in this paper that a universal quantifier PPIs may scope below it and thus appear in disguise; their PPI-like behaviour only becomes visible once they morpho-syntactically precede their anti-licenser. Another conclusion of this paper is that Dutch *iedereen* (‘everybody’), opposite to English *everybody*, is actually a PPI.

1 Introduction

Following recent lines of thinking (Kadmon & Landman 1993, Krifka 1995 and Chierchia , 2013), Negative Polarity Items (NPIs) are only fine in Downward Entailing contexts, since outside such contexts their semantics would give rise to a contradiction. According to Chierchia’s (2006, 2013) implementation, this is due to the fact that NPIs are equipped with a feature $[\sigma]$ that ensures (i) obligatory introduction of domain alternatives and (ii) that the proposition that this NPI is part of must be covertly exhaustified. To see this, take (1).

- (1) *I read any book

For Chierchia, the underlying syntax of (1) is (2).

- (2) $[\text{EXH}_{[\sigma]} [\text{I read } [\text{any book}]_{[\sigma]}]]$

Suppose that the domain quantification is the set of books $\{a,b,c\}$. Then $[[\text{I read any book}]]$ denotes $\exists x.[x \in \{a,b,c\} \ \& \ \text{read}(I, x)]$. Domain alternatives of $[[\text{I read any book}]]$ are then, for instance, $\exists x.[x \in \{a,b\} \ \& \ \text{read}(I, x)]$ or $\exists x.[x \in \{c\} \ \& \ \text{read}(I, x)]$. Now, if (1) is obligatorily exhaustified, as in (2), all stronger domain alternatives, like the ones above, are false. But if that is the case, then $[[\text{(2)}]]$ has the denotation in (3), which forms a logical contradiction and rules out the sentence. Since in Downward Entailing (DE) contexts logical entailments relations are reversed, all stronger alternatives become weaker, and therefore no contradiction arises.

- (3) $\exists x.[x \in \{a,b,c\} \wedge \text{read}(I, x)] \wedge \neg \exists x.[x \in \{a,b,c\} \wedge \text{read}(I, x)]$

But no language in the world seems to have a word meaning *all*, *everybody* or *everything* that is a PPI. Within the domain of quantifiers over individuals, most PPIs are actually existential quantifiers (e.g. English *some*), never universal quantifiers. This would suggest that for some unknown reason the approach by Kadmon & Landman, Krifka and Chierchia would not extend to universals.

However, in the domain of modals, universal quantifier PPIs are indeed attested. As has been pointed out by Israel (1996), Iatridou & Zeijlstra (2010, 2013) and Homer (t.a.) universal modals that take wide scope with respect to negation, like English *must*, *should* or *ought to*, are indeed PPIs. That they are PPIs ensure that these modals outscope negation.

- | | | |
|-----|------------------------|------------------|
| (4) | She must not leave | $\square > \neg$ |
| (5) | She should not leave | $\square > \neg$ |
| (6) | She ought not to leave | $\square > \neg$ |

The fact that these modals precede negation at the surface does not guarantee their scopal behaviour with respect to negation; other modals, e.g. deontic can or may, scope under negation, even though they surface in a higher position.

- | | | |
|-----|-------------------|-------------------|
| (7) | She cannot leave | $\neg > \Diamond$ |
| (8) | She may not leave | $\neg > \Diamond$ |

That it is the PPI-hood of modals like *must*, *should* and *ought to* that is responsible for the scopal differences between (4)-(6) as oppose to (7)-(8) is well established in Iatridou & Zeijlstra (2010, 2013) and Homer (t.a.) and I refer the reader to those articles for concrete evidence for the PPI status of such universal modals.

The existence of such universal PPI modals thus forms evidence in favour of the approach that takes polarity effects to result from logical contradictions: the predicted elements are indeed attested. But it gives rise to a new question as well: why have universal quantifier PPIs only been attested in the domain of modal auxiliaries and never in the domain of quantifiers over individuals? It is this question that I address in this paper.

2 Modal Universal PPIs

Iatridou & Zeijlstra (2013) briefly point out to what extent current approaches to NPI- and PPI-hood (Kadmon & Landman 1993, Krifka 1995, Postal 2000, Szabolcsi 2004, Chierchia 2006, 2013) may apply to account for the PPI-hood of these universal quantifier modals. They do, however, not provide any account themselves (though see Iatridou & Zeijlstra 2013: endnote xliii for a sketch). Rather they state in general terms that the reason why certain existentials/indefinites may be prone to become NPIs should extend to universals being prone to becoming PPIs.

To see this, take (9), which has the semantics in (10):

- | | |
|------|---|
| (9) | (according to the law,) John must leave |
| (10) | $\forall w[\text{the law is satisfied in } w \rightarrow \text{John leaves in } w]$ |

Now, assume that *must* carries a feature $[\sigma]$ too that obligatorily introduces domain alternatives in its first argument and requires it to be checked against a higher, c-commanding covert exhaustifier. This treatment of *must* is fully analogous to the treatment of NPIs, such as *any*, by Chierchia (2006,

2013). Then, the syntax of (9) is as in (11):

- (11) [EXH_[σ] [John must_[σ] leave]]

Now, think of a model where the only relevant worlds are w1, w2 and w3. Saying that each of w1, w2 and w3 is a world where John leaves provides a stronger statement than saying that John only leaves in a subdomain of these worlds, for instance in w1 and w2 only. Therefore, exhaustifying the *must*-clause in a positive context has no semantic effect. The semantics of (11) is still the one in (10).

However, once *must* is put under negation, things change dramatically:

- (12) John must not leave.

Now the question arises as to which takes widest scope, the modal or negation. Let us start with the case where the modal scopes under negation, and moreover, both scope under the required exhaustifier:

- (13) EXH > NEG > MUST

Interpreting (12) with the scopal order of (13) has as result that the set {w1, w2, w3} is not a subset of the set of worlds where John leaves, since the universal quantifier of word variables is under the scope of negation. At the same time, all domain alternative expressions of this assertion are stronger: for instance, saying that {w1, w2} is not a subset of the set of worlds where John leaves makes a stronger statement than the original assertion. Therefore, the proposition that {w1, w2} is not a subset of the set of worlds where John leaves must be negated, which in turn entails that w1 and w2 are worlds where John leaves. Since the same mechanism applies to all subdomains of {w1, w2, w3}, including {w1}, {w2} and {w3}, interpreting the modal under the scope of negation with the strengthening operator applying above it yields a contradiction.

This contradiction disappears once the modal takes scope above negation again, since expressions of the form “must (not (p))” are stronger than their alternatives and will therefore not be contradicted by negated stronger alternatives. Hence, by applying the [σ] feature mechanism to a universal modal quantifier, this quantifier becomes a PPI in exactly the same way as existentials/indefinites become NPIs when equipped with such a feature. At first sight, this opens up a way to account for the PPI-hood of those universal modals that generally outscope negation. However, if this feature [σ] can be applied to universal modal quantifiers (i.e. universal quantifiers over possible worlds), the question immediately arises as to why it cannot be applied to other universal quantifiers, such as quantifiers over individuals. After all, nothing in the above analysis hinges on the fact that these quantifiers bind possible worlds, so by the same logic a universal quantifier (*all*, *everybody* or *everything*) that carries a feature [σ] should be able to be a PPI.

To see this, take the imaginary word *pevery* that would be semantically identical to English *every* apart from being equipped with a feature [σ]. A negative sentence containing *pevery*, like (14), would have the syntax as in (15) and therefore the denotation as in (16), a clear contradiction.

- (14) I did not read *pevery* book

- (15) [EXH_[σ] [I did not read [pevery book]_[σ]]]

- (16) $\neg \forall x. [x \in \{a, b, c\} \rightarrow \text{read}(I, x)] \& \neg \neg \forall x. [x \in \{a, b, c\} \rightarrow \text{read}(I, x)]$

But, as of yet, such non-modal universal quantifiers PPIs have not been attested.

3 Non-modal Universal PPIs

3.1 Proposal

In this section I argue that the reason why only universal PPIs have been attested among quantifiers over possible worlds and not among quantifiers over individuals lies in their syntactic differences rather than in their semantic differences, in particular in their syntactic position in the sentence. More concretely, I argue that both universal modals and universal quantifiers over individuals with a feature $[\sigma]$ can actually be attested, but that the syntactic properties of universal quantifiers over individuals with such a feature may obscure their diagnostic PPI properties. To see, this, take again the scopal ordering of a universal quantifier with a feature $[\sigma]$, negation and the covert exhaustifier that gives rise to the logical contradiction, the ordering in (17):

$$(17) \quad ^\# \dots \text{EXH} > \text{NEG} > \forall_{[\sigma]}$$

If negation intervenes between the exhaustifier and the universal, a contraction arises. But nothing guarantees that a universal quantifier with a feature $[\sigma]$ (henceforward $\forall_{[\sigma]}$) has its exhaustifier scope higher than the negation: the feature $[\sigma]$ only requires that the exhaustifier c-commands the $\forall_{[\sigma]}$ and therefore has scope over it, but does not require that it has immediate scope. An alternative underlying syntactic configuration for such a universal quantifier, *pevery*, carrying a feature $[\sigma]$ would be (18).

$$(18) \quad [\text{NOT} [\text{EXH}_{[\sigma]} [\text{I read } [\text{pevery book}]_{[\sigma}]]]]$$

But (18) does not give rise to a logical contradiction. In (18) the proposition *I read pevery book*, denoting $\forall x.[x \in \{a,b,c\} \rightarrow \text{read}(I, x)]$, would be exhaustified (a vacuous operation, since it is already stronger than any of its alternatives) *before* it gets negated. The denotation of (18) is then just simply (19). The exhaustifier acts as an intervener.

$$(19) \quad \neg \forall x.[x \in \{a,b,c\} \rightarrow \text{read}(I, x)]$$

Consequently, a universal PPI (or to be more precise: a universal quantifier that obligatorily introduces domain alternatives and that must be exhaustified) is fine in a negative / downward entailing context as long as the exhaustifier is in between the negation or any other downward entailing operator and the universal quantifier itself. Universal quantifier PPIs may thus appear under negation without being ungrammatical and therefore be unrecognizable as such.

3.2 Universal quantifiers and self-intervention effects

How do we know then if such universal quantifier PPIs that bind individuals exist in the first place. And why is it that we do find clear modal universal PPIs? Since the recognisability of universal PPIs depends on the possibility of an intervening EXH, these questions arise as to exactly when EXH can intervene. Since, EXH is a covert operator whose present depends on an overt marker of it, the element carrying the uninterpretable feature $[\sigma]$ that must be checked against the interpretable $[\sigma]$ feature on EXH, this question depends on when and where such covert operators may be posited in general. Zeijlstra (2012) has argued for a number of different phenomena that covert operators must be included in a position c-commanding the highest overt marker(s) of an abstract operator. If this is correct, then the only restriction on the position of EXH in a sentence containing a universal quantifier

carrying uninterpretable $[\sigma]$ ($\forall_{[\sigma]}$ henceforward), is that it c-commands the surface realization of $\forall_{[\sigma]}$. This entails that the only orders where $\forall_{[\sigma]}$ may not appear under the scope of negation are exactly those cases where either $\forall_{[\sigma]}$ precedes negation or where it forms a morpho-syntactic unit with it; in both cases it is not possible for EXH to c-command $\forall_{[\sigma]}$ while still be under the scope of negation.

This applies to both types of modals in (20) and (21). In (20) negation is below the modal (and therefore already under its scope) and the modal is in turn c-commanded and outscoped by EXH. In (21) the negative marker is attached to the modal. Since EXH cannot intervene between negation or be inserted below the modal, it must be in a position c-commanding both of them. However, this would already give rise to a contradiction, since the scopal order would then be EXH>NEG> $\forall_{[\sigma]}$, yielding a contradiction. The only way to rescue this sentence is then to first have the modal undergo QR (along the lines of Iatridou & Zeijlstra 2013) and then have it c-commanded by EXH. In both (20)-(21) the scopal order can only be EXH > $\forall_{[\sigma]}$ > NEG, which gives rise to the PPI-effect.

- (20) John mustn't leave

- (21) Juan no-debe ir
Juan neg-must go
'Juan mustn't go'

By contrast, if a modal would appear (sufficiently lower) than negation, it is predicted that EXH might intervene between the negation and the modal. This is indeed the case. In a language like Dutch in main clauses a modal must precede negation (due the verb second property of the language), but in subordinate clauses the verb remains in situ and follows the negation. The prediction that this analysis makes is that modals that must take scope over negation in main clauses (due to their $[\sigma]$ feature), should be able to take scope both below and above negation in a subordinate clause. In a subordinate clause EXH can either be posited in between negation and the modal, or above negation. In the former case, the modal may scope under negation; in the later case, it may not and must raise across the negation. This prediction is indeed born out, as the data in (22)-(25) show.

- (22)*Jan moet niet vertrekken, maar het mag wel

Jan must neg leave, but it may prt
'John mustn't leave, but it is allowed'

- (23)Jan moet niet vertrekken, omdat het verboden is

Jan must neg leave, because it forbidden is
'John mustn't leave, because it is forbidden'

- (24)Ik weet dat Jan niet moet vertrekken, maar dat het wel mag

I know that Jan neg must leave, but that it prt may
'I know that John doesn't have to leave, but that it is allowed'

- (25)Ik weet dat Jan niet moet vertrekken, omdat het verboden is

I know that Jan neg must leave, because it forbidden is
'I know that John mustn't leave, because it is forbidden'

Applying this to universal quantifiers over individuals, it is entailed that those quantifiers that follow negation, can also be under the scope of negation or any other proper anti-liscenser, since EXH is able to intervene. On the basis of those cases one cannot even tell whether the universal quantifier carries a feature $[\sigma]$ or not.

Only on the basis of examples where a morphologically independent negation precedes a universal quantifier one cannot tell whether a universal quantifier like *everybody* is a PPI or not by investigating its scopal behavior when it precedes negation. In that case the surface scope order would be EXH > $\forall_{[\sigma]} >$ NEG. Under such a configuration, the universal quantifier that is equipped with a feature $[\sigma]$ cannot reconstruct below negation (as this would give rise to a logical contradiction), but a universal quantifier that is lacking $[\sigma]$, would be able to reconstruct below negation. Interestingly, variation between universal quantifiers that may and that may not reconstruct under negation when preceding it at surface structure, has indeed been attested (and never been properly explained). In the remainder of this paper I show that the only distinction between such quantifiers is the presence or the absence of a feature $[\sigma]$ on \forall .

4 \forall reconstruction

Following this line of reasoning, we can actually establish that English *everybody* is not a PPI, but that Dutch *iedereen* ('everybody') is a PPI, a novel observation to the best of my knowledge. In English (and most other languages), for almost all speakers a universal quantifier that precedes negation may reconstruct under negation. This shows that English everybody cannot carry a feature $[\sigma]$.

- (26) Everybody didn't leave $\forall > \neg; \neg > \forall$

However, for most speakers of Dutch (and several Northern German varieties), this reconstructed reading is not available (cf. Zeijlstra 2004, Abels & Marti 2011). This observation has never received a satisfactory explanation, but directly follows once universal quantifiers in Dutch are taken to be PPIs.

- (27) Iedereen vertrok niet (Everybody left not) $\forall > \neg; * \neg > \forall$
 Everybody left not
 'Nobody left'

If *iedereen* is a PPI, it must be c-commanded by EXH at surface structure and reconstructing it below negation would result in the contradictory reading EXH>NEG> $\forall_{[\sigma]}$, thus providing a simple solution for this hitherto unsolved problem. Moreover, if *iedereen* is a PPI the prediction that Chierchia's analysis makes with respect to universal quantifiers, namely that universal quantifiers equipped with a feature $[\sigma]$ should also be attested (since nothing principled rules them out, is also confirmed).

However, how can we independently investigate whether Dutch *iedereen* is indeed a PPI? As pointed out by Szabolcsi (2004), PPI-hood can be diagnosed in four different ways. First, PPIS should be fine under metalinguistic negation. This is indeed the case for Dutch *iedereen*, which may take scope under metalinguistic negation:

- (28) Speaker A: Iedereen moet de kamer uit
 Everybody must the room out
 'Everybody must leave the room'
 Speaker B: Nee, onzin. Iedereen moet niet de kamer uit; alleen Jan en Piet
 No, nonsense. Everybody must neg the room out; only Jan and Piet
 'No, nonsense. Everybody mustn't leave the room, only John and Piet
 must'

Also, PPIs can take scope under clause-external negation. Again this applies to *iedereen* as well.

- (29) Ik zeg niet dat iedereen moet vertrekken; alleen Jan moet vertrekken
I say not that everybody must leave; only Jan must leave
'I'm not saying that everybody must leave; only John must leave'

Third, PPIs can scope under negation if a proper intervener scopes between the PPI and its anti-licenser. In a way, we already saw that this is the case for those PPIs that appear under the surface scope of negation (since EXH then acts as an intervener), but more examples of intervention effects can be attested. Example (30) can be true in a situation where when serious family matters are discussed Jan and Piet always leave the room, but the rest doesn't, a reading that requires the universal quantifier to scope below negation.

- (30) Iedereen gaat niet altijd de kamer uit, alleen Jan en Piet
Everybody goes neg always the room out, only Jan and Piet
'Everybody doesn't always leave the room, only Jan and Piet do'

Finally, Szabolcsi (2004), following Baker (1970), shows that PPIs can be rescued again under two anti-licensors (with the highest one being a non-anti-additive anti-licenser). Again, this is the case for Dutch *iedereen*. Take (31). The most salient reading of this sentence is the one where the speaker is surprised that some people left (i.e. that not everybody stayed). Again this reading is only possible if *iedereen* is allowed to reconstruct under negation.

- (31) Het verbaast me dat iedereen niet blijft
It surprises me that everybody neg stays
'It surprises ma that everybody doesn't stay'

So, Dutch *iedereen*, when preceding negation, exhibits all the diagnostics of PPI-hood, thus allowing us to safely conclude that it is indeed a PPI in the classical sense. The fact that it may appear under the scope of surface negation when negation precedes it, simply follows because this PPI introduces an exhaustifier, which in turn may act as an intervener between the PPI and its anti-licenser.

5 Conclusions

To conclude, universal quantifier PPIs do exist, both in the domain of quantifiers over individuals and in the domain of quantifiers over possible worlds, as is predicted by the Kadmon&Landman-Krifka-Chierchia approach to NPI-hood. However, since the exhaustifier that is induced by these PPIs can act as an intervener between the PPI and its anti-licenser, universal quantifier PPIs often appear in disguise. Their PPI-like behaviour only becomes visible once they morpho-syntactically precede their anti-licenser.

References

- Abels, K. & L. Marti. (2010). ‘A unified approach to split scope.’ In *Natural Language Semantics* 18: 435-170.
- Baker, C. (1970). ‘Double negatives.’ In *Linguistic Inquiry* 1: 169-186.
- Chierchia, G. (2006). ‘Broaden your views. Implicatures of domain widening and the “Logicality” of language.’ *Linguistic Inquiry* 37: 535-590.
- Chierchia, G. (2013). *Logic in Grammar: polarity, Free Choice, and Intervention*. Oxford: OUP.
- Homer, V. (t.a.). ‘Neg-raising and Positive Polarity: The View from Modals.’ In *Semantics & Pragmatics*.
- Iatridou, S. & H. Zeijlstra (2010). ‘On the scopal interaction of negation and deontic modals.’ In: M. Aloni, H. Bastiaanse, T. De Jager & K. Schulz (eds). *Logic, language and Meaning*. Berlin: Springer.
- Iatridou, S. & H. Zeijlstra (2013). ‘Negation, polarity and deontic modals.’ *Linguistic Inquiry* 4: 529-568.
- Israel, M. (1996). ‘Polarity Sensitivity as Lexical Semantics.’ In *Linguistics & Philosophy* 19: 619-666.
- Kadmon, N. & F. Landman (1993). ‘Any.’ In *Linguistics and Philosophy* 16: 353-422.
- Krifka, M. (1995). ‘The semantics and pragmatics of polarity items in assertion.’ *Linguistic Analysis* 15: 209-257.
- Postal, P. (2000). ‘An introduction to the grammar of *squat*.’ Ms. NYU.
- Szabolcsi, A. (2004). ‘Positive polarity-Negative polarity.’ In *Natural Language and Linguistic Theory* 22: 409-452.
- Zeijlstra, H. (2004). *Sentential negation and Negative Concord*. PhD Dissertation, University of Amsterdam. Utrecht: LOT Publications.
- Zeijlstra, H. (2012). ,There is only one way to Agree.’ *The Linguistic Review* 29: 491-539.