

Composition and projection in speech and gesture

by

Maria Esipova

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Linguistics

New York University

May, 2019

Lucas Champollion

Acknowledgements

I've read so many dissertation acknowledgements over the years, and now that I'm about to write mine, I'm overcome with dread. How do I do this? How do I convert into words all the warmth, respect, admiration, and gratitude I am feeling towards all the people who supported me throughout this perilous journey? I wish I could gesture my way through this section, but oh well, here goes.

I owe this dissertation and my coming into being as a linguist and an academic to the members of my committee, so I start with them.

Thank you to Lucas Champollion, the chair of this dissertation and my first qualifying paper, my advisor, my mentor, my co-teacher, and my co-author. Thank you, Lucas, for your consistently thoughtful and detailed advice and for giving me numerous opportunities to hone my research and teaching skills. Thank you for introducing me to the marvelous world of event semantics and for making sure I always dot my *i*'s (but don't dot my *l*'s). Thank you also for always listening to me and respecting my perspective—this means so much to me. And thank you for joining us for 'The Room' movie night (and all the spoon throwing), Shakespeare in the Park, the roof post-defense party, and many other events. Thank you, Lucas!

Thank you to Ailís Cournane for diving headfirst, no questions asked, into the deep of gesture research as the chair of my second qualifying paper and then a committee member of my dissertation. Thank you for your own gestures, for raising big picture questions, for 'Language & Mind', and for dinosaurs! Thank you, Ailís!

Thank you to Kate Davidson for welcoming me at Harvard and making me feel like what I do matters. Kate, your masterful ability to balance lambdas with *p*-values, as well as your vast knowledge about words, signs, and gestures have always amazed and inspired me. Thank you also for being so supportive during my dissertation and job market woes. Thank you, Kate!

Thank you to Stephanie Harves for... oh gosh, where do I start? Thank you for being the voice of syntactic reason on this committee and for being a mentor and a friend to me throughout these five years. Thank you for guiding me through my first experience as a TA—one can't wish for a better guide than you! Thank you for making sure I had a place to go to on my first (and second!) New Year's Eve in New York, for my first Brighton Beach escapade, for trips to NetCost Market, for soccer viewing parties, and for all the food and wine. And above all, thank you for always-always-always being there for me when I would find myself at a low point. Shortly after I joined NYU, I heard one of the students say, 'Every department needs a Stephanie Harves'. Well, NYU Linguistics is sure as hell lucky to have one. Спасибо, Стефани!

Thank you to Philippe Schlenker. Philippe, thank you for making me what I am on so many levels, from the summer of 2010, when your NYI lectures on sign language semantics made me fall in love with both sign language and semantics, to now in 2019, when I define myself not only through what I took from you, but also through how I am different from you. This whole dissertation grew from me being inspired by you, but also from me disagreeing with you. Your sharp mind and your passion for pushing the boundaries have always galvanized me, and I thank you for that. I also have very fond memories of you as a person, from 'Produkty' on Fontanka to 'Löwenbräu' in Berlin, and I thank you for them, too. Спасибо, Филипп!

Parts of this dissertation and the work leading to it have been presented at various venues, whose audiences I thank profusely: 'LFRG' at MIT, 'Meaning and Modality Lab' at Harvard, 'Week of Signs and Gestures' at the University of Stuttgart, Sinn und Bedeutung 22, LSA 92, 'Linguistic Explorations Beyond Language' at CreteLing 2, 'Linguistic Investigations Beyond Language' at ZAS Berlin, 'Super Linguistics Research Group' at the University of Oslo, and SALT 29. Special thanks to 'Semantics Group' and the 'Mean & W(h)ine' reading group at NYU.

I also thank everyone who has ever provided their judgements for any of the English, Russian, French, or Spanish examples in this dissertation. Special thanks to our amazing Khoekhoe consultant Gerdut Hevita.

Thank you to Anna Szabolcsi, whose keen awareness of interface interactions and commit-

ment to always keeping in mind the cross-linguistic picture helped me define myself as a linguist. Anna, thank you for always being ready to discuss how Russian and Hungarian are basically the same language and for steering me towards the right research topics. Thank you for once saying, ‘Math is easy, semantics is hard.’, because it is true. Thank you also for providing tips (as in two pages of tips jotted down by Julia) on what to see in New York during my first winter break (the Tenement Museum is still my favourite NYC museum!) and for all the pies. Спасибо, Анна!

Thank you to Maria Gouskova, who has been a role model for me over these past five years in so many respects. Maria Vyacheslavovna, thank you for Phonology I, for your handouts on how not to suck as an academic, for all the fun with Russian diminutive suffixes, for MorphBeer knitting, for driving us to and from NELS 49 in one piece, for FASLs 24 and 29 (my first and last years!), for bringing Wednesday to the department, and for the weightlifting belt. And thank you for quite literally coming to my defense and casually saving me by miraculously producing a VGA-to-HDMI connector when I needed it the most. Спасибо, Маша!

Thank you to all the other faculty at NYU: Mark Baltin, Chris Barker (for being on my first qualifying paper committee and for teaching me that it’s the essential things—such as coffee—that have to be of the highest quality), Renée Blake, Sam Bowman, Dustin Chacón (for letting me cat-sit Ernie), Chris Collins, Lisa Davidson, Gillian Gallagher (for being a great DGS and for bringing Carrot to ‘Snack & Share’), Greg Guy, Richie Kayne (for all the SILENCE and for once saying ‘Sometimes one needs to say things that are true... and tautologies do the job.’), Hadas Kotek (for raising serious issues about questions and questions about serious issues), Laurel MacKenzie, Alec Marantz, Liina Pylkkänen, John Singler (for running the Department’s 2018 World Cup pool, of which I am so proud to be the winner, and for giving me one of your retirement ties), Scott Seyfarth, Juliet Stanton, Gary Thoms (for bringing Molly to the department and for sending job-related info my way).

Thank you to the many non-NYU faculty, who I have been so lucky to interact with. Thank you to the MIT and Harvard faculty who welcomed me during my semester in Cambridge: Isabelle Charnavel, Gennaro Chierchia, Kai von Fintel, Edward Flemming, Irene Heim, Sabine Iatridou,

David Pesetsky, and Roger Schwarzschild. Thank you to Jason Bishop (for the seminar on international phonology), Barbara Partee (for introducing me to many puzzles in semantics back in Moscow and for letting me stay with you, Volodja, Corey, my namesake Mashka, and Tyushka during my first NELS), Pritty Patel-Grosz and Patrick Grosz for the warm welcome in Oslo and for your dedication to super-linguistics. Huge thanks to John Bailyn for letting me discover what I want to be at NYI (and thanks to all the NYI faculty!), for hosting me during my NYU Open House visit, and for saving me when I got stranded in Port Jefferson at the latest FASL.

Thank you to our department administrators Hannah Katz and Teresa Leung for making my life at NYU so much easier. Thank you to our security guards Mike and Eddie for making my comings and goings to and from the department so much brighter.

Thank you to my #613 officemates. Thank you to Adina Williams for showing me first-hand how one survives their fifth year, for oysters, for movie nights and scarab poop balls, for ‘Semantics Group’, for conference trips, and for the picnic blanket. And thank you so much, Adinochka, for hosting the post-defense party for me on your awesome roof! Thank you to Paloma Jeretić for the little support group that we created for ourselves, for your trilingual judgements, for ‘Semantics Group’ (too!), and for all the fun from St. Petersburg to France. Thank you to Alicia Parrish for the camping trip, for all the coffee, and for letting me feed your cats. Thank you to Sarah Phillips for your no-bullshit take on everything, for all the pics, and for the NWAV whiskey.

Thank you to my cohort-mates: WooJin Chung (for the snow lion in our first year and for bringing Miu and Jiu to the department), Maria Kouneli (for all the parties and for all our conversations during our fifth year), and Natalie Povilonis de Vilchez.

Thank you to (some of) my older peers: Dan Duncan (for GoT viewing parties and for always killing it at the music trivia round), Itamar Kastner (for all your puns), Sonia Kasyanenko (for all the late night conversations), Dani Szeredi (for all the Happy Hours, Georgian food, and for always making me feel like I’m not the only jaded Eastern European in the room), Dunja Veselinović (for the BHFF festival and for impossible burgers), and Vera Zu (for soup dumplings).

Thank you to the current 4th years: Haoze Li, Yining Nie (for co-teaching with me and

for not letting me kill myself while benching), Mary Robinson (for making me acquire the *wh-all* construction in English), Ildikó Emese Szabó (for all the hugs), Sheng-Fu Wang (for making us all enjoy life). Thank you to the current 3rd years: Hagen Blix, Maddie Gilbert, Maxime Tulling, and Sora Yin. Thank you to the current 2nd years: Omar Agha, Ioana Grosu, Chris Klehm, Naomi Lee, and Alex Warstadt. Thank you to the current 1st years: Anna Alsop (for being such a great gesture model), Kimberley Baxter, Alicia Chatten, Kate Mooney, Jailyn Pena, and Guy Tabachnick. Thank you to our visiting students: Timothée Bernard, Thom van Gessel, Karo Lohiniva, and Kata Wohlmuth.

Thank you to all my non-NYU fellow students. I can't possibly list you all, so I will have to limit myself to a selected few. Thank you, Amir Anvari (for our adventures in Toulouse and in Chartrettes), Tania Bondarenko (for hosting me for SALT 28), Deniz Özyıldız (for our Crete shenanigans; you are my cohort-mate, I don't care what they say), Rob Pasternak (for your expert gestures and for all our conversations, from Dumskaya to Checkpoint Charlie), Mitya Privoznov (for chechil on a night boat trip), Jon Rawski (for your not-so-expert, but very fun gestures). Thank you, Nina Feygl-Tzipora Semushina Leyn. Feygl, Neenious, Professor—видите, Вас так много, а меня так мало—I've known you under different names, but you have always remained my dearest friend. We started it together, back in the hot summer of 2010, when I dragged you to NYI with me for the first time. Remember how we were sitting on the embankment on Vasilyevsky Island, looking over the Gulf of Finland (back when there was an embankment and you could still see the Gulf of Finland from there), and how I told you that I finally knew what I wanted to do with my life, but that it was impossible? Well, here we are now, nine years later, and it wasn't all that impossible, after all.

Thank you to my dexterous DnD alter ego Héloïse the Ranger and her fearless companion Lambda the Wombat. Thank you to my fellow adventurers, who all have been mentioned by now, for fighting by my side, starting fires, hiding behind a table, smashing rocks, hanging on the wall, beating the drum, knocking on doors, committing accidental genocide, and doing other fun things. Thank you to Andy Huemmer, for guiding our adventures.

Thank you to Quentin Tarantino and J.K. Rowling, for providing most of the names used in my examples here and elsewhere.

Thank you and sorry to everyone who I forgot to mention or had to omit due to limitations of space.

Finally, thank you to мама and кошка. This dissertation is dedicated to you.

Abstract

Some types of semantic content have a potential or a requirement to project, i.e., to get interpreted outside the semantic scope of various operators such as negation, conditional and question operators, modals, etc., despite appearing to be in their syntactic scope. But what is the exact classification of projective content, and how do we know which type of content projects in which way? In this dissertation I argue for a composition-driven approach to projection of pieces of content that have their own nodes in the morphosyntax, and I argue that this approach applies to both spoken and gestural content. Under this approach, how a given piece of content projects is uniformly determined by how it composes in the syntax/semantics, as opposed to a variety of alternative analyses for different types of content. I show that this approach is efficient in explaining the projection behavior of a range of content types, including spoken adjectives and appositives, hand gestures, facial expressions, *phi*-features on pronouns, and certain iconic properties of directional hand gestures. This dissertation, thus, has two main implications. First, it offers a general economical and explanatory approach to investigating projection of compositionally integrated content. Second, it establishes the need for treating content-bearing gestures as bona fide linguistic objects at all levels of representation.

Note on the versions of this dissertation

This version of my dissertation is different from the one published on ProQuest. While the two versions don't contradict each other in any major way, this is the more complete (and more typo-free) version intended for distribution.

Contents

Acknowledgements	ii
Abstract	viii
Note on the versions of this dissertation	ix
List of Figures	xiv
List of Tables	xv
List of Appendices	xvi
1 Introduction	1
1.1 Contextualizing this dissertation	1
1.2 Posing the question	2
1.3 Sketching the answer	6
1.4 Outline of the dissertation	9
2 Modifiers and supplements	13
2.1 Modifiers	14
2.1.1 Composition of modifiers	14
2.1.2 Projection of modifiers	20
2.1.2.1 Restricting vs. non-restricting modifiers	20
2.1.2.2 Non-restricting modifier inferences as semantic presuppositions (Leffel 2014)	24

2.1.2.3	Proposal: non-restricting modifier inferences as pragmatic pre-suppositions	26
2.1.2.4	Non-restricting modifier inferences and cosuppositions	39
2.1.2.5	Preferably non-restricting modifiers	46
2.2	Supplements	48
2.2.1	Composition of supplements	48
2.2.2	Projection of supplements	52
2.3	Summary of the chapter	54
3	Adjectives, appositives, and gestures: experiment	56
3.1	Background and goals of the study	57
3.2	Methods	61
3.3	Results	65
3.4	Discussion	67
3.4.1	Desiderata for a theory of co-speech gestures	67
3.4.2	Projecting non-restricting interpretations and <i>Minimize Restrictors!</i>	68
3.4.3	Note on “high” acceptability of unacceptable interpretations	69
4	Gestures	71
4.1	Supplemental analysis of co-speech gestures	72
4.2	Cosuppositional analysis of co-speech gestures	73
4.2.1	The gist of the cosuppositional analysis	73
4.2.2	NP-level gestures under the cosuppositional analysis	78
4.2.3	DP-level gestures under the cosuppositional analysis	80
4.3	Proposal: gestures can be modifiers or supplements	86
4.3.1	The principle of no gesture-specific composition	86
4.3.2	Integrated and non-integrated gestures	87
4.3.3	Deriving the data at hand	92

4.3.4	Gestures and linearization	95
4.4	Beyond hand gestures: facial expressions	102
4.5	The role of convention	112
4.6	Concluding remarks	117
5	<i>Phi</i>-features on pronouns	119
5.1	Projection of <i>phi</i> -features	120
5.2	Lexical presupposition analyses of <i>phi</i> -features	121
5.3	Proposal: <i>phi</i> -features as obligatorily non-restricting modifiers	125
5.4	Gender, local contexts, and indexicals	128
5.5	<i>Phi</i> -features and <i>Minimize Restrictors!</i>	140
5.6	Height specifications on gestures	141
6	Modifiers and supplements in the verbal domain	145
6.1	Spoken modifiers and supplements in the verbal domain	145
6.2	Gestural modifiers and supplements in the verbal domain	148
6.2.1	Projection of co-verbal gestures	148
6.2.2	Appositive-like gestures	151
7	Conclusion	157
7.1	Summary of the dissertation	157
7.1.1	Summarizing the main points of the dissertation	157
7.1.2	Summarizing the contributions of the dissertation	159
7.2	Directions for future research	164
7.2.1	Metalinguistic focus	164
7.2.2	Projective content under ellipsis and in attitude reports	166
7.2.3	Composition and projection of tense and aspect	171
7.2.4	Speech–gesture integration as code-switching/code-blending	173
7.2.5	Linearization constraints in multi-modal composition	176

Appendices **177**

Bibliography **194**

List of Figures

3.1 Experiment: layout of a typical trial.	61
3.2 Experiment: bar chart showing % acceptability of different interpretations for each content type.	66

List of Tables

2.1	Modifiers vs. supplements: a preliminary summary.	55
3.1	Design of the experiment.	62
3.2	Experiment: statistics for acceptability of different interpretations for each content type.	67
5.1	Results of a mini-poll on local context sensitivity of gender on pronouns.	133
7.1	Modifiers vs. supplements: a final summary.	158

List of Appendices

Appendix A Notational conventions	177
Appendix B Glossary of terms	179
Appendix C Instructions for Experiment in Chapter 3	183
Appendix D Items for Experiment in Chapter 3	187

Introduction

1.1 Contextualizing this dissertation

As semanticists, we are interested in how meaning works in natural language. However, when we have an utterance, different parts of its meaning will behave differently with respect to a range of dimensions, such as ability to address questions under discussion, ability to be targeted by direct responses of various kinds, ability to be ignored under ellipsis, and projection from under semantic operators. We want to explain these different patterns of behavior for different types of content and the connections between them in a principled way, based on independent semantic, syntactic, pragmatic, and prosodic factors.

Furthermore, when we have an utterance in spoken language, it contains various components other than a string of spoken words that contribute to its meaning, such as hand gestures, facial expressions, iconic voice modulations, etc. We want to know how exactly this content contributes to the meaning of utterances. Much of the work in formal semantics and pragmatics of gesture (and some other secondary modality content) has focused on how gestures interact with semantic operators such as negation, conditional and question operators, modals, quantifiers, etc. (Ebert & Ebert 2014; Ebert 2017; Schlenker 2018a,b; Tieu et al. 2017, 2018; Esipova 2019a). The main question all this literature asks is if, when, how, and why the content contributed by gestures projects from under various operators. Much of this literature, however, has been relying on pre-theoretical categories, for example, juxtaposing “co-speech” and “post-speech” gestures

(i.e., gestures co-occurring with and following the spoken expressions they associate with, respectively), when approaching the projection problem for gestures. In my recent work (Esipova 2017, 2018a,b,c, 2019b,c) I have been arguing that in order to answer the question about gesture projection, we need to systematically ask theoretically meaningful questions about how gestures integrate into spoken utterances at various levels (syntax, compositional semantics, linearization, prosody, articulation, discourse) and how this integration is constrained at the interfaces between these levels. To address these questions in a rigorous way, we need to approach gestures as bona fide linguistic objects that abide by linguistic rules at all levels of representation. The goal of an interested researcher is thus to determine what these rules are, when they are the same as the rules for spoken expressions, and when they are different.

These two lines of research—an interface-oriented approach to varied behavior of different types of content and a linguistic approach to multi-modal integration of content at various levels of representation—come together in this dissertation, which investigates projection of different types of content, both spoken and gestural, as determined by how they compose in the syntax/semantics.

1.2 Posing the question

Some types of semantic content have a potential or a requirement to PROJECT from under semantic operators (such as negation, question or conditional operators, modals, etc.), i.e., to get interpreted outside their semantic scope despite appearing to be in their syntactic scope. In (1.1) I illustrate the projective potential of five types of NON-SUBLEXICAL CONTENT, by which I mean content that has its own node in the morphosyntax and, thus, has to integrate compositionally in the semantics—unlike the arguably sublexical projective content, such as presuppositions of words like *stop*, *know*, or *regret*. The five types of content in (1.1) are adnominal adjectives, adnominal appositives, CO-NOMINAL GESTURES (i.e., content-bearing gestures co-occurring and associating with spoken nominal expressions), *phi*-features on pronouns, and height specifications on CO-VERBAL GESTURES (i.e., content-bearing gestures co-occurring and associating with spoken verbal expres-

sions). In (1.1), all these five types of content contribute inferences that end up projecting from under *if*.¹

- (1.1) a. *Context: We are going on a group tour and want to rent a van. The speaker just learned that Stephanie might bring along her only dog.*²

If Stephanie's bringing...

- (i) her large dog ADJECTIVE
- (ii) her dog, {a large animal, who is large} APPOSITIVE
- (iii)  her dog_{LARGE} CO-NOMINAL GESTURE

..., we should get a bigger van.

→ Stephanie's dog is large.

- b. If Skyler_i brings her_i dog, I'll give you \$10. PHI-FEATURE

→ Skyler is female.

- c. *Context: Zoe is a stuntwoman. The crew just filmed a scene in which she was fighting an extra, while the director of the movie Uma was away. Uma originally wanted Zoe to stab the extra in that scene, but she just learned that Zoe might have punched the extra in the face instead. Uma says:*



If Zoë punched the extra_{PUNCH-HIGH}, we'll have to reshoot the scene.

HEIGHT SPECIFICATION

→ The extra that Zoë was fighting is taller than Zoë.

If you ask a random semanticist how many types of projecting inferences are illustrated in (1.1), they will probably say between three and five, depending on who you ask.

The least controversial cases would probably be (1.1a-ii) and (1.1b). For (1.1a-ii), most se-

¹See Appendix A for the notational conventions used throughout this dissertation.

²The van-renting scenario for the paradigm in (1.1a), which I have been using extensively in my work, was proposed to me by Caroline Heycock (p.c.) as a means to justify the use of the size adjuncts.

manticists would tell you that appositives (i.e., nominal appositives and appositive relative clauses) have to project because they belong in a special class of expressions, sometimes called SUPPLEMENTS, which have a special projection mechanism associated with them. What exactly this mechanism is and what other types of content make use of it will depend on the specific analysis one assumes, of which there are plenty (e.g., Potts 2005; AnderBois et al. 2013; Koev 2013; Schlenker 2013). For (1.1b), you are most likely to get the answer that the gender inference is a lexical presupposition triggered by the feminine feature on the pronoun *her* (e.g., Cooper 1983; Heim & Kratzer 1998; Elbourne 2005; Sudo 2012).

The three remaining cases will be less clear-cut. When faced with (1.1a-i), one might think that the inference contributed by the adjective *large* projects as part of the existence inference of the definite description *her large dog* (combined with the extralinguistic context that assures said dog's uniqueness). The existence inference might indeed have an effect in (1.1a-i), but, as we will see later, inferences contributed by modifiers that don't restrict the expressions they modify (i.e., truth-conditionally vacuous modifiers), as is the case for *large* in (1.1a-i), (i) project stronger than existence inferences of definite descriptions, and (ii) arise outside of definite descriptions, in particular, in quantifier DPs (as was noticed in Leffel 2014). In other words, inferences of non-restricting modifiers constitute their own projecting inference type. At this point, if you accept the stories for (1.1a-ii) and (1.1b) above, we already have three types of projecting inferences, for spoken content alone.

Enter gestures. Ebert & Ebert (2014), in the first attempt to explain the projection behavior of CO-SPEECH GESTURES, i.e., gestures co-occurring and associating with spoken expressions, like the one in (1.1a-iii), argue that such gestures are akin to appositives and thus give rise to the same type of projecting inferences as the latter. If you are happy to adopt this view, the projecting inference type counter remains at three. Schlenker (2018a) argues against this approach and proposes instead that co-speech gestures trigger presuppositions of a special kind, which he calls COSUPPOSITIONS. Schlenker's cosuppositions are presuppositions in that they are inferences that have to be entailed by the local context of the triggering expression, but they have to be very dif-

ferent from, say, the purported presuppositions of *phi*-features on pronouns. First, cosuppositions have a different form, namely, they are assertion-dependent, conditional presuppositions. Second, for empirical reasons, Schlenker has to maintain that gestural cosuppositions are WEAK in that they don't always have to project from under semantic operators; presuppositions of *phi*-features, however, are typically taken to be STRONG, i.e., they always have to project. Third, the two types of presuppositions are triggered differently. In the case of *phi*-features, the presupposition is usually taken to be a lexical property thereof; in the case of co-speech gestures, it has to be something about the configuration in question (whereby a gesture co-occurs and associates with a spoken expression) that triggers the cosupposition. So, if you adopt Schlenker's analysis of co-speech gestures, the projecting inference type counter goes up to four.

Finally, very few people will have a well-formed opinion about the inferences contributed by height specifications on co-verbal gestures as in (1.1c). The null hypothesis would be that they constitute their own projecting inference type, which would bring our counter up to four or five, depending on what analysis we adopt for co-speech gestures. Schlenker & Chemla (2018), however, assimilate height specifications on verbal gestures with their own time slot to *phi*-features on pronouns, a view that extends straight-forwardly to height specifications on co-verbal gestures. If you are ready to accept this view, the projecting inference type counter will remain at three or four.

Basically, the picture we are getting at this point for the projecting inference types represented in (1.1) looks like this:

- (1.1)' a. *Context: We are going on a group tour and want to rent a van. The speaker just learned that Stephanie might bring along her only dog.*

If Stephanie's bringing...

- | | |
|---|----------------------------|
| (i) her large dog | NON-RESTRICTING MODIFIER |
| (ii) her dog, {a large animal, who is large} | SUPPLEMENT |
| (iii) <u>her dog</u> _{LARGE} | SUPPLEMENT? COSUPPOSITION? |

..., we should get a bigger van.

- Stephanie's dog is large.
- b. If Skyler_i brings her_i dog, I'll give you \$10. LEXICAL PRESUPPOSITION
 → Skyler is female.
- c. *Context: Zoe is a stuntwoman. The crew just filmed a scene in which she was fighting an extra, while the director of the movie Uma was away. Uma originally wanted Zoe to stab the extra in that scene, but she just learned that Zoe might have punched the extra in the face instead. Uma says:*
 If Zoe punched the extra_{PUNCH-HIGH}, we'll have to reshoot the scene.
 SOMETHING NEW? LEXICAL PRESUPPOSITION?
 → The extra that Zoe was fighting is taller than Zoe.

Can we clean up this picture? How many projecting inference types are in fact represented in (1.1)? And, more importantly, what determines which type of content gives rise to which type of inference?

1.3 Sketching the answer

In this dissertation I show that we can keep the projecting inference type counter at two, not only for the five types of content in (1.1), but also as we start adding other types of non-sublexical projective content to the picture. I also show that we can have a principled way of determining which inference type arises when that works uniformly for spoken and gestural content. I propose that we can accomplish these two things by adopting a composition-driven approach to projection of non-sublexical content whereby how a given piece of content projects is determined by how it composes in the syntax/semantics.

In this dissertation I focus on two types of compositionally integrated content: (i) (SUBSective) MODIFIERS, as exemplified by (subjective) adjectives, which combine with an expression β yielding an expression that entails β , and (ii) SUPPLEMENTS, as exemplified by appositives, which combine with an anchor yielding a proposition of a special kind about that anchor. The two

composition strategies are associated with two distinct patterns of projection behavior.

Specific instances of subsective modifiers can be NON-RESTRICTING, i.e., intended by the speaker as truth-conditionally vacuous. If the addressee thinks that the speaker intends a certain instance of a modifier as non-restricting, they will pragmatically generate an inference that the expression being modified entails the result of modification. This inference will project as a presupposition, i.e., relative to a local context. I furthermore propose to equate Schlenker's cosuppositions with non-restricting modifier inferences (both for spoken and gestural content), which simultaneously constrains cosuppositions in a way that prevents overgeneration of unattested interpretations and allows them to apply in cases in which Schlenker's original formulation can't (namely, cases of non-intersective modification).

Supplements project, and very strongly so, because they are supplements; any of the analyses cited in the previous section will do. I will not propose any new analysis of supplements, but I will provide arguments that it is not possible to reduce their projection to that of non-restricting modifiers.

Adjectives uncontroversially exemplify the modifier strategy, and appositives uncontroversially exemplify the supplement strategy. In this dissertation I will show how we can reduce other types of content, spoken and gestural, to one or the other.

In particular, I will argue that both the modifier and the supplement strategies are available to co-speech gestures, but there is a pragmatic pressure for them to be truth-conditionally vacuous. Supplemental co-speech gestures satisfy this requirement because they project as supplements; like most other supplements, they project very strongly and can't be interpreted locally. Modifier co-speech gestures, in turn, prefer to be non-restricting (and will thus give rise to non-restricting modifier inferences), although this preference can be overridden under pressure. I, thus, unify the two analyses of co-speech gestures in Ebert & Ebert 2014 and Schlenker 2018a, but I also go beyond them by linking how gestures project to how they integrate into utterances compositionally instead of relying on their linearization. I also adduce novel experimental data showing that neither of these two analyses can account for available and unavailable interpretations of co-speech

gestures on its own, but the composition-driven analysis proposed here can.

As for *phi*-features on pronouns, I argue that they are obligatorily non-restricting modifiers, which explains why they always project. This proposal has an edge over the standard lexical presupposition analysis, since it doesn't need to stipulate that *phi*-features on pronouns lexically trigger presuppositions that are, furthermore, strong, nor does it need to assume any new composition strategies. Treating *phi*-features on pronouns as modifiers is also more plausible from the morphosyntactic point of view.

Finally, height specifications on gestures can be analyzed as modifiers as well. In the case of verbal gestures like in (1.1c), height specifications modify incorporated nominal arguments. Depending on the nature of this nominal argument, height specifications can be further assimilated to *phi*-features on pronouns or to ordinary modifiers (such as adjectives).

In other words, I propose that when it comes to projecting inference types represented, (1.1) looks like this:

- (1.1)" a. *Context: We are going on a group tour and want to rent a van. The speaker just learned that Stephanie might bring along her only dog.*

If Stephanie's bringing...

- | | |
|--|--|
| (i) her large dog | NON-RESTRICTING MODIFIER |
| (ii) her dog, {a large animal, who is large} | SUPPLEMENT |
| (iii) her dog _{LARGE} | NON-RESTRICTING MODIFIER or SUPPLEMENT |

...., we should get a bigger van.

→ Stephanie's dog is large.

- b. If Skyler_i brings her_i dog, I'll give you \$10. NON-RESTRICTING MODIFIER
→ Skyler is female.

c. *Context: Zoe is a stuntwoman. The crew just filmed a scene in which she was fighting an extra, while the director of the movie Uma was away. Uma originally wanted Zoe to stab the extra in that scene, but she just learned that Zoe might have punched the extra instead.*

extra in the face instead. Uma says:

If Zoe punched the extra_{PUNCH-HIGH}, we'll have to reshoot the scene.

NON-RESTRICTING MODIFIER

→ The extra that Zoe was fighting is taller than Zoe.

Much of this dissertation will focus on adnominal content such as in (1.1), but I also extend the composition-driven approach to projection of degree modifiers, facial expressions, and modifiers and supplements in the verbal domain.

The composition-driven approach to projection should be understood as first and foremost a heuristic method; I take motivating and developing this method to be one of the two main goals of this dissertation. When faced with a given piece of non-sublexical content that has a potential or a requirement to project, before asking the question of when and how it projects, one should first ask the question of how it composes. The second main goal of this dissertation is to apply this approach to spoken and gestural content equally, furthering the idea that gestures can and should be analyzed as linguistic objects, and many of the same linguistic rules apply to spoken and gestural content alike at various levels of representation. As I hope to show in this dissertation, adopting both ideas leads to significant conceptual and empirical gains.

1.4 Outline of the dissertation

This section is meant to help the reader navigate through this dissertation.

Chapter 2 lays out the two composition strategies this dissertation focuses on, which are modification and supplementation. It also talks about the projection behavior of modifiers and supplements. In particular, I argue that apart from a mechanism assuring projection of supplements, the need for which has been recognized and fulfilled in the literature, we also need an independent mechanism assuring projection of inferences contributed by non-restricting, i.e., truth-conditionally vacuous, modifiers. I build on the definition of non-restricting modifiers in Leffel 2014 to propose that non-restricting modifier inferences are triggered and project as pragmatic

presuppositions. This view of non-restricting modifier inferences is more conceptually appealing and more general than the one proposed by Leffel himself. I furthermore propose to equate non-restricting modifier inferences as defined by me and cosuppositions, i.e., assertion-dependent presuppositions proposed in Schlenker 2018a,b,c. I show, in particular, that this move allows cosuppositions to apply both to intersective and non-intersective modifiers, which is not the case in Schlenker's formulation. While doing so, I also make a slight, independently motivated adjustment to the classical treatment of degree modification in Kennedy & McNally 2005, which makes degree modifiers actual modifiers and renders the internal composition of modifier phrases more similar to that of nominals and clauses.

Chapter 3 reports experimental results on how co-nominal gestures can and cannot be interpreted, as compared to spoken adnominal adjectives and appositives. As expected, adjectives can easily be restricting or non-restricting, and when they are non-restricting, they have to project. Appositives cannot be restricting, and they always have to project. Co-nominal gestures can be non-restricting, and when they are, they have to project; restricting interpretations are also available to them, but they are degraded as compared to non-restricting ones, which sets them apart from both adjectives and appositives.

Chapter 4 focuses on explaining the projection behavior of gestures. I start by showing why neither of the two existing analyses of projection of co-speech gestures, by Ebert & Ebert (2014) and Schlenker (2018a), can account for the gestural data from Chapter 3 on its own. I then proceed to applying the composition-driven approach to projection of gestures. First, I discuss some general principles of how gestures integrate into utterances compositionally; in particular, I posit the ‘no gesture-specific compositionality’ principle. I then argue that both the modifier and the supplement composition strategies are in principle available to co-nominal gestures, but co-speech gestures prefer to be truth conditionally vacuous, so when they are modifiers, they are preferably interpreted as non-restricting. Next, I discuss how language-specific linearization and prosodic grouping rules may constrain the use of gestures with their own time slot as modifiers in otherwise spoken utterances. I then adduce some preliminary data on co-speech and co-gesture facial

expressions encoding surprisal. Those data show that, much like spoken adverbs with similar semantics, such as *surprisingly* or *impressively*, such facial expressions can compose either as degree modifiers, in which case they don't seem to project by default, or as sentence-level supplements, in which case they project as such. I thus strengthen the main point of the chapter that how secondary modality content projects is determined by how it composes, not by how it is linearized or by whether it is syntactically optional, contra Schlenker 2018a,b. I also discuss the role of convention in how a given piece of gestural content projects and conclude that level of conventionalization doesn't seem to be a directly relevant factor in how a given piece of content projects, nor is treating it as such a theoretically appealing choice. I conclude by adding some methodological remarks on how to investigate projection of secondary modality content.

Chapter 5 applies the composition-driven approach to projection of inferences contributed by *phi*-features on pronouns. In particular, I argue that *phi*-features on pronouns are modifiers, which are obligatorily non-restricting, because they always modify a property whose extension is a singleton set. That explains why contributions of *phi*-features on pronouns project similarly to other non-restricting modifier inferences. I compare this approach to two versions of the lexical presupposition analysis of *phi*-features on pronouns, in Heim & Kratzer 1998 and Elbourne 2005, and argue that the modifier approach is superior conceptually and more plausible morphosyntactically. The argument about morphosyntactic plausibility is supported by novel data on morphosyntactic composition of pronouns in Khoekhoe coming from Lee (2019) and our joint fieldwork. I also add a discussion on why the inferences contributed by gender features on pronouns are less sensitive to local contexts than other non-restricting modifier inferences. I adduce novel data on inter- and intra-speaker variation for this phenomenon, likening gender on pronouns to T-V features on pronouns in Russian, and outline the direction for analyzing both as FORM INDEXICALS, which are interpreted relative to indices within the context parameter that keep track of appropriate forms of expressions (unifying and enriching some of the previous suggestions in the literature, such as Yanovich 2010 and Schlenker 2007). This treatment of gender and T-V on pronouns is independent of their non-restricting modifier status. The final section of this chapter applies the

composition-driven approach to height specifications on verbal gestures. I argue that a natural construal of such configurations involves a nominal argument incorporated into the gesture, which the height specification modifies. Depending on the exact nature of this nominal argument, height specifications can then be assimilated to *phi*-features on pronouns or ordinary modifiers.

Chapter 6 shows that the modifier vs. supplement distinction exists for content that combines with verbal expressions as well, with similar consequences for projection. I also discuss a crucial difference between supplements in the nominal and in the verbal domains and how it bears on the status of appositive-like gestures.

Chapter 7 summarizes the main points and contributions of the dissertation and outlines potential directions for future research.

Modifiers and supplements

Abstract

This chapter sets the theoretical stage for the rest of the dissertation by laying out two of the composition strategies used by non-sublexical content, the (SUBSECTIVE) MODIFIER strategy and the SUPPLEMENT strategy, and the associated patterns of projection behavior.

Roughly, subsective modifiers are pieces of content that compose with an expression β yielding an expression α that entails β . Supplements are pieces of content that compose with an expression and return a proposition of a special kind about that expression.

Subsective modifiers are always RESTRICTIVE in that they have the compositional potential to restrict the expression they modify, i.e., to pick out a potentially smaller part of its denotation; however, specific instances of modifiers can be NON-RESTRICTING. Following Leffel 2014, I define non-restricting modifiers as modifiers that are intended by the speaker as truth-conditionally vacuous. However, I refine Leffel's original definition of non-restricting modifiers by divorcing their truth-conditional vacuity, which holds at the level of the entire utterance, from the inference that the expression being modified entails the result of modification, which I call the non-restricting modifier inference. I observe that non-restricting modifier inferences project like presuppositions, i.e., relative to local contexts, and propose that these presuppositions are triggered pragmatically. This view of non-restricting modifier inferences is more conceptually appealing and more general than the one proposed by Leffel himself whereby non-restricting modifier inferences are semantic

presuppositions. I furthermore propose to equate non-restricting modifier inferences as defined by me and cosuppositions, i.e., assertion-dependent presuppositions proposed in Schlenker 2018a,b,c, which, among other things, allows cosuppositions to apply to non-intersective modifiers, which is not the case in Schlenker’s formulation. In the process I also make an independently motivated adjustment to the analysis of degree modifiers in Kennedy & McNally 2005, which makes composition of degree modifiers more similar to that of modifiers in the nominal and verbal domains.

Supplements don’t have the compositional potential to be restricting to begin with, i.e., they are NON-RESTRICTIVE. They also typically have to project, and they project very strongly, which can be and has been captured in a variety of ways. Crucially, projection of supplements is more regular and more conventional than that of non-restricting modifier inferences, so one cannot be reduced to the other.

2.1 Modifiers

2.1.1 Composition of modifiers

From the semantic point of view, the term MODIFIER can be used broadly to describe an expression γ that combines with an expression β of a type τ yielding an expression α of the same type τ . This type-based view of modification is, for example, assumed as a useful approximation in Morzycki 2015 and is different from the use of this term by Potts (2005), who calls any adjuncts, including, for example, appositives, “modifiers”, regardless of their semantics. While I will in general adopt the type-based notion of modification, throughout most of this dissertation I will only be discussing SUBSECTIVE MODIFIERS, so I will often use the term MODIFIER sloppily, as equivalent to SUBSECTIVE MODIFIER.³

What are subsective modifiers? Usually this term is used descriptively to refer to modifiers such that the result of modification entails the expression being modified. For example, *blond* is a

³Note also that under the purely type-based view of modification, negation, for example, is also a modifier, but it’s definitely not the kind of content I am focusing on in this dissertation.

subsective modifier, because, intuitively, *blond NP* always entails *NP*, as illustrated in (2.1):

(2.1) Zoe is a blond stuntwoman.

→ Zoe is a stuntwoman.

Formalizing this intuition, we can say that a subsective modifier is an expression γ that combines with an expression β yielding an expression α such that α entails β . What does it mean for two sub-propositional expressions to entail one another? It means that α entails β via GENERALIZED ENTAILMENT ($\alpha \Rightarrow_{\forall} \beta$), which simply means that α entails β via GENERALIZED MATERIAL IMPLICATION ($\alpha \Rightarrow \beta$) for any values of the arguments α and β take. The full definitions are given below. I will call this entailment relation between the result of subsective modification and the expression being modified the SUBSECTIVE ENTAILMENT.

(2.2) **SUBSECTIVE MODIFIER (descriptive definition, to be revised)**

In a tree α whose daughters are γ and β , γ is a SUBSECTIVE MODIFIER iff $\alpha \Rightarrow_{\forall} \beta$.

(2.3) **GENERALIZED MATERIAL IMPLICATION⁴**

$$[\![\alpha_{(\tau_1 \dots \tau_n, t)} \Rightarrow \beta_{(\tau_1 \dots \tau_n, t)}]\!] = \lambda X_{\tau_1}^1 \dots \lambda X_{\tau_n}^n . [\![\alpha]\!](X_{\tau_1}^1) \dots (X_{\tau_n}^n) \rightarrow [\![\beta]\!](X_{\tau_1}^1) \dots (X_{\tau_n}^n)$$

(2.4) **GENERALIZED ENTAILMENT**

$$[\![\alpha_{(\tau_1 \dots \tau_n, t)} \Rightarrow_{\forall} \beta_{(\tau_1 \dots \tau_n, t)}]\!] = 1 \text{ iff } \forall X_{\tau_1}^1 \dots \forall X_{\tau_n}^n [\![\alpha \Rightarrow \beta]\!](X_{\tau_1}^1) \dots (X_{\tau_n}^n)]$$

For example, the formalization of the intuitive entailment relation from (2.1) is given in (2.5).

(2.5) *blond stuntwoman* \Rightarrow_{\forall} *stuntwoman*

$$\forall x \forall w [\![\text{blond stuntwoman}]\!](x)(w) \rightarrow [\![\text{stuntwoman}]\!](x)(w)]$$

Now, (2.2) aims to define SUBSECTIVE MODIFIER as a natural class of content. However, I said at the outset of this chapter that I would discuss subsective modification as a composition strategy,

⁴Generalized material implication is a lifted version of material implication that can apply to sub-propositional expressions, just like generalized conjunction (Partee & Rooth 1983) is a lifted version of logical conjunction. The definition in (2.3) is adopted from Schlenker 2018a, fn. 25, with minor notational changes.

namely, a composition strategy used to yield the configuration in (2.2) characterized by subsective entailment. (2.6) lays out this composition strategy. Under the composition-based definition, a SUBSECTIVE MODIFIER is an expression γ that combines with an expression β in such a way that the truth conditions in the denotation of the resulting expression α contain a conjunct that applies β to all the arguments α is going to take.

(2.6) **SUBSECTIVE MODIFIER (composition-based definition, adopted here)**

In a tree $\alpha_{\langle \tau_1 \dots \tau_n, t \rangle}$ whose daughters are γ and $\beta_{\langle \tau_1 \dots \tau_n, t \rangle}$, γ is a SUBSECTIVE MODIFIER iff:

$$[\![\alpha]\!] = \lambda X_{\tau_1}^1 \dots \lambda X_{\tau_N}^n. (\dots) \wedge [\![\beta]\!](X_{\tau_1}^1) \dots (X_{\tau_N}^n)$$

It is easy to see that the definition in (2.6) ensures the subsective entailment from (2.2). This is the definition I will use throughout this dissertation.

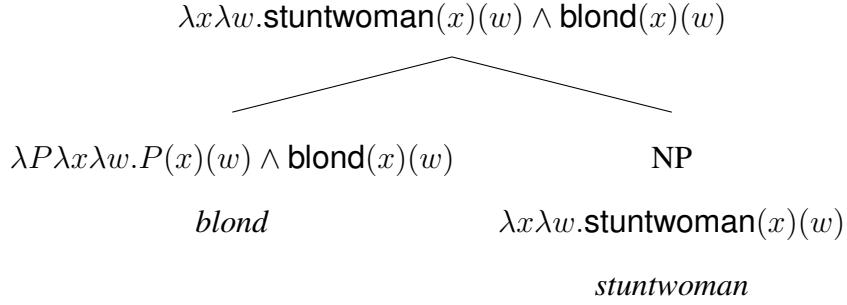
Note, however, that I keep talking about subsective modification as a composition “strategy” rather than a “rule” and have ellipsis in place of γ ’s contribution in (2.6). This vagueness is on purpose. Different modifiers make their contributions in different ways.

For simple cases like *blond stuntwoman*, we can perform subsective modification in two ways. Those are PREDICATE MODIFICATION, when we combine two expressions of the same type by conjoining them, and FUNCTIONAL APPLICATION, when one expression (the modifier) takes the other (the one being modified) as an argument. The two ways are illustrated in (2.7).

(2.7) a. **Composing *blond stuntwoman* via predicate modification**

$$\begin{array}{c}
 \lambda x \lambda w. \text{stuntwoman}(x)(w) \wedge \text{blond}(x)(w) \\
 \diagdown \qquad \diagup \\
 \lambda x \lambda w. \text{blond}(x)(w) \qquad \text{NP} \\
 \text{blond} \qquad \qquad \lambda x \lambda w. \text{stuntwoman}(x)(w) \\
 \qquad \qquad \qquad \text{stuntwoman}
 \end{array}$$

b. Composing *blond stuntwoman* via functional application



In (2.7), the two results are identical, and in both cases the subsective entailment is assured by the presence of the $\text{stuntwoman}(x)(w)$ conjunct in the result of modification. However, predicate modification introduces this conjunct straight-forwardly, via the mechanics of generalized conjunction. Giving the modifier a higher type and having it take the expression being modified as an argument makes the subsective entailment a lexical property of the modifier (the $P(x)(w)$ conjunct in the entry for *blond* in (2.7b)), which lacks the elegance of predicate modification.

However, some modifiers simply can't be combined with the expressions they modify via conjunction. The famous case in point are NON-INTERSECTIVE subsective adjectives. For example, both *blond* and *skillful* are subsective adjectives, as evidenced by the fact that the entailment that Zoe is a stuntwoman goes through both in (2.8a) and (2.8b). However, if the extension of *blond* in a given world can be taken to denote the set of individuals who are blond in that world, which allows us to think of the extension of *blond stuntwoman* as the intersection of the set of blond individuals and the set of stuntwomen, *skillful* doesn't lend itself to such treatment. This is evidenced by the fact that in (2.8a) we can infer that Zoe is a blond cellist from Zoe being a blond stuntwoman and a cellist, but in (2.8b) we cannot infer that Zoe is a skillful cellist from her being a skillful stuntwoman and a cellist.

- (2.8) a. Zoe is a blond stuntwoman.

Zoe is a cellist.

→ Zoe is a stuntwoman.

- Zoe is a blond cellist.
- b. Zoe is a skillful stuntwoman.
- Zoe is a cellist.
-
- Zoe is a stuntwoman.
- ↗ Zoe is a skillful cellist.

A standard solution is to have *skillful* denote a function from sets to subsets thereof (see Morzycki 2015, section 2.3 for an overview of non-intersective subsective adjectives). Whether all modifiers need to be generalized to the worst case and treated as functions that take the expression they modify as input is, however, immaterial for my purposes, which is why in the rest of the dissertation I will use predicate modification to compose modifiers with the expressions they modify, when it's possible, for the sake of simplicity. I will, however, maintain the schema in (2.6) with an explicit conjunct in the truth conditions assuring the subsective entailment for both intersective and non-intersective modifiers instead of, say, treating the subsective entailment of *skillful* as an idiosyncratic property of the **skillful** relation.

Now, one problem with using predicate modification is that we can't maintain a purely semantic notion of modification; if we completely ignore syntactic labels, there is no way of telling which of the two conjoined expressions is being modified and which one is the modifier. Thus, either we need to go with higher type modifiers and functional application across the board (which might still not be enough), or we need to let some syntactic information seep into our definition of a modifier (namely, which expression syntactically projects its label/features up the tree). Coming up with a 100% foolproof definition of a modifier is not one of the goals of this dissertation, however, so I will not attempt to fine-tune it any further. I will just assume that there is always a way to tell which expression is the modifier in a modification configuration.

Now, let me add a quick note about NON-SUBSECTIVE MODIFIERS. Adjectives like *alleged* in (2.9) notoriously don't give rise to the subsective entailment:

(2.9) Daisy is an alleged criminal.

↗ Daisy is a criminal.

The standard way to handle non-subsective modifiers is to make use of the worlds in our lexical entries (which so far have not been doing much work). For example, *criminal* would denote a relation that holds between an individual x and a world w if and only if x is a criminal in w , and *alleged* would take this relation and return a relation that holds between an individual x and a world w if and only if x is a criminal in all the worlds that are compatible with what's alleged in w (see, e.g, Morzycki 2015 for details). I will not touch upon non-subsective modifiers at all, however. Their projection behavior, to my knowledge, has not been studied, but when they project, they certainly don't give rise to the same inferences as non-restricting subsective modifiers, so they don't fall within the immediate purview of this dissertation. Also, just like with the term MODIFIER, I will often use the term ADJECTIVE to mean SUBSECTIVE ADJECTIVE.

Now, before I proceed to talking about projection of modifiers, let me add a few more examples of modifiers.

Adjectives are one type of adnominal modifiers. Other types include, for example, NP-adjoining PPs (as in *the beer on the table*) and restrictive relative clauses (as in *the dog that is large*). For the verbal domain I'm assuming standard Neo-Davidsonian treatment of modification (Carlson 1984; Parsons 1990, a.o.), under which all non-supplement composition within the verbal domain before existential closure happens via intersecting sets of events. Under this approach, there is no argument vs. adjunct distinction with respect to semantic composition, and in a sentence like *Zoe stabbed the extra with a knife*, both *the extra* and *with a knife* will be modifiers from the compositional point of view.

What about degree adverbs like *completely*, as in (2.10)?

(2.10) The glass is completely full.

→ The glass is full.

Intuitively, *completely* should be a subsective modifier, as the subsective entailment goes through. However, under the classical treatment of degree adverbs in Kennedy & McNally 2005, *completely* isn't a modifier under the definition assumed here, because *full* and *completely full* are not of the same semantic type. I will propose an—independently motivated—adjustment to Kennedy and McNally's analysis in subsection 2.1.2.4, which will allow me to treat degree expressions as bona fide modifiers and, as a result, apply my analysis of non-restricting modifier inferences to them.

2.1.2 Projection of modifiers

2.1.2.1 Restricting vs. non-restricting modifiers

All subsective modifiers are RESTRICTIVE, which means that they have the compositional potential to restrict the denotation of the expression they modify, i.e., to yield logically stronger expressions. However, not all specific instances of modifiers realize this potential within a given utterance, i.e., not all instances of modifiers are RESTRICTING. The tension between having a potential to be restricting and realizing it has been noted at least as early as in Potts 2005 and was further discussed, for example, in Leffel 2014 and Schlenker To appear.

The intuitive distinction between restricting vs. non-restricting instances of modifiers is illustrated in (2.11). Both definite descriptions in (2.11) pick out the same person in the actual world, Allison Anders. Both contain the adjective *female*, which in both cases composes as a subsective modifier and is thus restrictive. However, in (2.11a) *female* is not restricting the expression denoted by *director of 'Mi Vida Loca'*, as this film was directed by Anders alone. ‘Four Rooms’, however, consists of four novellas, directed by four different people, of which Anders was the only woman. Thus, in (2.11b) *female* does restrict the expression denoted by *director of 'Four Rooms'*.

- (2.11) a. the female director of ‘Mi Vida Loca’
 b. the female director of ‘Four Rooms’

Intuitively, restricting modifiers are used to pick out smaller parts of the denotations of the expres-

sions they modify while non-restricting modifiers are used to add extra information about these denotations without changing them. I will call this extra information added by non-restricting modifiers the NON-RESTRICTING MODIFIER INFERENCE. For example, in (2.11a) the non-restricting modifier inference is that the director of ‘Mi Vida Loca’ is female. Now, how do we capture this intuitive distinction?

Leffel (2014) talks about non-restricting modifiers as truth-conditionally vacuous modifiers, i.e., modifiers that don’t affect the truth conditions of the utterance in which they appear. He illustrates this observation via examples like (2.12). Under the world-knowledge based assumption that not all chemicals are harmful, but all toxins are, (2.12a) doesn’t entail a version of itself without the adjective, but (2.12b) does.

- (2.12) a. I will eliminate every harmful chemical.
 $\not\rightarrow$ I will eliminate every chemical.
 b. I will eliminate every harmful toxin.
 \rightarrow I will eliminate every toxin.

More specifically, Leffel (2014) defines NON-RESTRICTING MODIFIERS⁵ as (i) (subsective) modifiers intended by the speaker as truth-conditionally vacuous (the non-restrictingness condition proper), (ii) the inference contributed by the non-restricting modifier is relevant in the discourse (the relevance constraint). The relevance constraint seems to be related to the *Minimize Restrictors!* pragmatic principle proposed in Schlenker 2005, whereby truth-conditionally vacuous modifiers are dispreferred unless they serve a pragmatic function such as expressing the speaker’s attitude, reminding the addressee about a contextually relevant piece of information, etc. Leffel’s (simplified) definition of non-restricting modifiers is given in (2.13).⁶

⁵He uses the terms “non-restrictive” and “non-restricting” differently from me.

⁶He formalizes the relevance constraint; for the purposes of this dissertation this is not necessary, so I am omitting his formalisms for simplicity.

(2.13) **Non-restricting modifiers in Leffel 2014, (3.58a) (simplified)**

An occurrence α_n of a modifier α in the phrase $[\text{DP} \dots \alpha_n \dots \text{N} \dots]$ or $[\text{DP} \dots \text{N} \dots \alpha_n \dots]$ is non-restricting with respect to index i iff:

- a. the speaker believes that $\llbracket \text{DP} \rrbracket^i = \llbracket \text{DP}[\alpha_n/\varepsilon] \rrbracket^i$,
where $X[a/b]$ is the expression just like X except with a token of a replaced by one
of b and ε is the empty string; and (non-restricting intention)
- b. there is a discourse relation between the non-restricting modifier inference and some
discourse unit in the context. (relevance constraint)

I propose a refined version of Leffel's original definition in (2.14), which introduces the following changes. First, Leffel's definition is restricted to adnominal modifiers; I make it more general. Also, Leffel eventually excludes non-subsective adjectives like *former* and *alleged* from consideration, but this restriction is not part of his original definition; I make it explicit in the definition. Most importantly, however, I switch to global, utterance-level equivalence from local, constituent-level one. As we will see in subsection 2.1.2.3, Leffel's original definition is at odds with the projecting behavior of non-restricting modifier inferences, and we need to separate the definitional utterance-level truth-conditional vacuity of non-restricting modifiers, which holds globally, from the inference about the local equivalence of the expression being modified and the result of modification, which is sensitive to local contexts.

(2.14) **NON-RESTRICTING MODIFIER (adopted here)**

If an utterance u contains a subtree α whose daughters are γ and β and γ is an occurrence of a subsective modifier, γ is NON-RESTRICTING at an index i iff:

- a. the speaker of i believes that $\llbracket u \rrbracket^i = \llbracket u[\gamma/\varepsilon] \rrbracket^i$,
where $X[a/b]$ is the expression just like X except with a token of a replaced by one
of b and ε is the empty string; and (non-restricting intention)
- b. there is a discourse relation between the non-restricting modifier inference and some

discourse unit in the context.

(relevance constraint)

It is tempting then to define a restricting modifier by simply negating the condition about the speaker's beliefs in (2.14a) and deleting (2.14b), as in (2.15).

(2.15) **RESTRICTING MODIFIER (weak)**

If an utterance u contains a subtree α whose daughters are γ and β and γ is an occurrence of a subsective modifier, γ is RESTRICTING at an index i iff it is not the case that the speaker of i believes that $\llbracket u \rrbracket^i = \llbracket u[\gamma/\varepsilon] \rrbracket^i$.

That would assure that a restricting modifier is one that affects the truth conditions of the utterance in which it occurs, but that would be too weak, however. Simply saying that a restricting modifier is one that is truth-conditionally non-vacuous clashes with our intuition that restricting modifiers are used by the speaker intentionally to distinguish among (non-empty or potentially non-empty) parts within the denotations of the expressions they modify. In other words, it is not enough for us to assure the lack of non-restricting intention; we need something like the “restricting intention”. A refined definition is thus given in (2.16).

(2.16) **RESTRICTING MODIFIER (strong)**

If an utterance u contains a subtree α whose daughters are γ and β and γ is an occurrence of a subsective modifier, γ is RESTRICTING at an index i iff:

- a. it is not the case that the speaker of i believes that $\llbracket u \rrbracket^i = \llbracket u[\gamma/\varepsilon] \rrbracket^i$; and
(lack of non-restricting intention)
- b. the speaker of i uses γ to pick out a potentially non-empty part of the denotation of β .
(restricting intention)

The restricting intention part of the definition in (2.16) is quite vague and eventually needs to be made more precise. I will not do so here, however, as I will mostly focus on non-restricting modifiers in the rest of the dissertation, but I will come back to the issue in subsection 2.1.2.3, and

it will become relevant again in section 5.3 when I talk about *phi*-features on pronouns. I will keep the two parts of the definition in (2.16) separate, however, to emphasize that I am allowing for the possibility that the speaker is not opinionated about whether γ picks out a proper part of β or not.

2.1.2.2 Non-restricting modifier inferences as semantic presuppositions (Leffel 2014)

Leffel raises the question of how non-restricting modifiers make their contribution to the meaning of the utterance, i.e., how the non-restricting modifier inference arises. He proposes that non-restricting modifiers compose differently from restricting ones, and the non-restricting modifier inference is generated as a presupposition during this composition. Under Leffel's story, non-restricting modifiers in fact compose with individuals, either because they compose with a kind-denoting N, as would be the case in (2.17a), or because they adjoin at the D' level within a DP-shell structure (Larson 1991), as would be the case in (2.17b).

- (2.17) a. Cigarettes contain harmful toxins.
 \rightarrow All toxins are harmful.
- b. I take care of my sick mother.
 \rightarrow My mother is sick.

Leffel posits the composition rule in (2.18), where the underlined part is a SEMANTIC PRESUPPOSITION, i.e., a presupposition that is triggered in the semantics (here it is hardcoded into the definition of a semantic composition rule). This contrasts with a PRAGMATIC PRESUPPOSITION, which arises via pragmatic reasoning about the speaker's beliefs and intentions.

(2.18) Type mismatch adjustment principle (TMAP) in Leffel 2014, (3.86)

Let $\alpha : \langle \sigma, t \rangle, \beta : \sigma$, and assume XP is not a clausal syntactic category.

Then $\llbracket_{[XP} \alpha \beta \rrbracket = [\underline{\alpha(\beta)}].[\beta]$.⁷

⁷Leffel's original formulation contains typos; the string after *Then* reads as $\lambda x_\sigma [\underline{\alpha(\beta)}].[\alpha]$. I change the definition to match the results in (2.19) and (2.20).

For the two cases in (2.17) the rule is supposed to apply as follows:

(2.19) **Composition of $[DP_1 [D_i \text{ my}] [DP_2 [\text{AP sick}] [D' t_i [NP \text{ mother}]]]]$ in Leffel 2014, (3.92)⁸**

$$\begin{aligned} [[DP_2]] &= \mathbf{TMAP}([[AP]])([[D']]) = \mathbf{TMAP}(\lambda x.\text{sick}(x))(\iota x.\text{mother}_{\text{me}}(x)) \\ &= [: \underline{\text{sick}(\iota x.\text{mother}_{\text{me}}(x))}].[\iota x.\text{mother}_{\text{me}}(x)] \end{aligned}$$

(2.20) **Composition of $[N' [N_1 [\text{AP harmful}] [N_2 \text{ toxins}]]]$ in Leffel 2014, (3.96)**

$$\begin{aligned} [[N_1]] &= \mathbf{TMAP}([[AP]])([[N_2]]) = \mathbf{TMAP}(\lambda x_{\langle s,e \rangle}.\text{harmful}(x))(\lambda w.\iota * \text{toxin}_w) \\ &= [: \underline{\text{harmful}(\lambda w.\iota * \text{toxin}_w)}].[\lambda w.\iota * \text{toxin}_w] \end{aligned}$$

In a way, Leffel makes composition of non-restricting modifiers very similar to Pottian (2005) composition of supplements (which we will see in subsection 2.2.1), except instead of the conventional implicature dimension, he has a semantic presupposition dimension (he himself acknowledges the parallels in section 5.1 of Leffel 2014).

Leffel also mentions cases like (2.21) from Larson & Marušič 2004, (31a), for which a non-restricting construal of the modifier is also possible.

(2.21) Every unsuitable word was deleted.

- ✓ Every word that was unsuitable was deleted. (restricting)
- ✓ Every word was deleted; they were unsuitable. (non-restricting)

Leffel doesn't provide a derivation for the non-restricting reading in (2.21), but says that in this case unsuitability would be "predicated of some contextually salient plurality of words (which could be conceived of as a subkind of "words")" (Leffel 2014, p. 78), in effect, partially following Morzycki (2008), who proposes that all non-restricting modifiers contribute Pottian (2005) conventional implicatures about contextually salient definite descriptions. I make this note here to highlight that even such a heavily syntactic analysis as Leffel's at the end of the day needs to resort to contextual domain restriction of some sort to account for non-restricting modifier inferences across the board.

⁸The assumption here is that D_i reconstructs.

That aside, there are two major issues with Leffel's analysis. One is the ad hoc nature of the TMAP composition rule (which Leffel himself acknowledges) and, more generally, the disconnect between the highly pragmatic nature of non-restricting interpretations of modifiers (recognized in the definition of non-restricting modifiers in (2.13), which relies on the speaker's beliefs and discourse relations) and the highly grammaticalized nature of Leffel's analysis of non-restricting modifier inferences.

The other issue is that it is unclear how this analysis would generalize to non-intersective or non-adnominal modifiers (e.g., degree modifiers or modifiers in the verbal domain⁹), given the nature of the presupposition yielding the non-restricting modifier inference in his analysis as well as how much his analysis relies on very specific syntax/semantics mapping within the nominal domain. The presupposition in (2.18) applies the modifier to the expression being modified yielding something of type t ; it is designed for cases when the modifier is of type $\langle(s)e, t\rangle$ and the expression being modified is of type $(s)e$. It's unclear how this presupposition would work in cases when a degree modifier applies to an adjective or a non-intersective subsective adjective applies to a noun.

2.1.2.3 Proposal: non-restricting modifier inferences as pragmatic presuppositions

In this subsection I propose to treat non-restricting modifier inferences as pragmatic presuppositions that are immediately tied to the definition in (2.14). In subsection 2.1.2.4 I will link these pragmatic presuppositions to COSUPPOSITIONS proposed in Schlenker 2018a,b,c.

This approach keeps the composition of restricting and non-restricting modifiers the same across the board, thus, embracing the restricting vs. non-restricting distinction as a pragmatic phenomenon. The proposed approach is also easily generalizable to non-intersective and non-adnominal modifiers. While developing this proposal, I also explore the projection behavior of non-restricting modifier inferences empirically and conclude that while they do seem to project like presuppositions (i.e., relative to local contexts), they cannot be interpreted locally, unlike many lexical presuppositions. I show that under the proposed analysis such unattested local interpretations

⁹Leffel gives examples of non-restricting modifiers in the verbal domain but doesn't show how to apply his analysis to them.

can't arise; if the non-restricting modifier inference can't project due to some contextual pressure, it simply won't be triggered.

Let me start by making explicit the point (reflected both in Leffel's original definition in (2.13) and my definition in (2.14)) that, unlike the subsective entailment, the definitional truth-conditional vacuity of non-restricting modifiers is context-sensitive, so it is crucial to specify the worlds relative to which it is assessed. For example, *director of 'Mi Vida Loca'* entails *female director of 'Mi Vida Loca'* in the actual world, but it is easy to imagine worlds in which that is not the case. Furthermore, we could imagine a speaker who knows that Anders is a woman, but they mistakenly assume that she co-directed 'Mi Vida Loca' with a man. If they uttered a sentence containing the definite description *the female director of 'Mi Vida Loca'*, that instance of *female* would not be non-restricting. Similarly, the inferential judgements reported in (2.12) hinge on one's beliefs about chemicals and toxins. For example, if the sentence in (2.12a) is uttered by someone who thinks that all chemicals are harmful (perhaps, due to their lexical entry for *chemical* being different from the one assumed in chemistry), the entailment will go through, and that instance of *harmful* will be non-restricting in this case.

Leffel (2014) intends to capture this context-sensitivity in his original definition in (2.13) by relativizing the equivalence of the expression being modified and the result of modification to the speaker's beliefs. In his analysis, he intends to derive this equivalence by generating a semantic presupposition assuring it. However, presuppositions project in a way that doesn't always assure that the speaker believes their content. As it turns out, the same is true about non-restricting modifier inferences. In particular, as we will see shortly, the speaker can use modifiers in a non-restricting way without globally subscribing to the belief that the denotations of the expression being modified and the result of modification are equivalent. This is why in my definition in (2.14) I switched from equivalence of the expression being modified and the result of modification to equivalence of the utterances with and without the modifier. The two equivalences need to be kept separate, because while the latter equivalence always holds globally in cases of non-restricting modification, the former equivalence is sensitive to its local context.

I propose that it is this equivalence between the expression being modified and the result of modification that is the source of the non-restricting modifier inference. We can write this equivalence as $\alpha \Leftrightarrow_{\forall} \beta$, where α is the result of modification and β is the expression being modified. Since we are only dealing with subsective modifiers, i.e., $\alpha \Rightarrow_{\forall} \beta$ necessarily holds whether the modifier is restricting or not, we are only really concerned with the entailment from the expression being modified to the result of modification, i.e., $\beta \Rightarrow_{\forall} \alpha$.

The $\beta \Rightarrow_{\forall} \alpha$ inference projects like a presupposition, i.e., relative to a context. As a first approximation, let us say that it needs to be entailed by the context set C (Stalnaker 1975 et seq.), which is, roughly, the set of worlds compatible with the joint beliefs of the speech act participants. Now imagine a situation in which the addressee of an utterance containing a modifier infers that the speaker intends for this instance of this modifier to be non-restricting, but the addressee did not previously believe that the expression being modified entails the result of modification. They will still interpret the contribution of this modifier instance, and the way they will do so is reminiscent of GLOBAL ACCOMMODATION of lexical presuppositions.¹⁰

An example of an utterance that is likely to produce such a result is given in (2.22). The adjective *obnoxious* has a strong evaluative component and is thus often non-restricting (I will come back to preferably non-restricting adjectives later in subsection 2.1.2.5). The addressee of (2.22) is therefore likely to infer that the speaker intends to use this adjective as a non-restricting one, i.e., one that doesn't affect the truth conditions of their utterance. The addressee will then infer that the speaker considers all philosophers obnoxious, even if they didn't previously believe that, and will tacitly update the context set accordingly. (Or, alternatively, if they aren't willing to globally accommodate this inference, they might challenge the speaker on the issue, e.g., by saying *Hm, I didn't know you don't like philosophers....*)

(2.22) I don't want any obnoxious philosophers at my talk next week.

→ The speaker doesn't want any philosophers at their talk next week.

¹⁰The notion of global accommodation was discussed as early as Karttunen 1974 and Stalnaker 1974; the term “accommodation” was introduced in Lewis 1979; the “global” part was added in Heim 1983 to distinguish it from local accommodation.

→ The speaker believes that all philosophers are obnoxious.

However, just like lexical presuppositions, non-restricting modifier inferences are (potentially) sensitive to their LOCAL CONTEXT, i.e., roughly, the context set as updated throughout the utterance (Karttunen 1974; Stalnaker 1974; Heim 1983; Schlenker 2009, a.o.). For example, in (2.23), the addressee is once again likely to think that the speaker intended for the adjective *obnoxious* to be truth-conditionally vacuous. However, this sentence doesn't give rise to a global inference that the speaker believes all philosophers to be obnoxious. Instead, the inference contributed by *obnoxious* is conditional, as it is interpreted relative to the original context set C updated with the content of the antecedent of the conditional.

(2.23) *Context: The speaker just observed a philosopher ask an annoying question at a talk.*

If all philosophers ask questions like this, I don't want any obnoxious philosophers at my talk next week.

→ If all philosophers ask questions like this, the speaker doesn't want any philosophers at their talk next week.

↗ The speaker believes that all philosophers are obnoxious.

→ The speaker believes that if all philosophers ask questions like this, they are all obnoxious.

I adduce a few more examples of local context sensitivity of non-restricting modifier inferences in (2.24). None of them gives rise to an inference that the speaker believes in the equivalence of the expression being modified and the result of modification, but in all the cases the utterance with the modifier is equivalent to a version of itself without the modifier and the equivalence between the expression being modified and the result of modification is established relative to the local context of the modification configuration.

(2.24) a. Either Pam's mother has recovered, or Pam is still taking care of her sick mother.

- Either Pam's mother has recovered, or Pam is still taking care of her mother.
- ↗ Pam's mother is sick.
- If Pam's mother hasn't recovered, Pam's mother is sick.
- b. Maybe processed meat causes cancer and I should never eat any deadly sausages again.
- Maybe processed meat causes cancer and I should never eat any sausages again.
- ↗ All sausages are deadly.
- If processed meat causes cancer, all sausages are deadly.

For comparison, local context sensitivity of lexical presuppositions is illustrated for *know* in (2.25).

- (2.25) a. Zoe knows that she can't continue the race.
- Zoe can't continue the race.
- b. If Zoe's car has been damaged, she knows that she can't continue the race.
- c. Either Zoe's car hasn't been damaged, or she knows that she can't continue the race.
- d. Maybe Zoe's car has been damaged and she knows that she can't continue the race.
for (b–d): ↗ Zoe can't continue the race.
for (b–d): → If Zoe's car has been damaged, she can't continue the race.

The hedging above in “(potentially) sensitive” is due to the fact that we do often get stronger inferences than one would expect if local contexts are always taken into account. For example, this issue arises for lexical presuppositions in the case of the infamous proviso problem (Geurts 1996 et seq.), illustrated in (2.26). Many local context-based analyses of presupposition projection (Heim 1983 et seq.) predict conditional inferences in both examples in (2.26), however, the inference is only conditional in (2.26a).

- (2.26) a. If Bridget is a spy for the Allies, she won't expose her British contact.
- If Bridget is a spy for the Allies, she has a British contact.

- b. If Bridget is smart, she won't expose her British contact.

$\not\rightarrow$ If Bridget is smart, she has a British contact.

\rightarrow Bridget has a British contact.

The proviso problem arises for non-restricting modifier inferences, as well, which further solidifies their status as presuppositions.

- (2.27) a. If processed meat causes cancer, you shouldn't eat so many deadly sausages.

\rightarrow If processed meat causes cancer, all sausages are deadly.

- b. If you care about your health, you shouldn't eat so many deadly sausages.

$\not\rightarrow$ If you care about your health, all sausages are deadly.

\rightarrow All sausages are deadly.

To sum up, inferences contributed by non-restricting modifiers do indeed project similarly to lexical presuppositions, i.e., relative to local contexts. However, there is no need to assume that they are triggered semantically. The addressee can reason about the speaker's intentions, relying, in particular, on the discourse relation mentioned in (2.14b). If they conclude that the speaker intended a given modifier instance as a non-restricting, i.e., a truth-conditionally vacuous one, they will generate the non-restricting modifier inference, i.e., the inference that the expression being modified β entails the result of modification α in the local context of α , which they can then globally accommodate or challenge.

We are now getting close to defining non-restricting modifier inferences formally. I will adopt Schlenker's (2009; 2010) general pragmatic approach to local contexts, which defines the local context c' of any expression d whose type ends in t as the strongest possible restriction one could make before interpreting d in a Stalnakerian context set C ; a full definition is given in (2.28).

- (2.28) **LOCAL CONTEXT (adopted with minor changes from Schlenker 2010, (7))**

The LOCAL CONTEXT of an expression d that occurs in a syntactic environment $a _ b$ in a

context set C is the expression with the strongest denotation c' which guarantees that for any expression d' of the same type as d , for all strings b' for which $a \ d' \ b'$ is a well-formed sentence, the following holds in each world w in C :

$$a (c' \text{ and } d') b' \leftrightarrow a d' b'$$

We will need to keep track of two local contexts. One is the propositional local context c' computed via the mechanism in (2.28) for the clause that contains the modifier (e.g., by updating the global context with an antecedent of a conditional, as in (2.23), the negation of the left disjunct, as in (2.24a), or the left conjunct under *maybe*, as in (2.24b)). The second local context is a sub-propositional local context c'' , which is computed within the clause for the result of modification and is always relativized to c' .¹¹ In simple cases (which are the only ones I will consider in this) this means that the world argument of c'' is restricted to the worlds of c' .

We can now define the NON-RESTRICTING MODIFIER INFERENCE as follows:

(2.29) **NON-RESTRICTING MODIFIER INFERENCE**

If a clause p contains a subtree α whose daughters are γ and β and the addressee infers that the speaker intends γ as a non-restricting modifier (as defined in (2.14)), the addressee will generate the NON-RESTRICTING MODIFIER INFERENCE:

$$c'' \Rightarrow_{\forall} (\beta \Rightarrow \alpha),$$

where c'' is the local context of α in the local context c' of p .

Now let us apply the definition above to (2.22) and (2.23). For now let us not worry too much about

¹¹The need to keep track of two local contexts, a propositional and a sub-propositional one, is not specific to non-restricting modifier inferences. The issue is not discussed explicitly in Schlenker's (or, to my knowledge, anyone else's) work on local contexts; he only discusses sub-propositional contexts for expressions in the scope of quantifiers and does not bring up cases in which both the propositional and the sub-propositional local contexts matter for projection of lexical presuppositions, such as (i), where to compute the projecting inference we need to first restrict our attention to the worlds in which the track is damaged and then to the individuals who are stuntwomen in those worlds.

- (i) If the track is damaged, each stuntwoman knows that she can't continue the race.
→ If the track is damaged, no stuntwoman can continue the race (i.e., each stuntwoman is such that she can't continue the race).

the sub-propositional local context c'' (I will come back to the issue in Chapter 4) and assume that it is maximally broad in both cases, i.e., that it denotes $\lambda x \lambda w. D_e(x) \wedge \llbracket c' \rrbracket(w)$, where D_e is the domain of individuals in the model. The addressee hears (2.22) and deduces that *obnoxious* is likely to be non-restricting due to its highly evaluative nature, so they generate the non-restricting modifier inference as defined in (2.29). The propositional local context c' of the clause containing *obnoxious philosophers* in (2.22) denotes the global context set C , and we have agreed to treat c'' as maximally broad, so the addressee will generate the inference in (2.29) and will globally accommodate or challenge it if they didn't believe it previously.

$$\begin{aligned}
(2.30) \quad & \forall x \forall w [\llbracket c'' \rrbracket(x)(w) \rightarrow (\llbracket \text{philosopher} \rrbracket(x)(w) \rightarrow \llbracket \text{obnoxious philosopher} \rrbracket(x)(w))] \\
& \forall x \forall w [(D_e(x) \wedge \llbracket c' \rrbracket(w)) \rightarrow \\
& \quad (\text{philosopher}(x)(w) \rightarrow (\text{philosopher}(x)(w) \wedge \text{obnoxious}(x)(w)))] \\
& \forall x \forall w [(D_e(x) \wedge C(w)) \rightarrow \\
& \quad (\text{philosopher}(x)(w) \rightarrow (\text{philosopher}(x)(w) \wedge \text{obnoxious}(x)(w)))] \\
& \forall x \forall w [(D_e(x) \wedge C(w)) \rightarrow (\text{philosopher}(x)(w) \rightarrow \text{obnoxious}(x)(w))]
\end{aligned}$$

The above is a simplification in that the lexical semantics of *obnoxious* should be more complex and linked to the attitudes of the speaker, but other than that, the result corresponds to the intuitive inference in (2.22).

In (2.23), however, the propositional context c' of the clause containing *obnoxious philosophers* does not denote C , but rather the intersection of C with the set of worlds denoted by the antecedent of the conditional, i.e., the set of worlds in which all philosophers ask questions like this, which I will abbreviate as Q (see Schlenker 2009, 2010 for the full proof of this being the strongest possible propositional restriction in this case). The inference arising in (2.23) is given in (2.31) (I omit the intermediate steps).

$$(2.31) \quad \forall x \forall w [(D_e(x) \wedge C(w) \wedge Q(w)) \rightarrow (\text{philosopher}(x)(w) \rightarrow \text{obnoxious}(x)(w))]$$

Now, one might be concerned about the non-deterministic nature of the pragmatic triggering in (2.29).¹² However, this is a good thing, as we often do observe a lot of ambiguity as to whether a given instance of a modifier is meant to be non-restricting. For instance, we have seen an example of this in (2.21), which, in the absence of further contextual information, is ambiguous.

Non-restricting modifier inferences project very STRONGLY, i.e., they usually cannot be interpreted locally under semantic operators, not even under pressure (e.g., if a local interpretation is the only way to prevent the utterance from being contradictory).

For example, non-restricting modifier inferences project more strongly than existence inferences of definite descriptions. This is illustrated in (2.32). In (2.32a), the inference of the definite description *her car* that Pam has a car cannot be satisfied globally, but it can be marginally interpreted locally under *if*, resulting in an imperfect but relatively acceptable utterance. Such local interpretation should then be possible in (2.32b) as well, however, either it isn't or it's not enough to save the day; the utterance is crashingly bad.

(2.32) *Context: Zoe, Lucy, and Pam are going on a camping trip. Pam is coming from Boston to New York to join the rest of the group. Zoe and Lucy are discussing how to get to the camping site from New York. ...*

- a. ...*They have agreed that they need a car, no matter how big. Lucy:*

?Well, Pam is coming from Boston. I don't know if she **has** a car, but if she's coming in her **car**, we can use it to get to the camping site.

'...if (she has a car and she's coming in her car)...'

- b. ...*They have agreed that they need a large car to fit all their supplies. Lucy knows that Pam has exactly one car. Lucy:*

#Well, Pam is coming from Boston. I don't know if her car is **large**, but if she's coming in {her **large** car, her large **car**}, we can use it to get to the camping site.

Intended: '...if ({she has a large car, her car is large}) and she's coming in her large

¹²Note also that Leffel's analysis is also non-deterministic, at least for pre-nominal adjectives in English, as it allows for both restricting and non-restricting compositional construals for one and the same string.

car)...’

I attribute the total unacceptability of (2.32b) to the fact that the non-restricting modifier inference of the adjective *large*¹³ is not satisfied globally, but, unlike existence inferences of definite descriptions, it cannot be interpreted locally. I thus conclude that even in cases of non-restricting modification within definites, the non-restricting modifier inference is independent of the existence inference of the definite.

The unacceptability of (2.32b) illustrates another important point, namely, the necessity to independently capture the “restricting intention” of restricting modifiers discussed at the end of section 2.1.2.1. The weak definition in (2.15) doesn’t prevent us from saying that *large* in (2.32b) is, in fact, a restricting modifier, since all we require there is that restricting modifiers aren’t truth-conditionally vacuous, which would be the case if Pam doesn’t have a large car in (2.32b). In that case, we would not have a non-restricting modifier inference, and locally accommodating the existence inference of the definite description should be enough to save the day. However, empirically it looks like we simply can’t use modifiers to distinguish between two alternatives, one in which the modification is vacuous and one in which the result of modification is empty. Thus, whatever the precise implementation of the restricting intention in (2.16) is, its consequence should be that we infer that the speaker believes that the modifier picks out a locally non-empty part of the denotation of the expression it modifies, where by “locally” I mean that these parts need to be non-empty relative to the local context of the modification configuration, as illustrated in (2.33).

- (2.33) If Stephanie has a small and a large dog, I would rather she brought her **large** dog, not her **small** dog.

¹³Here and for all similar examples I am assuming that the modifier is operating on the set of contextually relevant individuals. In this case it would be Pam’s cars only, which is a singleton set, so the adjective is non-restricting. However this domain restriction is operationalized, it seems to be empirically solid: we don’t get an inference about all the cars in the world, only Pam’s. It’s also possible to have a story similar to Leffel’s (see (2.19)), where the head noun in a possessive DP bears an index linking it to the possessor, except I would maintain that *car_i* is still of type *et*. Note also that, as mentioned before, Leffel still has to assume non-syntactic, contextual domain restriction for some other cases of non-restricting modification, such as in (2.21). See also the discussion in Schlenker 2018a, fn. 31, where the potential need for an assignment-sensitive domain restriction is acknowledged for gestures.

In other words, restricting modifiers come with their own “restricting modifier inferences”, which arise pragmatically as a result of the function of restricting modifiers and project relative to local contexts. I will not explore these inferences any further, since this dissertation will mostly focus on non-restricting modifier inferences. However, the fact that restricting modifier inferences exist and that they are independent of simple existence inferences of definites assures that, under the assumption that each instance of a subsective modifier is either restricting or non-restricting, whenever we have a subsective modifier modifying an expression whose extension is believed by the speaker to be a singleton set, this instance of a modifier has to be non-restricting, as the only way for it to be truth-conditionally non-vacuous would be if the extension of the result of modification could be the empty set, contra the restricting modifier inference. This consequence will become important later in section 5.3 when we talk about *phi*-features on pronouns as obligatorily non-restricting modifiers.

Now, going back to non-restricting modifier inferences, they project very strongly in other cases, too. Thus, (2.34) is only acceptable if the adjective *obnoxious* is interpreted as a restricting one. If *obnoxious* was non-restricting, the *I don't know...* clause would have prevented the non-restricting modifier inference that all philosophers are obnoxious from projecting, so it would have to be interpreted locally, as a conjunct under *if*, but this is not a possible interpretation.¹⁴

- (2.34) I don't know if all philosophers are obnoxious or if very few are, but if every obnoxious philosopher comes to my talk, I'll be annoyed.
- ↗ All philosophers are obnoxious.
- ✗ ‘...if (all philosophers are obnoxious and every philosopher comes to my talk), I'll be annoyed.’

In presupposition literature, local interpretations of normally projecting inferences are standardly derived via LOCAL ACCOMMODATION (e.g., Heim 1983; Schlenker 2009), which is a last-resort

¹⁴The local interpretation isn't inconceivable either; one could easily imagine a scenario in which the speaker would only be annoyed if the number of obnoxious philosophers in the audience is large.

mechanism that is in principle available for presuppositions. Local accommodation can be operationalized differently, but at the end of the day yields the result where the presupposition is interpreted as a conjunct at some local level, which is what seems to happen in (2.32a).

The presupposition literature typically distinguishes between WEAK and STRONG presuppositions, i.e., those that in principle can be locally accommodated under some pressure and those that can't. The terms WEAK and STRONG are also often used to describe the presupposition triggers that trigger weak and strong presuppositions, respectively. For example, when it comes to lexical presuppositions, *start* is usually considered a weak trigger, as shown in (2.35), and *again* is usually considered a strong trigger, as shown in (2.36).

- (2.35) I don't know if Jackie smokes, but if she starts now, it'll be hard for her to quit.

↗ Jackie doesn't smoke.

≈ ‘...if (she doesn't smoke and starts now)...’

- (2.36) ??I don't know if Jackie has been in jail before, but if she goes to jail again, she's unlikely to be released on parole.

↗ Jackie has been in jail before.

Intended: ‘...if (she has been in jail before and goes to jail again)...’

This distinction can be seen as a continuum from very weak to very strong presuppositions, depending on how acceptable local accommodation thereof is. As things stand, however, the weak–strong distinction for presupposition triggers is a descriptive, not an explanatory one.

In this dissertation I am using the weak–strong distinction more generally, for any projecting content, regardless of whether or not it is presuppositional. Non-restricting modifier inferences project very strongly, so if they are to be analyzed as presuppositions of some kind, we could simply stipulate that they are strong without losing any explanatory power. If we could, however, have a principled reason why local interpretations of non-restricting modifier inferences are not possible, we could gain some.

Under the pragmatic triggering mechanism I have proposed in this subsection, local accommodation is not a meaningful option for non-restricting modifier inferences. The addressee only generates non-restricting modifier inferences if they believe that the speaker intends the given instance of a modifier as a non-restricting, i.e., a truth-conditionally vacuous one, and if they have a reason to believe that this is not the case, no inference will be generated in the first place. In (2.32b) and (2.34), the target modifiers are clearly meant to affect the truth conditions, so no inference is generated in the first place and the modifiers can only be interpreted as restricting. In (2.32b), the restricting interpretation is impossible, as it is specified in the context that Pam only has one car, and also this interpretation clashes with the question under discussion (QUD; Roberts 2012 et seq.) set up by the speaker.

I would imagine that Leffel (2014) could say something similar to explain the lack of local accommodation of the semantic presupposition in (2.18), but in his system the link between the triggering of non-restricting modifier inferences and the definition of non-restricting modifiers is more opaque, so it is less clear how the argument would go.

Now, let me add a quick note about focus. It has been observed before that focus on an adjective forces a restricting interpretation thereof (the observation was originally made in Umbach 2006 for German and extended to English in Leffel 2014). For example, (2.37a) is a reasonable utterance that doesn't commit one to the existence of harmless carcinogens, but (2.37b) is weird because focus on *harmful* signals that the speaker believes harmless carcinogens exist, which is at odds with our world knowledge.

- (2.37) a. Cigarettes contain harmful carcinogens. (Leffel 2014, (3.10))
 b. #Cigarettes contain **harmful** carcinogens. (Leffel 2014, (5.66))

Leffel (2014) argues that the reason that happens is because focus on a modifier evokes contrasting alternatives that, furthermore, need to be defined (this second part is a stipulation). In (2.37b) a natural contrasting alternative to *harmful* would be *harmless*, and the set denoted by *harmless carcinogens* then would need to be non-empty.

Whether or not Leffel's explanation is correct, this property of focused modifiers alone could not explain the unacceptability of (2.32b) or the unavailability of the (local) non-restricting interpretation in (2.34), as these facts obtain regardless of whether the adjective is focused or not.

Furthermore, it is not obvious to me that focus does in fact always force restricting interpretations of modifiers. In Esipova 2018a, I discuss instances of focus that only marks contrast between two expressions but does not relate them to any QUD. Curiously enough, instances of non-QUD-marking focus don't seem to force restricting interpretations, as shown in (2.38), where the instances of *harmful* and *healthy* are likely to be non-restricting given our world knowledge.

- (2.38) We need to distinguish between **harmful trans** fats and **healthy monounsaturated** fats.
- All trans fats are harmful.
 - All monounsaturated fats are healthy.

The observation that focus on modifiers only seems to force a restricting interpretation thereof when it relates them to some QUD leads to an alternative explanation of the effect observed in Umbach 2006 and Leffel 2014. A piece of content that addresses a QUD often can't be truth-conditionally vacuous (the link is claimed to be categorical in Simons et al. 2010; see, e.g., Koev 2013 for some counterexamples involving appositives). The only way for a modifier not to be truth-conditionally vacuous is to be restricting; non-restricting modifiers are truth-conditionally vacuous by definition. This would explain the correlation between being focused and being restricting. However, the potentially non-categorical nature of the link between QUD-addressing and truth-conditional vacuity leaves space for further exploration (which I will not pursue here).

2.1.2.4 Non-restricting modifier inferences and cosuppositions

The pragmatic triggering mechanism for non-restricting modifier inferences developed in subsection 2.1.2.3 is reminiscent of COSUPPOSITIONS, i.e., assertion-dependent inferences that project like lexical presuppositions but aren't necessarily triggered as such. Cosuppositions were originally proposed in Schlenker 2018a to account for inferences triggered by co-speech gestures across the

board and extended to other types of content in Schlenker 2018b,c. In this dissertation I propose to fully equate cosuppositions with non-restricting modifier inferences. I argue that by doing so we will both constrain cosuppositions in the right way and let them apply in the cases in which Schlenker's original formulation doesn't. I will develop the first argument and provide empirical support for it in Chapter 4, when I talk about gestures and facial expressions. As for the second argument, I will develop it now by showing that my formulation of non-restricting modifier inferences applies to non-intersective modifiers while Schlenker's formulation of cosuppositions doesn't.

Schlenker's original motivation for introducing cosuppositions as a mechanism of generating projecting inferences contributed by co-speech gestures (in Schlenker 2018a) was to assure that their contribution is truth-conditionally vacuous. He then extended the notion of cosuppositions to other types of content. In (2.39), I give his full generalized definition of cosuppositions from Schlenker 2018c.¹⁵

(2.39) **Schlenker's unified theory of cosupposition generation (Schlenker 2018c, (67))**

- a. A cosupposition is triggered when an elementary expression pp' has an entailment p' which is presented as being unimportant, and for this reason the global Context Set C should guarantee that, relative to its local context c' , pp' should be equivalent to p , i.e.:
 - (i) $c' \Rightarrow_{\forall} (pp' \Leftrightarrow p)$
- b. (i) is equivalent to the standard definition of cosuppositions in (ii):¹⁶
 - (ii) $c' \Rightarrow_{\forall} (p \Rightarrow p')$
- c. An entailment p' might be presented as unimportant for different reasons:
 - (i) for reasons of manner, in case p' is contributed by a co-speech or co-sign gesture (which is parasitic and thus should not make an essential contribution);
 - (ii) for conceptual reasons, in case p' is understood not to matter given the context

¹⁵I change Schlenker's notation a bit to better match mine. Also, let me point out that while (2.39) only makes reference to the global context set C and a single local context c' , Philippe Schlenker (p.c.) has confirmed to me that for a non-propositional pp' we need to compute two local contexts, a propositional one and a sub-propositional one, as is made explicit in my formulation in (2.29).

of the conversation.

- d. Because cosuppositions are presented as unimportant, they can to some extent be disregarded under ellipsis resolution. (This does not preclude the possibility that co-speech/sign cosuppositions are more easily ignored for syntactic reasons.)¹⁷

Schlenker's algorithm in (2.39) is, on the one hand, unconstrained in that it isn't linked to any specific compositional configuration and can even apply if p and p' don't combine compositionally and are instead two pieces of a single lexical entry pp' .¹⁸ On the other hand, Schlenker assumes that p and p' are conjuncts within pp' (whether or not they combine compositionally) and that p' is thus an entailment of pp' , which is too restrictive, as this prevents cosuppositions from arising in cases where p and p' are of different semantic types (pp' can't entail p' if they are of different types, nor can p and p' conjoin in this case).

In Chapter 4, I will show that attempts to constrain the cases in which cosuppositions arise based on how a given piece of content is linearized without making reference to its compositional integration (featured most prominently in 2018b, but also in Schlenker 2018a) are misguided.

I propose instead to constrain cosuppositions by equating them with non-restricting modifier inferences, as defined in (2.29). Schlenker's take on what constitutes an “uninimportant entailment” in (2.39c) can then be re-conceptualized as some of the reasons why a given instance of a modifier might prefer to be non-restricting. I will discuss some other reasons why certain instances of modifiers might be preferably or obligatorily non-restricting in subsection 2.1.2.5.

Now, Schlenker's cosupposition algorithm generates the same inferences for non-restricting intersective modifiers as the algorithm in (2.29) (modulo the need for two local contexts rather than one). However, Schlenker's algorithm can't apply to non-intersective subsective adjectives like *skillful*, illustrated before in (2.8b) and repeated below in (2.40), which are of higher type than the expression they modify, because the result of modification cannot in this case entail the

¹⁶Here Schlenker makes reference to the original formulation of cosuppositions for co-speech gestures in Schlenker 2018a.

¹⁷I will come back to ellipsis related issues in subsection 7.2.2.

¹⁸In fact, the definition refers to pp' as an “elementary expression”; however, Schlenker himself applies cosuppositions to cases when p and p' within pp' combine compositionally.

modifier (*skillful stuntwoman* doesn't entail *skillful*).

(2.40) Zoe is a skillful stuntwoman.

Zoe is a cellist.

→ Zoe is a stuntwoman.

↗ Zoe is a skillful cellist.

Under my formulation in (2.29), the modifier doesn't have to be of the same semantic type as the expression it modifies or the result of modification. The only requirement is that the expression being modified is of the same type as the result of modification, which is already assured by the very basic type-based definition of a modifier (see subsection 2.1.1). (2.41) exemplifies how this view works for the NP *skillful stuntwoman*, yielding the intuitively correct result that if *skillful* is non-restricting, we get an inference that in the local context of the whole NP, being a stuntwoman entails being a skillful stuntwoman. The exact lexical semantics of *skillful* doesn't matter here, I simply treat *skillful* as a relation connecting properties, individuals, and worlds.

(2.41) a.

NP₂

$\langle e, st \rangle$

$\lambda x \lambda w. \text{skillful}(\text{stuntwoman})(x)(w) \wedge \text{stuntwoman}(x)(w)$

AdjP

NP₁

skillful

stuntwoman

$\langle \langle e, st \rangle, \langle e, st \rangle \rangle$

$\langle e, st \rangle$

$\lambda P \lambda x \lambda w. \text{skillful}(P)(x)(w) \wedge P(x)(w) \quad \lambda x \lambda w. \text{stuntwoman}(x)(w)$

b. non-restricting modifier inference:

$c'' \Rightarrow_{\forall} (\text{NP}_1 \Rightarrow \text{NP}_2)$

$\forall x \forall w [\llbracket c'' \rrbracket(x)(w) \rightarrow (\text{stuntwoman}(x)(w) \rightarrow \text{skillful}(\text{stuntwoman})(x)(w))]$

In fact, one of the cases Schlenker (2018c) himself discusses involves degree modification, which is non-intersective. The case comes from Aristodemo 2017 and involves a maximal-degree modifier ‘completely’ in Italian Sign Language (LIS) and spoken Italian. In LIS it is argued to be an obligatory part of signs such as *FULL* and *BALD*; in spoken Italian it’s a separate co-speech gesture combining with closed-scale adjectives such as ‘full’, ‘bald’, ‘empty’, and ‘straight’. Aristodemo argues that both give rise to projecting conditional inferences (e.g., ‘full’ entails ‘completely full’) and analyzes both as cosuppositions, positing that in LIS ‘completely’ is a “co-sign gesture”.

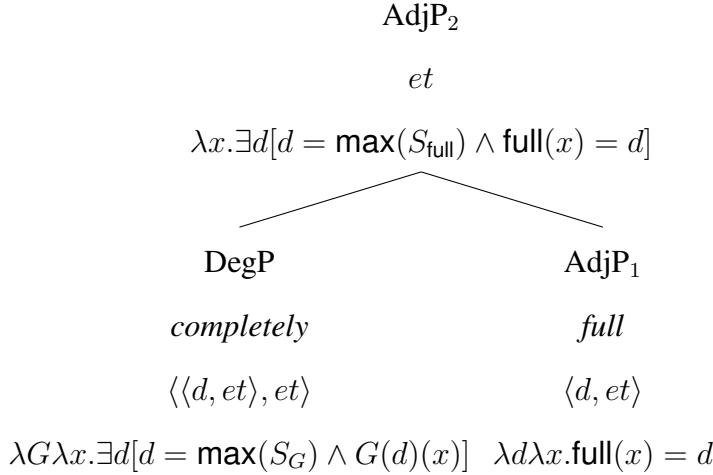
Schlenker’s algorithm in (2.39) cannot assure this result, though, since ‘completely full’ would need to entail ‘completely’ for the algorithm to apply. However, ‘completely’ isn’t of the same type as ‘completely full’, so this is not possible.

Under my view, ‘completely’ in the two cases above would be a preferably non-restricting modifier, whose type doesn’t matter.¹⁹ However, even under my view of cosuppositions as non-restricting modifier inferences, we still need to maintain that the expression being modified is of the same type as the result of modification, i.e., in order for it to apply to degree modifiers, we need to assure that degree modifiers are, in fact, modifiers. Curiously, this is not the case in the classical analysis of degree modifiers in Kennedy & McNally 2005, under which degree modifiers uniformly take a gradable adjective of type $\langle d, et \rangle$ and return a predicate of individuals of type *et*. I provide an example of how degree modifiers compose with adjectives in this system in (2.42) (*d* ranges over degrees; $d = \max(S_{\text{full}})$ reads roughly as ‘*d* is the maximal degree on the scale of fullness’); composition of non-maximal-degree modifiers is similar, it’s the lexical semantics only that will be different.

AdjP_1 and AdjP_2 above have different types, because in Kennedy and McNally’s analysis, degree modifiers systematically do two things: they actually modify the degree of the incoming adjective (i.e., they add some information about the degree variable), and they existentially close off the degree variable. But there is no reason why they should do both. In fact, there is a reason

¹⁹The fact that in LIS it is an obligatory part of the signs does not mean there is no composition in the morphosyntax; bound morphemes exist. Aristodemo’s “co-sign gesture” analysis, in fact, assumes there is composition in the morphosyntax, too.

(2.42) **Composition of degree modifiers in Kennedy & McNally 2005**



why they shouldn't. While it is uncommon to have multiple degree modifiers modifying the same predicate, it is not impossible:

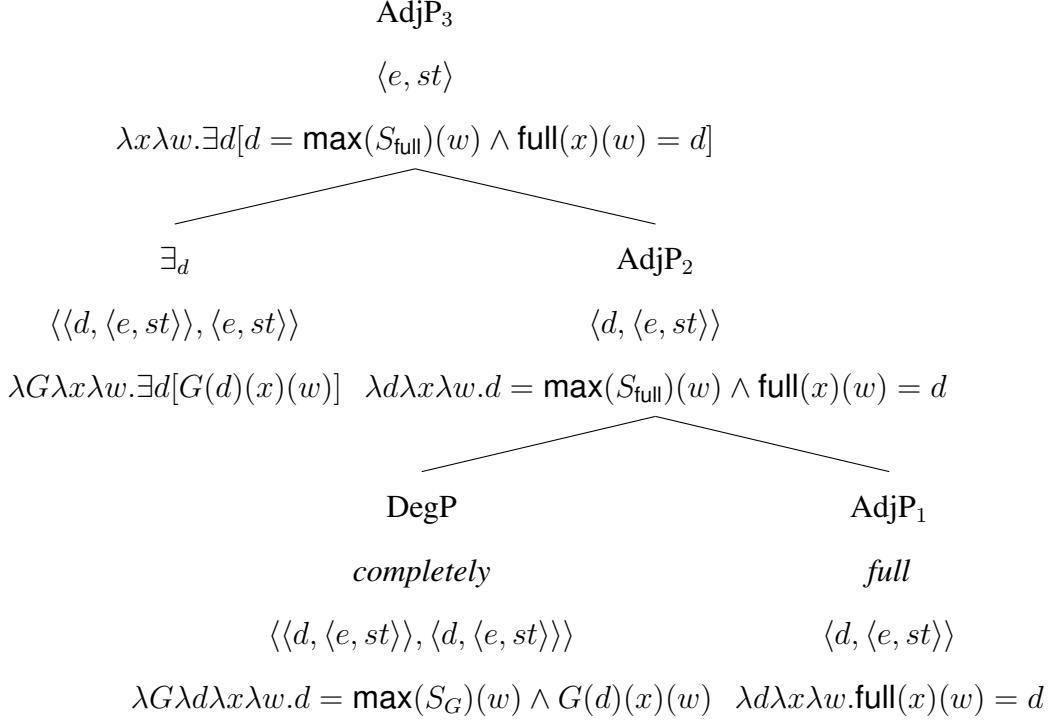
- (2.43) Zoe is an [incredibly, [extremely, [amazingly talented]]] stuntwoman.

The reason why it might be uncommon to have multiple degree modifiers is perhaps that since they modify the predicate on the same scale, multiple modification is likely to be either redundant or contradictory. In (2.43), however, redundancy has an emphatic effect; plus, the modifiers have different evaluative flavors, so they are not entirely redundant.²⁰

No overt coordinator between the degree modifiers is required in (2.43), so it stands to reason that the compositional structure is as indicated by the bracketing. Kennedy and McNally's analysis is incompatible with multiple degree modification, because once we've existentially closed off the degree variable, we can't access it anymore. A natural solution is to divorce degree modification proper from existential closure, which is what I do in (2.44).

²⁰Similarly, multiple degree modifiers can occur in different modalities, in which case they might be perceived as less redundant, although such examples are likely to be ambiguous with respect to their exact compositional structure (see, for instance, the example in footnote 64). Also, Aristodemo doesn't talk about combining gestural and non-gestural degree modifiers in Italian, but I would expect it to be possible.

(2.44) **Proposed composition of degree modifiers**



We thus make modification of adjectives more similar to modification in the nominal and verbal domains, whereby we start out with an expression of a certain type ($\langle e, st \rangle$, $\langle v, st \rangle$, or $\langle d, \langle e, st \rangle \rangle$), and we can modify it by taking it and returning an expression of the same type however many times we need, until we merge a determiner-like element (determiner or existential closure) and turn it into a different type.

Now we can generate non-restricting modifier inferences for degree modifiers if necessary. For example, if *completely* is non-restricting in (2.44), we will have the inference in (2.45), which, at the end of the day, assures that (relative to the local context) if an individual is full to some degree, that degree is maximal.

(2.45) non-restricting modifier inference for *completely full*:

$$\begin{aligned}
c'' \Rightarrow_{\forall} (\text{AdjP}_1 \Rightarrow \text{AdjP}_2) \\
\forall d \forall x \forall w [\llbracket c'' \rrbracket(d)(x)(w) \rightarrow ((\text{full}(x)(w) = d) \rightarrow (d = \max(S_{\text{full}})(w)))]
\end{aligned}$$

Now, let me emphasize that simply switching to the ‘daughter entails parent’ formulation of cosup-

positions from ‘parent entails daughter’ (or ‘sister entails sister’) as a technical tweak to be able to generate cosuppositions in cases of non-intersective modification doesn’t in and of itself address the need for constraining cosuppositions, which I take to be the main issue with Schlenker’s cosuppositions as is. Re-conceptualizing cosuppositions as non-restricting modifier inferences, however, both constrains them in the right way (as we will see in Chapter 4) and allows generating them in cases of non-intersective modification. In other words, fixing the technical bug doesn’t address the bigger issue, but addressing the bigger issue fixes the technical bug along the way.

Is it possible that there exist assertion-dependent inferences that project like presuppositions, but are not non-restricting modifier inferences? For example, can such inferences be triggered lexically rather than pragmatically, or can they be associated with a different compositional configuration? I believe that this in principle possible, and one might even want to call such inferences cosuppositions, but they would be a different class of inferences with properties distinct from those of non-restricting modifier inferences. Crucially, I will argue that none of the cases covered in the rest of this dissertation (gestures, facial expressions, *phi*-features) have so far warranted positing cosuppositions outside of non-restricting modifier inferences—and, in fact, failing to constrain cosuppositions to the modifier configuration in these cases overgenerates.

2.1.2.5 Preferably non-restricting modifiers

Before I move on to discussing supplements, I would like to add a few notes on what can make a given instance of a modifier preferably or obligatorily non-restricting. Instances of modifiers that modify expressions whose extensions are known to be singleton sets will be obligatorily non-restricting, which would account, for example, for the preference for a non-restricting modifier inference arising in (2.17b), as people typically have one mother.

Certain modifiers may prefer to be non-restricting, however, due to their lexical semantics. In (2.22) and (2.23) above I exploited the fact that adjectives that have an evaluative meaning component often exhibit a preference for non-restricting readings by default.²¹ For example, (2.46a)

²¹I thank Alicia Parrish (p.c.) for urging me to think about such adjectives.

is more likely to be uttered in a context in which Stephanie has one dog (or at least a context in which only one of Stephanie's dogs is salient, and the adjective is not used to identify the referent of the definite description), especially as compared to (2.46b). This is not a categorical preference, however, as shown in (2.46c); such adjectives can still in principle be used in a restricting way.

- (2.46) a. Stephanie is bringing her {ginormous, obnoxious, awesome} dog.
- b. Stephanie is bringing her brown dog.
- c. A: Which of her dogs is Stephanie bringing?
 B: The {ginormous, obnoxious, awesome, brown} one.

The reason why (2.46a) tends to default to the non-restricting reading of the adjectives while (2.46b) tends to default to the restricting reading seems to be two-fold. By default, it is hard (although, of course, not impossible) to imagine a context in which the relevance constraint for *brown* would be justified. Modifiers with an evaluative component, however, always have a salient pragmatic function in the discourse, that of expressing the speaker's attitude, which makes them good non-restricting modifiers by default. Apart from that, modifiers whose lexical meaning is mostly or entirely evaluative, such as *awesome* or *obnoxious*, likely don't make very good restricting modifiers because they are highly subjective. Yet, they can still be used in a restricting way, even if it's not an incredibly helpful strategy to identify referents.

Interestingly enough, there are adjectives that can't be used in a restricting way at all, namely, expressives.²² For example, in (2.47b), only literal, non-expressive interpretations of the adjectives are available.

- (2.47) a. Stephanie is bringing her {fucking, bloody, damn} dog.
- b. A: Which of her dogs is Stephanie bringing?

²²Note that Potts (2005) actually refers to positively evaluative adjectives like *lovely* or *brilliant* as "expressives" as well. I use this term more narrowly here. Under this more narrow use, the most important property of true EXPRESIVES is that they can convey attitudes about things other than what they adjoin to syntactically. This is actually noted by Potts for adjectives like *fucking* or *bloody*, but he doesn't make this the definitive property of expressives. So far I have not been able to find expressions in English or other languages that have this property, but are positive by default.

B: #The {fucking, bloody, damn} one.

Could it be that expressive adjectives are just an extreme case of modifiers that are very bad at being restricting due to total lack of any descriptive meaning component? I have reasons to believe that this can't be the end of the story for expressives, one of them being the categorical nature of the judgements in (2.47b). I will briefly come back to expressives in sections 4.4, where I talk about facial expressions, and 5.4, where I talk about the potential performative component of gender features on pronouns, but I will not provide a new analysis of expressives in this dissertation.

Now, let me make it clear that having a preference for being non-restricting, even a categorical one, isn't the same as being NON-RESTRICTIVE. At the beginning of the subsection I defined RESTRICTIVE as having the potential to be restricting, and when I did so, I meant compositional potential. Subsective modifiers have this compositional potential by virtue of having the left conjunct in (2.6).²³ I will reserve the term NON-RESTRICTIVE for content that can't possibly restrict whatever it combines with due to how it combines with it. Thus, whether, for example, expressives are restrictive but aggressively non-restricting or simply non-restrictive depends on what composition strategy one assumes for them, not on whether they can ever be restricting empirically. One example of non-restrictive content are supplements, which is what I will talk about next.

2.2 Supplements

2.2.1 Composition of supplements

The composition strategy used by appositives (both nominal appositives and appositive relative clauses) is different from that of modifiers. Appositives associate with anchors (individuals in the case of adnominal appositives) and contribute some propositional content about those anchors.

Most analyses of appositives assume they adjoin to their anchors in narrow syntax (e.g.,

²³Of course, one could come up with an α that makes the left conjunct in (2.6) trivially true (for example, as a function throwing away its argument and returning “true”), but as a composition strategy, subsective modification as defined here always has the potential to restrict the expression being modified.

Potts 2005; AnderBois et al. 2013; Koev 2013, etc.). For adnominal appositives, whose anchors are individuals, this means that they adjoin to DPs (Determiner Phrases), as opposed to adjectives, which adjoin to NPs (Noun Phrases). Further differences among specific analyses of appositives are immaterial for the purposes of this dissertation, but I will assume the majority view that adnominal appositives adjoin to DPs in narrow syntax.²⁴

For example, in (2.48), the appositive (*who is*) *a stuntwoman*²⁵ composes with the DP *Zoe* denoting the individual Zoe and contributes the proposition of a special type that Zoe is a stuntwoman.

- (2.48) I invited Zoe, (who is) a stuntwoman.

→ *Zoe is a stuntwoman.*

Whichever specific account of appositives one assumes, the general outcome is the same: appositives are not restrictive. The way they compose with their anchors doesn't allow them to restrict anything. For example, Potts (2005) assumes bidimensional semantics whereby the anchor fills the argument slot of the appositive yielding a proposition of a special, *conventional implicature* type. Koev (2013) and AnderBois et al. (2013) assume a dynamic unidimensional setup instead, whereby appositives introduce propositional discourse referents of a special kind; the link between appositives and their anchors is anaphoric.

In (2.49), I illustrate the composition of appositives in Potts 2005 (although, once again, I myself don't commit to this or any other analysis). Under (a simplified version of) Potts's story, (*who is*) *a stuntwoman* composes with the DP *Zoe*, which is of type *e*, passing on the denotation of *Zoe* unchanged and contributing the conventional implicature that *Zoe* is a stuntwoman.

²⁴See McCawley 1998; Schlenker 2013 for an alternative view on the syntax of appositives.

²⁵Here I assume that a nominal appositive, like *a stuntwoman*, has some amount of silent syntactic structure that makes it similar to a full-blown appositive relative clause, like *who is a stuntwoman*; in other words, I assume that nominal appositives are essentially reduced relative clauses (cf. *Zoe, (who is) often called the most talented stuntwoman*). This assumption isn't crucial, however, and whether there are any syntactic differences between the two structures is immaterial for the purposes of this dissertation.

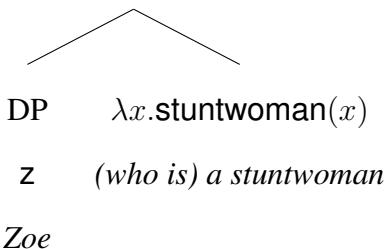
(2.49) **Composition of appositives in Potts 2005 (simplified)**

I invited Zoe, (who is) a stuntwoman.

z (at-issue)

•

$\text{stuntwoman}(z)$ (conventional implicature)



The consequence of this composition that appositives are non-restrictive is borne out empirically, as appositives cannot have restricting interpretations. I will adduce experimental evidence to support this claim in Chapter 3. At this point let me just discuss two purported exceptions to this empirical generalization.

Wang et al. (2005), and subsequently Nouwen (2014), give examples like (2.50), claiming that *one*-appositives are, in fact, restricting (and, thus, restrictive).

(2.50) If a professor, a famous one, publishes a book, they will make a lot of money.

↗ If a professor publishes a book, they are famous.

≈ ‘If a famous professor publishes a book...’

First, let me point out that (2.50) doesn’t really empirically distinguish between a restricting interpretation of the *one*-appositive and a non-projecting one, because the appositive there is adjoined to a low-scoping indefinite: ‘If a famous professor publishes a book...’ is equivalent to ‘If (a professor publishes a book and they are famous)...’. But setting that aside, I agree with the intuition in AnderBois et al. (2013) that *one*-appositives are more like corrections, or second thought elaborations, rather than bona fide restricting modifiers. AnderBois et al. (2013) adduce evidence that *one*-appositives are “appositive in prosody only”; in particular, they don’t require discourse refer-

ents and can be easily targeted by direct responses in the discourse. However, these observations, albeit well-taken, prove that *one*-appositives are “at-issue”, but don’t prove that they are corrections or elaborations rather than bona fide restricting modifiers. This can be proven, however, by contrastive examples, which clearly show that *one*-appositives can’t be used as planned restrictors, as in (2.51c)—as opposed, for example, to adjectives, as in (2.51a), or restrictive relative clauses, as in (2.51b).

- (2.51) a. If a **famous** professor publishes a book, they will make a lot of money, but if an **unknown** professor does so, they will make nothing.
- b. If a professor that’s **famous** publishes a book, they will make a lot of money, but if a professor that’s **unknown** does so, they will make nothing.
- c. #If a professor, a **famous** one, publishes a book, they will make a lot of money, but if a professor, an **unknown** one, does so, they will make nothing.

A second alleged exception is examples like (2.52), in which nominals following other nominals seem to help identify the referents.

- (2.52) (_{PrP} I invited Zoe the **stuntwoman**), (_{PrP} not Zoe the **politician**).

I don’t take these cases to be exceptions either. Following Matushansky 2008, I assume that proper names contain NP-layers denoting properties of being named in a certain way, and in (2.52), *the stuntwoman* and *the politician* in fact adjoin at the NP level and compose as ordinary modifiers (denoting the properties of being the contextually relevant stuntwoman and the contextually relevant politician, respectively).²⁶ This is evidenced, in particular, by the fact that these nominals

²⁶Another possible construal is one where it is technically the same Zoe, but I am talking about the two facets of the same individual, in a way reminiscent of Szabolcsi’s (1980) example (the original example is in Hungarian; the English translation works in the same way):

- (ii) It was [my **friend**]_i that I invited, not [the minister’s **wife**]_{i/j}.

Whatever the details, I would analyze this construal as involving restricting modifiers as well.

don't form their own prosodic phrases, unlike actual appositives.

I will refer to the composition strategy used by appositives—whatever the details thereof—as the SUPPLEMENT strategy. The term SUPPLEMENT is used, in particular, in Potts 2005 and some subsequent work to describe a certain class of expressions, all of which Potts analyzes as conventional implicatures. Apart from appositives, Potts includes parentheticals and sentence-level adverbs, such as *surprisingly* or *unfortunately*, in this class. However, unlike appositives and sentence-level adverbs, full-clause parentheticals, as in (2.53), don't have anchors, and even when they contain a pronoun anaphoric to something in a non-parenthetical clause, the link between that pronoun and its antecedent is weaker than that between, say, an appositive and its anchor (see Esipova 2017 for further details).

- (2.53) I invited Zoe (she's a stuntwoman I told you about yesterday) to the party.

More generally, parentheticals are less restricted from the morphosyntactic perspective and are arguably late merged. Because of that, full-clause parentheticals are not supplements under my use of the term. This is not to say that one cannot have the same projection mechanism for appositives and parentheticals (e.g., that of conventional implicatures). If one wanted to maintain a terminological similarity between appositives and parentheticals, they could introduce a broader definition of supplements than the one assumed in this dissertation and then distinguish between anchored and anchorless supplements within this broader class. I will keep my terminology simple, however.

2.2.2 Projection of supplements

Supplements project, and typically they project very strongly:

- (2.54) a. If you invite Zoe, a stuntwoman, you should show her your muscle car.
→ Zoe is a stuntwoman.
- b. #I don't know if Zoe is a stuntwoman, but if you invite Zoe, a stuntwoman, you should show her your muscle car.

Intended: ‘...if (Zoe is a stuntwoman and you invite her)...’

How exactly one accounts for supplement projection is immaterial for the purposes of this dissertation. For example, in both Potts 2005 and AnderBois et al. 2013 projection of supplements is assured by the propositions they contribute being special in a way that makes them impervious to semantic operators. In Potts 2005, these propositions have a special semantic type, and in AnderBois et al. 2013 they are a special type of discourse referents. Either option works for me.

Schlenker (2013) discusses apparent exceptions to the projection requirement on appositives, such as (2.55).

(2.55) If tomorrow I call the Chair, who in turn calls the Dean, we will be in deep trouble.

↗ If tomorrow I call the Chair, they will call the Dean.

≈ ‘If (tomorrow I call the Chair and they call the Dean)...’

Such interpretations are very constrained, however. Jasinskaja & Poschmann (2018) argue (convincingly) that appositive relative clauses only have such local interpretations when the discourse coherence relation between the matrix clause and the appositive one is a coordinating one. For example, in (2.55) we are dealing with an ordered sequence of events (Jasinskaja and Poschmann subsume this case under the discourse coherence relation of “Narration”). The cases discussed in this dissertation, however, don’t ever involve coordinating coherence relations between the matrix clause and the appositive one and, thus, don’t ever give rise to such local interpretations.

The crucial difference between modifiers and supplements with respect to projection is, thus, that for a modifier not to project means to be restricting, and this interpretation is in principle available to any modifier by default. Whether a given instance of a modifier is interpreted as restricting (and, therefore, non-projecting) or non-restricting (and, therefore, projecting) is determined by a variety of utterance-external factors. In supplements, however, projection is systematically triggered by a certain compositional configuration (e.g., that of an appositive or a sentence-level adverb), and local interpretations are only available under very limited circumstances.

This difference leads me to conclude that we do need two separate projection mechanisms for non-restricting modifiers and supplements. A reductionist story like in Morzycki 2008, whereby non-restricting modifiers contribute Pottsonian (2005) conventional implicatures, fails to capture this contrast in availability of local interpretations. Even if we were to use the same mechanics for assuring lack of interaction with semantic operators for the two types of projecting inferences (i.e., projection proper), we would still need two separate triggering mechanisms. In subsection 2.1.2.3 above, I proposed that triggering of non-restricting modifier inferences is pragmatic in nature. I am not proposing any new mechanism for supplements, but it is clear that triggering of projecting inferences in supplements is much more systematic, much more conventional, and, thus, much more semantic, which is how it is implemented in most existing analyses of supplements.²⁷

2.3 Summary of the chapter

So, let's take a tally. We have (at least) two composition strategies available to non-sublexical content in natural language, modification and supplementation. Adnominal subsective adjectives exemplify modifiers. They adjoin to NPs (type $\langle e, (s)t \rangle$); they are restrictive, but not always restricting; when they are non-restricting, they have to project, and they do so as pragmatically triggered presuppositions. Adnominal appositives exemplify supplements. They adjoin to DPs (type $(s)e$); they are non-restrictive, so they are never restricting; they have to project, and they do so in a way that is more systematic and, thus, more semantic than modifiers. I summarize the modifier vs. supplement distinction in Table 2.1. By the end of this dissertation several more spoken and gestural content types will be added to the example lists in the two columns.

The next two chapters are about gestural content. Co-nominal gestures, as in (2.56), are one of the types of gestural content that I'll be discussing.

(2.56) *Context: We are going on a group tour and want to rent a van. The speaker just learned*

²⁷Even Jasinskaja & Poschmann (2018), despite relying a lot on discourse relations, assume different syntactic construals for globally vs. locally interpreted appositives—which makes sense, given that the crucial distinction for them is between subordination and coordination.

Table 2.1: Modifiers vs. supplements: a preliminary summary.

(Subsective) modifiers	Supplements
Compose with β , yielding α such that α entails β	Compose with β , return a proposition of a special kind about β
Can be restricting or non-restricting	Can never be restricting
Trigger projecting inferences when non-restricting; triggering is pragmatic	Almost always trigger projecting inferences; triggering is conventional
Examples (target content bolded):	Examples (target content bolded):
<ul style="list-style-type: none"> • adnominal adjectives <i>(a blond stuntwoman)</i> • restrictive relative clauses <i>(her dog that's large)</i> • degree modifier adverbs <i>(Mia got surprisingly drunk)</i> • modifiers of verbal projections <i>(Zoe shot at the target with a longbow)</i> 	<ul style="list-style-type: none"> • adnominal appositives <i>(her dog, who is a large animal)</i> • appositives with verbal anchors <i>(Zoe shot at the target, which [= this event of shooting at the target] she did with a longbow; Zoe ran a marathon, which [= running a marathon] is not an easy thing to do)</i> • sentence-level adverbs <i>(Surprisingly, Mia got drunk)</i>

that Stephanie might bring along her only dog.

If Stephanie is bringing her dog_{LARGE}, we should get a bigger van.

Are such co-nominal gestures more like adjectives or more like appositives? Or are they their own thing? In Chapter 3, I report the results of an acceptability judgement experiment comparing co-nominal gestures to adnominal adjectives and appositives. Those results show that, when it comes to available and unavailable interpretations, gestures pattern differently from both adjectives and appositives. In Chapter 4, I show that the two analyses of co-speech gestures currently on the market, Ebert & Ebert 2014; Ebert 2017 and Schlenker 2018a, cannot account for these data on their own, and propose an alternative that can.

Adjectives, appositives, and gestures: experiment

Abstract

In this chapter I report the results of an acceptability judgement experiment looking at how co-nominal gestures can and cannot be interpreted, as compared to spoken adnominal adjectives and spoken adnominal appositives. I show that, as expected, when justified by the context, adjectives can have restricting and non-restricting interpretations equally easily, but when they are non-restricting, the content they contribute has to project. Appositives, in their turn, cannot be restricting, and the content they contribute always has to project (again, as expected). Co-nominal gestures are similar to both adjectives and appositives in that they can easily have non-restricting interpretations, and when they do, they have to project. Restricting interpretations are in principle available to co-nominal gestures, which sets them apart from appositives; however, such interpretations are degraded, as compared to non-restricting ones, which sets them apart from adjectives. The data, R code, and videos pertaining to the experiment can be found at <https://doi.org/10.17605/OSF.IO/5UJYX>.

3.1 Background and goals of the study

As mentioned before in Chapter 1, much of the recent literature on formal semantics and pragmatics of co-speech gestures has been raising the question of how such gestures project (Ebert & Ebert 2014; Ebert 2017; Schlenker 2018a,b; Tieu et al. 2017, 2018; Esipova 2019a). Most of this literature makes an assumption that there is a single uniform answer to this question, i.e., there is a single uniform way in which co-speech gestures project—an assumption that the next chapter will ultimately challenge and reject—and doesn't relate the question of projection of co-speech gestures to the question of their composition.

The existing experimental studies of co-speech gesture projection, thus, yield an incomplete empirical picture. For example, Tieu et al. (2017, 2018) report the results of inference-based tasks (truth-value judgements, picture selection tasks, inferential judgements) looking at co-speech gestures embedded under various semantic operators and comparing them to unambiguously “at-issue” (i.e., non-projecting) modifiers. Examples of both item types are given in (3.1).

- 
- (3.1) a. The boy will not use the stairs_{UP}.
- b. The boy will not use the stairs in this direction_{UP}.

Target inference: If the boy were to use the stairs, he would go up the stairs.

Tieu et al. find that the endorsement rates for the target inference are systematically higher for examples like (3.1a) than for examples like (3.1b), which leads them to the conclusion that co-speech gestures make projecting contributions by default.

They, furthermore, conclude that the projection patterns of inferences contributed by co-speech gestures resemble those of lexical presuppositions more so than those of spoken appositives. I will not rehash the details of this conclusion here but will come back to it in subsection 4.2.1. I will note, however, that Tieu et al. don't adduce comparable experimental data for either lexical presuppositions or spoken appositives; their conclusions are based on introspective data for these

types of content.

While Tieu et al.'s studies contribute important baseline results (in particular, that co-speech gestures contribute projecting inferences by default), they are limited in several respects. For one thing, they rely on inference-based tasks. One can, however, investigate projection patterns via acceptability tasks by fixing the interpretation of a given sentence and asking the participants how acceptable this sentence is. I believe the acceptability-based approach is methodologically superior to the inference-based approach. First, it yields explicit data on whether the test items are acceptable in the first place; it is unclear how participants behave in an inference-based task when the item is not acceptable or is marginally acceptable. Second, by giving every potential interpretation "its best shot", the acceptability-based approach mitigates the concern that the participants don't get a certain interpretation because it is hard for them to construct a context that would support it. Third, and most importantly for our purposes, it allows distinguishing among various interpretations when the target inference does not project.

Regarding this latter point, the endorsement rates of the target inferences in Tieu et al.'s data were not always very high in examples with co-speech gestures. That led Tieu et al. to the same conclusion as in Schlenker 2018a that co-speech gestures are weak presupposition triggers, which allow for local accommodation relatively easily under pressure. For example, they say, local accommodation might be preferred in some cases because the inferences contributed by the gestures might sometimes be too pragmatically odd to accommodate globally.

However, there are at least two conceivable local interpretations for co-speech gestures. One is that of a restricting modifier and the other is that of a locally interpreted non-restricting modifier or a locally interpreted supplement (I will label this second hypothetical interpretation as 'Non-projecting non-restricting'), as illustrated in (3.2).²⁸

- (3.2) If Stephanie's bringing her dog_{LARGE}, we should get a bigger van.

Projecting interpretation: → Stephanie's only dog is large.

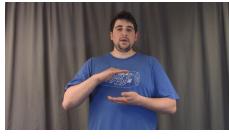
²⁸Chapter 6 establishes that the same possibilities exist for co-verbal gestures, which are the ones that Tieu et al.'s focus on.

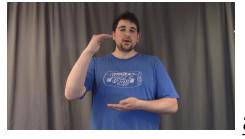
Restricting interpretation: \approx ‘If Stephanie’s bringing her large dog, but not her small one...’

Non-projecting non-restricting interpretation: \approx ‘If (Stephanie’s only dog is large and she is bringing her large dog)...’

Zlogar & Davidson 2018 do use an acceptability task, but the questions they ask are very different from the questions asked in Tieu et al. 2017, 2018 or in this dissertation, so I will not review this work here.

In Esipova 2019a, I mitigated some of the issues with Tieu et al.’s studies listed above, as well as one more issue, which is that Tieu et al. ignore the role of focus (more specifically, that *this* is systematically focused in their controls like the one in (3.1b), which is likely what forces the truth-conditionally non-vacuous interpretation, however, there is no reason why the gesture should be focused in examples like (3.1a)). In particular, I looked at the relative acceptability of examples like in (3.3).

(3.3) John might order  a beer_S...

a. ...or  a beer_L.

b. ...or  a cocktail_S.

In examples like (3.3a) the contributions of the gestures can’t possibly project because the gestures are the only locus of contrast between the two juxtaposed alternatives. In examples like (3.3b), however, the spoken expressions (but not the gestures) are contrastive, so the contribution of the gestures is allowed to project (it’s unclear, of course, if it actually does). I found that, overall, examples of the former type are less acceptable than examples of the latter type, however, speakers

vary in an internally consistent way in their rating patterns. These results provide evidence both against the view that the contributions of co-speech gestures don't project by default (more convincingly so than Tieu et al.'s results, due to the focus-related considerations) and any view that posits a fixed bias against non-projecting interpretations of co-speech gestures (due to the large amount of inter-speaker variation).

That said, the study in Esipova 2019a still doesn't eliminate all of the concerns above. First, it does not use the full power of the acceptability-based approach to investigating projection, as it does not provide contexts that would support the target interpretations. Second, and most importantly, it still doesn't allow us to get at the potential contrast between the two types of local interpretations (restricting and non-projecting non-restricting), because it only looks at gestures co-occurring with indefinites, and the target distinction is impossible to see for indefinites, as mentioned previously in subsection 2.2.1 and illustrated once again in (3.4).

- (3.4) ‘It might be that John will order a small beer, but not a large one.’ = ‘It might be that (John will order a beer and that beer will be small, but not large).’

The main goal of the present study is, thus, to distinguish between these two possible local interpretations while using the full power of the acceptability-based approach to investigating projection. More specifically, this study aims to answer the following questions:

Q1: Can co-speech gestures have restricting interpretations?

Q2: When a co-speech gesture is not restricting, can it be interpreted locally under semantic operators?

Answering these questions will allow us to see how co-speech gestures fit, if at all, into the modifier vs. supplement distinction summarized in Table 2.1 at the end of the previous chapter. To have clear baselines for comparison, I also collect the relevant data for spoken adjectives and appositives.

3.2 Methods

Participants were recruited on Amazon Mechanical Turk and paid \$1 each for completing the task. Those who failed the attention check or reported being non-native speakers of English were excluded. The final number of participants was 122 (33 female, 89 male).

Participants were asked to read context paragraphs, watch videos of sentences uttered in those contexts, and assess these sentences by dragging a slider to the desired position on a pseudo-continuous scale from ‘Totally unnatural’ to ‘Totally natural’ (mapped to 0–100). See Appendix C for the instructions given to the participants. The layout of a typical trial is shown in Figure 3.1 (not to scale).

Context: We are going on a group tour. Anna and Maria are responsible for renting a van. Maria just told Anna that **Stephanie, who has two pets, a small cat and a large dog, is** planning to bring along one of her pets. **Anna, who has seen both Stephanie's pets before, says:**

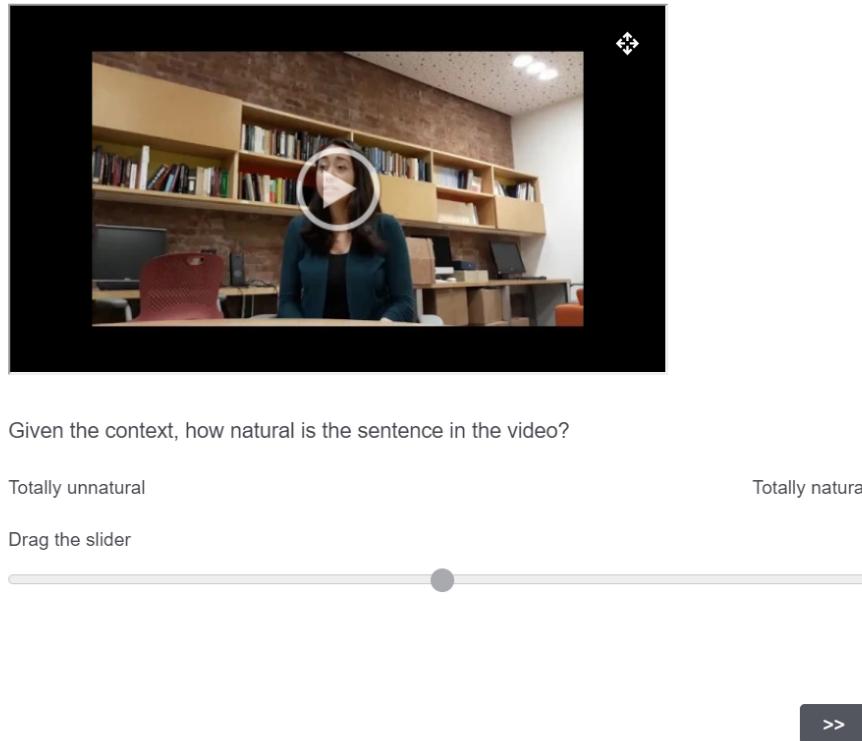


Figure 3.1: Layout of a typical trial.

The items differed along two fully-crossed factors: Content Type (Adjective, Appositive,

Gesture) and Interpretation (Projecting non-restricting, Restricting, Non-projecting non-restricting), resulting in 9 conditions. The target interpretation was enforced within each video by setting up the QUD and explicitly contrasting two alternatives;²⁹ the written contexts were meant to maximally support this interpretation. Four complete test paradigms (for four different scenario types) were constructed; see Appendix D for the list of all the test items used in the experiment. The design is summarized in Table 3.1.

Table 3.1: Design of the experiment.

		Interpretation		
		<i>Projecting non-restricting</i>	<i>Restricting</i>	<i>Non-projecting non-restricting</i>
Content Type	<i>Adjective</i>	4	4	4
	<i>Appositive</i>	4	4	4
	<i>Gesture</i>	4	4	4

A complete sample paradigm is given in (3.5). All the written contexts (in italics) in this paradigm started in the same way, setting up the van-renting scenario. They then continued in three different ways, setting up the subscenario supporting one of the three target interpretations. Three videos were then recorded for each of the three written contexts, with adjectives, appositives, and gestures, respectively, as the expressions of interest.

(3.5) *Context: We are going on a group tour. Anna and Maria are responsible for renting a van. Maria just told Anna that...*

- a. *Projecting non-restricting: ...Stephanie, who has two pets, a small cat and a large dog, is planning to bring along one of her pets. Anna, who has seen both Stephanie's pets before, says:*

Do you know which one of Stephanie's pets is coming with us? 'Cause if she's bringing...

²⁹The alternatives were made explicit after a pilot study revealed that some participants might be “rescuing” some of the items with appositives by interpreting the latter as describing a kind, not an instance thereof. For example, applying this strategy to *If Stephanie's bringing her dog, a large animal, we should get a bigger van* yields an inference that dogs are in general large.

- (i) her small **cat**
- (ii) her **cat**, a **small** animal



- (iii)  her **cat**_SMALL

..., we'll be fine, but if she's bringing...

- (i) her large **dog**
- (ii) her **dog**, a **large** animal



- (iii)  her **dog**_LA_GR_AE

..., we should get a bigger van.

- b. *Restricting: ...Stephanie, who has two dogs, a small Pug and a large Great Dane, is planning to bring along one of her dogs. Anna, who has seen both Stephanie's dogs before, says:*

Do you know which one of Stephanie's dogs is coming with us? 'Cause if she's bringing...

- (i) her **small** dog
- (ii) her dog, a **small** animal
- (iii) her **dog**_SMALL

..., we'll be fine, but if she's bringing...

- (i) her **large** dog
- (ii) her dog, a **large** animal
- (iii) her **dog**_LA_GR_AE

..., we should get a bigger van.

- c. *Non-projecting non-restricting interpretation: ...Stephanie is planning to bring along her dog. Anna knows that Stephanie only has one dog, but has never seen it. She says: Do you know how big Stephanie's dog is? 'Cause if she's bringing...*

- (i) her **small** dog

- (ii) her dog, a **small** animal
 - (iii) her dog_{SMALL}
- ..., we'll be fine, but if she's bringing...
- (i) her **large** dog
 - (ii) her dog, a **large** animal
 - (iii) her dog_{LARGE}
- ..., we should get a bigger van.

(3.5a) corresponds to the ‘Projecting non-restricting’ column of Table 3.1. Here the expressions of interest are not restricting the expressions they combine with, as the extensions of those are singleton sets (Stephanie only has one cat and one dog), hence the ‘non-restricting’ part in the interpretation label. Furthermore, the QUD is about which of her two pets Stephanie is bringing, not their size, so the contributions of the expressions of interest are allowed to project, hence the ‘projecting’ part of the interpretation label.

(3.5b) corresponds to the ‘Restricting’ column of Table 3.1. Here the expressions of interest are restricting the expressions they combine with, as Stephanie has two dogs and these expressions are meant to pick out one of them based on size.

(3.5c) corresponds to the ‘Non-projecting non-restricting’ column of Table 3.1. Here the expressions of interest are not restricting the expressions they combine with, as their extensions are singleton sets (Stephanie only has one dog), hence the ‘non-restricting’ part of the interpretation label. Furthermore, the QUD is about the size of Stephanie’s dog, and it’s clear from the question Anna asks that she doesn’t know what size it is, so the contributions of the expressions of interest have to be interpreted locally under *if*, hence the ‘non-projecting’ part.

The design was within subjects. Each participant saw one randomly selected item per cell (thus, 9 test items) and 2 additional items (thus, 11 items in total), presented in random order. One of the additional items was an attention check where the participant was instructed to drag the slider all the way to the left or to the right. The other additional item had a mismatch between the

written context and the video sentence; it was always a context corresponding to the non-projecting non-restricting interpretation and a sentence with restricting adjectives (randomly selected from the four scenario types). This item was added to assess whether participants were paying attention to the written contexts and as an informal baseline for degradedness.

3.3 Results

A spreadsheet with the raw data (including from the excluded participants) can be found at <https://osf.io/s43n8/>. A spreadsheet with the cleaned up data can be found at <https://osf.io/qymr2/>. All statistical tests and plots were done using R (2019); the code for the regression models and charts can be found at <https://osf.io/v39q2/>. The data were subsetted based on Content Type, and a mixed effects linear regression model was run for each type of content with Interpretation as a fixed effect, and Participant and Scenario Type as random effects. Once the significant effect of Interpretation was established, pairwise comparisons between different interpretations were performed via similar models.

The results are visualized in Figure 3.2, and the statistics are summarized in Table 3.2. The mean % acceptability for mismatch items was 39, which is not significantly different from that of non-projecting non-restricting items with adjectives ($\text{Beta} = -.041$, $t = -.858$, $p = .393$), but no meaningful comparisons can be made with the other content types.

As expected, projecting non-restricting and restricting interpretations of adjectives are both equally and highly acceptable. Non-projecting non-restricting interpretations of adjectives are much less acceptable—again, as expected, since, as we have observed before, based on introspective data, non-restricting modifier inferences have to project.

Only projecting non-restricting interpretations are highly acceptable for appositives. Both restricting and non-projecting and non-restricting interpretations are much less acceptable, and there is no significant difference between the two. This is again expected, as appositives don't compose in a way that would allow them to have restricting interpretations, and they have to project.

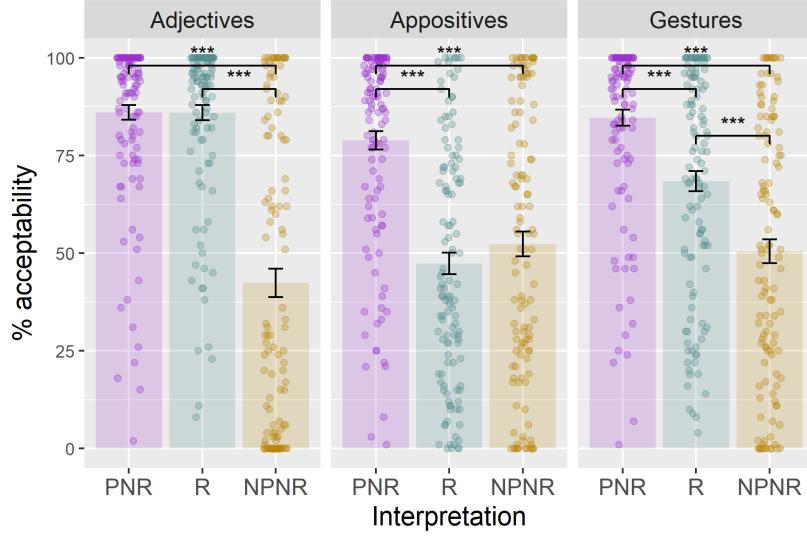


Figure 3.2: Bar chart showing % acceptability of different interpretations for each content type. Error bars show standard error. Dots represent individual responses (jitter was added for better visualization).

Projecting non-restricting interpretations of co-nominal gestures are highly acceptable, just like for adjectives or appositives. Non-projecting non-restricting interpretations of co-nominal gestures are much less acceptable, again, just like for adjectives and appositives. Crucially, however, restricting interpretations of co-nominal gestures are marginal; they are significantly more acceptable than non-projecting non-restricting ones (in contrast to appositives) but significantly less acceptable than projecting non-restricting ones (in contrast to adjectives).

Going back to the two questions posed in section 3.1, we can now provide the following answers:

Q1: Can co-speech gestures have restricting interpretations?

A1: Yes, marginally.

Q2: When a co-speech gesture is not restricting, can it be interpreted locally under semantic operators?

A1: No.

Table 3.2: Statistics for acceptability of different interpretations for each content type.

Content	Mean % acceptability				Comparisons	
	PNR	R	NPNR	PNR/R	PNR/NPNR	R/NPNR
Adjectives	86.0	85.9	42.4	NS	***	***
				Beta = .001	Beta = .565	Beta = .568
				t = .03	t = 10.71	t = 10.842
				p = .976	p < 2e-16	p < 2e-16
Appositives	78.9	47.4	52.4	***	***	NS
				Beta = .49	Beta = .398	Beta = -.077
				t = 10.35	t = 8.057	t = -1.628
				p < 2e-16	p = 6.28e-13	p = .106
Gestures	84.6	68.4	50.5	***	***	***
				Beta = .301	Beta = .513	Beta = .277
				t = 6.298	t = 10.95	t = 5.459
				p = 5.43e-09	p < 2e-16	p = 2.59e-07

3.4 Discussion

3.4.1 Desiderata for a theory of co-speech gestures

A comprehensive analysis of co-nominal gestures should, thus, satisfy the following conditions:

1. It should allow for restricting interpretations of co-speech gestures.
2. It should explain why restricting interpretations of co-speech gestures are still degraded as compared to non-restricting ones.
3. It should have a potential to capture the gradient and variable nature of the judgements on restricting interpretations of co-speech gestures. The design of the experiment discussed in this chapter doesn't allow us to assess inter-speaker variability in a meaningful way, but in Esipova 2019a I show that speakers vary in an internally consistent way in how acceptable they find non-projecting interpretations of co-nominal gestures.
4. It should not allow for non-projecting non-restricting interpretations of co-speech gestures.

In the next chapter I show that the theories of co-speech gestures currently on the market

(Ebert & Ebert 2014; Ebert 2017 and Schlenker 2018a) can't fulfill all these desiderata on their own. I then propose a view of gestures that can and is furthermore conceptually appealing.

3.4.2 Projecting non-restricting interpretations and *Minimize Restrictors!*

When presenting the results for adjectives, I said that the fact that projecting non-restricting and restricting interpretations of adjectives were equally and highly acceptable was expected. In principle, we might have expected non-restricting adjectives to be somewhat degraded, due to the *Minimize Restrictors!* principle (mentioned before in subsection 2.1.2.1), however, their presence was clearly justified in each case (e.g., in (3.5a) the size of the pets matters for van selection purposes).

This observation might shed the light on the experimental results in Zlogar & Davidson 2018, where sentences with co-speech gestures were degraded across the board. It is reasonable to assume that something like *Minimize Restrictors!* applies not only to modifiers (non-restricting modifiers need to satisfy the discourse relevance condition), but also to supplements (they, too, need to be linked to the matrix speech act via some sort of discourse relation). If, as we see in Figure 3.2, restricting co-speech gestures are degraded and the presence of truth-conditionally vacuous gestures (be they non-restricting modifiers or supplements) needs to be justified, it stands to reason that without such a justification co-speech gestures might be perceived as not entirely natural.

Crucially, in my results, examples with gestures were not degraded across the board; projecting non-restricting gestures were almost as acceptable as projecting non-restricting adjectives (84.6 vs. 86, respectively) and more acceptable than projecting non-restricting appositives (84.6 vs. 78.9, respectively). While there might be some methodological reasons for the contrast between my results and the ones in Zlogar & Davidson 2018,³⁰ the need to justify truth-conditionally vacuous material might have played a role, as well.

³⁰In particular, I took measures to mitigate the potential prescriptive bias against examples with gestures. My instructions contained examples with gestures, thus, “normalizing” them. Also, the speaker in my video stimuli was in a less formal setting than the speaker in Zlogar and Davidson’s stimuli and was pretending to be talking to someone off screen rather than looking straight at the camera (in this respect my earlier study in Esipova 2019a patterns with Zlogar & Davidson 2018). See also Duncan 2019 for a relevant methodological discussion.

3.4.3 Note on “high” acceptability of unacceptable interpretations

Something that might seem suspicious in Figure 3.2 is that the ratings of the interpretations I deemed unacceptable are still pretty high in absolute terms, and also, plenty of individual judgments in these columns are high. I have several thoughts on the matter.

First, let me emphasize that the individual responses in Figure 3.2 are just that; they are not mean ratings for a given condition for a given participant, since each participant only saw one item per condition. Because of this property of the design, it is impossible to talk about individual variation (grammatical or behavioral) in a meaningful way (cf. the design and analysis of variation in Esipova 2019a).

Second, it is worth pointing out that all sentences in all items were syntactically well-formed; the expected degradedness was always semantic and/or pragmatic. Furthermore, the items were complex and involved (unlike the examples in the instructions, for instance). If a given participant was only assessing syntactic well-formedness or simply wasn’t paying attention to all the aspects of the items, they could very well judge all or most of the items fairly high. The fact that the context mismatch item was judged fairly high (mean rating of 39) suggests that indeed some people were not paying attention to all the aspects of the items.³¹

There can also be “rescuing” strategies for some of the items that I deemed unacceptable. One such obvious strategy would be treating some of the purportedly unacceptable local interpretations as instances of metalinguistic focus, which I discuss separately in subsection 7.2.1. Note, however, that this cannot explain the contrast between the restricting and the non-projecting non-restricting interpretations of gestures, as if this strategy is available, it should be equally available in both cases.

Another potential “rescuing” strategy could have been used specifically for non-projecting non-restricting items with adjectives and gestures. If a participant ignored the written context (in which it was specified that Stephanie only has one dog), the question *Do you know how*

³¹Let me point out, however, that the written context wasn’t crucial for enforcing the target interpretation; the interpretation was enforced by setting up the QUD and contrasting two alternatives explicitly within the video sentences.

big Stephanie's dog is? could have been interpreted similarly to *Do you know how big the dog Stephanie's going to bring is?*, which is compatible with Stephanie having two dogs and the adjectives/gestures in the conditionals having restricting interpretations.

Finally, note that judgements for items with appositives and gestures are more scattered overall. This might be because the participants have less experience with these two types of content and are less certain in their judgements.

Gestures

Abstract

In this chapter I discuss how projection of gestures should be analyzed based on the data from the previous chapter and on further empirical observations. The main overarching point of this chapter is that in order to understand how gestural expressions project, we first need to understand how they integrate into utterances at various levels of representation; simply looking at how a given expression is linearized or whether it is syntactically optional is insufficient.

First, I focus on CO-SPEECH GESTURES, i.e., content-bearing gestures that co-occur and associate with spoken expressions. I show that a theory that treats co-speech gestures as supplements across the board undergenerates by failing to derive restricting interpretations of co-nominal gestures. The theory that treats co-speech gestures as triggering cosuppositions across the board, on the other hand, overgenerates unattested interpretations. I propose an alternative whereby how (compositionally integrated) gestures project is determined by how they compose. In particular, co-nominal gestures can compose as modifiers or as supplements. When they compose as supplements, they project as supplements. When they compose as modifiers, they only project if they are non-restricting, and they do so like other non-restricting modifier inferences. However, there is a further pragmatic preference for co-speech gestures to be truth-conditionally vacuous, which in the case of modifier gestures results in a (gradient and potentially variable) preference for non-restricting interpretations.

I subsequently discuss PRO-SPEECH GESTURES, i.e., content-bearing gestures with their own time slot that are compositionally integrated into otherwise spoken utterances. I show that articulatory and prosodic grouping considerations constrain the linearization possibilities for such gestures, which, in turn, constrains the syntactic construals available to them and, as a result, their interpretation and possible projection patterns. I furthermore adduce some data from French and Russian showing that prosodically independent gestures integrating into otherwise spoken utterances abide by the same language-specific rules for linearization, prosodic grouping, and displacement as spoken content. I conclude that, once again, the projection behavior of gestures with their own time slot is not determined by their linearization, and, in particular, the distinction between “pro-speech” and “post-speech” gestures introduced in Schlenker 2018a is not a linguistically meaningful one.

Next, I show that, based on some preliminary data on co-speech and co-gesture facial expressions encoding surprisal, projection of facial expressions, too, is determined by how they compose, not by how they are linearized.

Finally, I discuss the role of convention in how various pieces of content project and conclude that a typology of projective content that relies on level of conventionalization is neither practical nor necessary.

I conclude with some methodological remarks that could be useful for future research on secondary modality content.

4.1 Supplemental analysis of co-speech gestures

In what, to my knowledge, is the first attempt to account for the projection behavior of co-speech gestures, Ebert & Ebert (2014) claim that such gestures are supplements, akin to appositives. They propose to analyze such gestures within the dynamic system from Koev 2013 and AnderBois et al. 2013, although, once again, for our purposes the specific analysis doesn’t matter.

More recently, Ebert (2017) proposed a more refined analysis. She claims that co-speech

gestures are usually supplements, but there also exist NP-level co-speech gestures with “exemplification” semantics. Under this view, her dog^{LARGE}, for example, typically has the same semantics as *her dog, who is large*, in which case the gesture adjoins to the DP *her dog*. But—if my understanding of her claims is correct—the gesture *LARGE* can also sometimes be interpreted as adjoining to the NP *dog* and exemplifying a typical entity in its denotation.

If co-speech gestures are supplements, it is unsurprising that they allow for projecting, but not for non-projecting non-restricting interpretations; the explanation would be the same as for appositives. Yet, neither the supplement nor the exemplification strategy can yield restricting interpretations of co-speech gestures, which we have seen to be marginally available. Thus, while it is possible that co-speech gestures are sometimes supplements (and maybe sometimes exemplifiers), this doesn’t give us the full range of available interpretations. In other words, the supplemental analysis of co-speech gestures undergenerates.

4.2 Cosuppositional analysis of co-speech gestures

4.2.1 The gist of the cosuppositional analysis

Schlenker (2018a) argues against the supplemental analysis and proposes instead that co-speech gestures trigger cosuppositions across the board.

One of Schlenker’s original arguments against the supplemental analysis is that co-speech gestures don’t seem to be subject to the same anaphoric constraints as adnominal appositives. More specifically, as shown in (4.1), while adnominal appositives need a discourse referent to serve as an anchor for them, similarly to ordinary pronouns requiring discourse referents as antecedents for cross-sentential anaphora, co-speech gestures don’t seem to require that.³² I’d like to note that in my experience the judgements on anchorless co-speech gestures are gradient and variable, and most people do find them somewhat degraded (which is what I report in (4.1)), but there is still a contrast between co-speech gestures and appositives to be explained. I will come back to the

³²A *dog* in (4.1) is meant to be a low scope indefinite.

potential reasons why (4.1a) and (4.1b) are somewhat degraded for many people in the next section.

- (4.1) a. %?Stephanie didn't bring a dog_{LARGE}.
- b. %?Stephanie brought no dog_{LARGE}.
- c. *Stephanie didn't bring a dog, {who was large, a large animal}.
- d. *Stephanie brought no dog, {who was large, a large animal}.
- e. Stephanie didn't bring a dog. *It was large.
- f. Stephanie brought no dog. *It was large.

Tieu et al. (2017) provide further experimental evidence that co-speech gestures don't always project like appositives either. In particular, they look at examples like (4.2a) (among other things).



- (4.2) a. None of these ten guys helped_{LIFT} his son.
- b. None of these ten guys helped his son, which [= helping one's son] involved doing this_{LIFT}.
- c. *None of these ten guys helped his son, which [= the event of helping x 's son] happened like this_{LIFT}.

The claim they put forward isn't that such examples should be ungrammatical under the supplement construal of the gesture, since spoken appositives in the verbal domain can have event predicates as anchors, as shown in (4.2b) (cf. (4.2c), where the anchor is meant to be an event, but since there is no available event discourse referent, the sentence is unacceptable).³³ Instead, they propose that, under the supplement construal of the gesture, sentences like (4.2a) should give rise to universal inferences (for all of these ten guys, if he was to help his son, that would have involved lifting) rather than existential ones (for some of these ten guys, if he was to help his son, that would have involved lifting), because, they claim, that's what sentences like (4.2b) do. They

³³There is further discussion in Tieu et al. 2017 involving indicative vs. subjunctive appositives; however, since it's unclear to me what it would mean for a gesture to bear subjunctive mood, I will not rehash it here.

find, however, some evidence in favor of the existential projection pattern for sentences like (4.2a), which is what is expected under some analyses of presupposition projection (e.g., Beaver 2001).

Now, as mentioned before in section 3.1, Tieu et al. (2017) don't provide comparable data either for spoken supplements or lexical presuppositions. Furthermore, the choice of the appositive for comparison in (4.2b) is not motivated in any obvious way. The fact that the appositive in (4.2b) bears the same tense as the matrix clause, to my mind, favors a reading in which the appositive describes homogeneous rules for all the ten guys within a specific situation. However, appositives with event predicate anchors can bear any tense or quantificational force, and it's unclear to me why gestures like *LIFT*, which contain no overt tense morphemes nor quantificational adverbs, should be compared to appositives like in (4.2b) rather than, say, generic appositives like in (4.3a), which, as far as I can tell, don't necessarily give rise to universal inferences regarding the ten guys in question (as some of them can be an exception to the generic rule), or even appositives with existential force like in (4.3b) or (4.3c).³⁴

- (4.3)
- a. None of these ten guys helped his son, which [= helping one's son] (<{usually, typically}) involves doing this_{LIFT}.
 - b. None of these ten guys helped his son, which [= helping one's son] sometimes involves doing this_{LIFT}.
 - c. None of these ten guys helped his son, which [= helping one's son] could(, for example,) involve doing this_{LIFT}.³⁵

That aside, however, even if we assume a robust contrast in this respect (supplement construals always give rise to universal inferences; presupposition construals can give rise to existential inferences), the results in Tieu et al. 2017 cannot be used as an argument that gestures in examples like (4.2a) are never construed of as supplements, only that they aren't always construed of as such, since just because participants allow the construal that gives rise to an existential (weaker)

³⁴See section 6.2 for further discussion of event predicate anchor construals for appositives in the verbal domain.

³⁵A comparison with this appositive would be in line with the claim put forward in Ebert 2017 that co-speech gestures can have exemplification uses.

inference, we cannot conclude that they don't also allow the construal that gives rise to a universal (stronger) inference.

The same can be said about Schlenker's observations about anahoric constraints illustrated in (4.1). In this respect, Schlenker and Tieu et al.'s data point in the same direction as my experimental data from the previous chapter: even if co-speech gestures are sometimes interpreted as supplements, the supplement strategy doesn't yield the entire range of possible interpretations for co-speech gestures.

However, as I will show next, co-speech gestures can't possibly generate cosuppositions across the board either. While applying cosuppositions to NP-adjoining gestures yields good results (the same results we would get if we treated NP-level gestures as preferably non-restricting modifiers), trying to apply them to DP-level gestures generates unattested interpretations.

The formulation of cosuppositions in Schlenker 2018a is slightly different from the one in Schlenker 2018c as cited in (2.39), but it is similar. As formulated in Schlenker 2018a, a gestural cosupposition of the configuration $[\underline{S}]^G$ has the form $c' \Rightarrow_{\forall} (S \Rightarrow G)$, where S is the spoken expression the gesture co-occurs and associates with, G is the gesture, and c' is the local (non-propositional) context. Schlenker also makes a standard assumption that presuppositions can in principle be locally accommodated under some (possibly minor) pressure, depending on the strength of the trigger. He maintains that gestural cosuppositions are relatively weak and can thus be locally accommodated fairly easily. He further specifies that when a gestural cosupposition is locally accommodated, the $S \Rightarrow G$ expression is conjoined to S : $S \& (S \Rightarrow G)$, which is equivalent to $S \& G$, where $\&$ is generalized conjunction. Local accommodation of presuppositions is typically taken to incur some cost,³⁶ whose amount can vary across triggers (the weak–strong scale) and—potentially—across speakers.

Now, the result of local accommodation of a gestural cosupposition is equivalent to treating the gesture as simply conjoining with the spoken expression. Schlenker's views on gestures

³⁶The psycholinguistic nature of this cost isn't usually discussed in the relevant literature; I am using this term simply to refer to whatever results in lower acceptability. However, see Chemla & Bott 2013 for experimental data on response times as a measure of processing cost incurred by local accommodation.

(Philippe Schlenker, p.c.) is broadly compatible with a variation on the story above whereby co-speech gestures start out as composing with the spoken expressions they associate with conjunctively and by default trigger cosuppositions. However, if for some reason the cosupposition is blocked or dispreferred (for example, the context is incompatible with the cosupposition, or the cosupposition is not supported by the context, but is hard to accommodate globally), cosuppositions aren't triggered. In other words, as things stand, local accommodation and non-triggering yield the same results in Schlenker's system, which is conjunctive composition of the spoken expression and the gesture. Which version of Schlenker's analysis one assumes will not matter for the problem I will raise in subsection 4.2.3, but the second version is closer in spirit to the view of cosuppositions as non-restricting modifier inferences advocated for in subsection 2.1.2.4.³⁷

As mentioned before in subsection 2.1.2.4, for Schlenker's cosuppositions to apply, it is crucial that S , G , and c' are all of the same semantic type, otherwise it is not possible to apply generalized entailment, generalized material implication, or generalized conjunction. Thus, the type of the gesture itself and the expression to which the gesture adjoins matters. In particular, for co-nominal gestures like in *her dog*^{LARGE} it will matter whether the gesture adjoins to the NP (*dog*), of type $\langle e, (s)t \rangle$, or the DP (*her dog*), of type $(s)e$ or $\langle \langle e, (s)t \rangle, (s)t \rangle$.

Note that the exact temporal alignment of a gesture is unlikely to be a reliable indicator of its adjunction level, especially for co-nominal gestures.³⁸ For example, the preparatory phase of a co-speech gesture often starts way before the spoken expression it “associates with”.³⁹ In many cases it is also hard to tell what the type of the gesture itself is, given how morphosyntactically poor gestures typically are. In particular, for size gestures like *LARGE* it is entirely plausible that they

³⁷The two are still very different, as Schlenker's view doesn't associate cosuppositions with the modifier composition strategy, but with simple conjunction, compositional or not, which is both too constrained, as we have seen in subsection 2.1.2.4, and not constrained enough, as we will see in subsection 4.2.3.

³⁸Of course, one can try to check things for gestures co-occurring with different parts of DPs with very long determiners, but I believe such investigations would need to be done experimentally, and I don't have any prior expectations about how sensitive non-signers are to precise gesture alignment. More importantly, the question of whether the exact temporal alignment of a gesture can predict its adjunction level is orthogonal to the issue I am raising for the cosuppositional analysis—or any analysis that implicitly or explicitly assumes that co-speech gestures compose in the same way across the board.

³⁹This can be seen even in the video stimuli for the experiment in Chapter 3 and the one in Esipova 2019a; in both cases the gesture models were instructed to produce the gestures in the way that seemed natural to them.

are construed of as representing properties (akin to the adjective *large*) or as representing objects (akin to the nominal *a large object*). Thus, I will assume that whenever a gesture roughly aligns with and associates with a nominal, the gesture might be in principle adjoining either to the entire DP or to an NP within it.

With this in mind, let's see what results the cosupposition mechanism above yields for co-nominal gestures adjoining to NPs and DPs, respectively.

4.2.2 NP-level gestures under the cosuppositional analysis

Schlenker's cosuppositional analysis yields good predictions for property-denoting co-speech gestures adjoining to NPs; these results are essentially equivalent to treating gestures as preferably non-restricting modifiers. For example, if in *Stephanie brings her dog_{LARGE}*, the gesture *LARGE* adjoins to the NP *dog*, we get the results in (4.4) (as before, C is the global context set), modulo some assumptions.

(4.4) *Stephanie brings her [[_{NP} *dog*]_{LARGE}].*

- a. spoken expression *S*:

$$[\![\text{dog}]\!] = \lambda x \lambda w. \text{dog}(x, w)$$

- b. gesture *G*:

$$[\![\text{LARGE}_{\text{NP}}]\!] = \lambda x \lambda w. \text{large}(x, w)$$

- c. local context *c'*:

$$\lambda x \lambda w. C(w) \wedge \text{bring}(\mathbf{s}, x, w) \wedge \text{poss}(\mathbf{s}, x, w)$$

- d. cosupposition $c' \Rightarrow_{\forall} (S \Rightarrow G)$:

$$\forall x \forall w [(C(w) \wedge \text{bring}(\mathbf{s}, x, w) \wedge \text{poss}(\mathbf{s}, x, w)) \rightarrow (\text{dog}(x, w) \rightarrow \text{large}(x, w))]$$

\approx ‘For all individuals *x*, for all worlds *w* in *C*: if *Stephanie brings x in w* and *x belongs to her in w*, then if *x is a dog in w*, *x is large in w*.’

- e. local accommodation or non-triggering *S & G*:

$$\lambda x \lambda w. \text{dog}(x, w) \wedge \text{large}(x, w)$$

The assumptions about the denotations of *dog* and an NP-level *LARGE* are standard: I assume both denote $\langle e, st \rangle$ -type properties (so does Schlenker).⁴⁰ A somewhat more liberal assumption is made about the local non-propositional context c' of *dog*^{LARGE}, which also has to be of type $\langle e, st \rangle$. Under the assumption (made in Schlenker 2009, 2010) that when computing a local context we can take into account at least all the content linearly preceding the trigger, we can assume that the local context c' of *dog*^{LARGE} in the target example is *Stephanie brings her* and denotes the property of being an x such that *Stephanie* brings x and x belongs to her in a world of C. This is not a crucial assumption, as we could also restrict the domain to *Stephanie*'s dogs from the get-go (see the discussion in footnote 13), but I will make it to make things more transparent.

With these assumptions in place, we get desirable results for the target sentence. If the cosupposition projects, we end up with a projecting inference that if *Stephanie* brings a dog of hers, that dog is large. Since size is typically an individual-level predicate, it is natural to conclude from that that all *Stephanie*'s dogs are large, which gives us the ‘Projecting non-restricting’ interpretation from Chapter 3. If we locally accommodate the cosupposition, or if we don’t trigger it in the first place, the predicate denoted by the gesture composes as an ordinary restrictive modifier without any additional inferences, which allows for the ‘Restricting’ interpretation from Chapter 3, as desired.

The ‘Restricting’ interpretation can be degraded for some speakers because of the cost of local accommodation of cosuppositions. It is less obvious if non-triggering should incur any cost, but there can be a general bias against truth-conditionally non-vacuous co-speech gestures (which is the explanation I will ultimately maintain for why restricting interpretations of co-speech gestures are degraded). In other words, if co-nominal gestures are allowed to adjoin to NPs, Schlenker’s analysis can account for the first two bars for gestures in Figure 3.2, and at this point we are not generating any unattested interpretations.

⁴⁰Of course, the predicate description in the gesture’s denotation is simplified; size gestures indicate size iconically, so a more accurate description would be along the lines of ‘roughly of this size’ (this is not to say that the mapping is always accurate), but for my purposes this simplified denotation will suffice. Furthermore, I will not touch upon the issue of how we arrive at the lexical semantics of a given gesture, and in particular, of how we go from an iconic demonstration (roughly in the sense of Davidson 2015) to an object that is compositionally integrable with the rest of the utterance. I believe this to be an important question, but this dissertation is concerned with what happens after such a compositionally integrable object has been obtained.

4.2.3 DP-level gestures under the cosuppositional analysis

Things change if we allow co-nominal gestures to adjoin to DPs and allow the cosupposition mechanism to apply to them. Let us look at *Stephanie brings her dog_{LARGE}* again, but with *LARGE* adjoining to the DP *her dog*. Schlenker doesn't discuss this possibility, so we will have to make our own assumptions.

First, we need to figure out what DP-level gestures denote. Once again, for the cosupposition mechanism to apply, they should be of the same type as the DPs they adjoin to. It is unclear what it would mean for generalized material implication to hold between two individuals, so let us assume that both are generalized quantifiers of type $\langle\langle e, st \rangle, t \rangle$. But what exactly would an $\langle\langle e, st \rangle, t \rangle$ -type gesture denote? As a first pass, let us try treating such gestures as existential quantifiers. So, a DP-level *LARGE* would mean roughly 'a large object', which intuitively is a plausible construal.

Next, we need to figure out what the local context of *her dog*_{LARGE} is; it needs to be of type $\langle\langle e, st \rangle, t \rangle$, as well. Based on the linearly preceding material, it is not unreasonable to stipulate a local context along the lines of 'something that Stephanie brings in a world of C'. Again, how justified this stipulation is is a separate issue that I am not planning to discuss here. Also, in this specific case, this assumption isn't crucial at all— c' can be as simple as $\lambda P \lambda w. C(w) \wedge \exists x [D_e(x) \wedge P(x, w)]$ (i.e., 'something in a world of C'). I will, however, make the preceding material part of c' for consistency.

The derivation for the target example is given in (4.5).

(4.5) Stephanie brings $[[_{DP} \text{her dog}]]_{LARGE}$. (Attempt 1.)

- a. spoken expression S :

$$[[\text{her dog}]] = \lambda P \lambda w. P(\iota x. \text{dog}(x, w) \wedge \text{poss}(\mathbf{s}, x, w))$$

- b. gesture G :

$$[[LARGE_{DP}]] = \lambda P \lambda w. \exists x [\text{large}(x, w) \wedge P(x, w)]$$

- c. local context c' :

- $\lambda P \lambda w. C(w) \wedge \exists x [\text{bring}(\mathbf{s}, x, w) \wedge P(x, w)]$
- d. cosupposition $c' \Rightarrow_{\forall} (S \Rightarrow G)$:
- $$\forall P \forall w [(C(w) \wedge \exists x [\text{bring}(\mathbf{s}, x, w) \wedge P(x, w)]) \rightarrow$$
- $$(P(\iota x. (\text{dog}(x, w) \wedge \text{poss}(\mathbf{s}, x, w)), w) \rightarrow \exists x [P(x, w) \wedge \text{large}(x, w)])]$$
- \approx ‘For all properties P , for all worlds w in C : if Stephanie brings an object that has P in w , then if Stephanie’s dog has P in w , there is a large object that has P in w .’
- e. local accommodation or non-triggering $S \& G$:
- $$\lambda P \lambda w. P(\iota x. (\text{dog}(x, w) \wedge \text{poss}(\mathbf{s}, x, w)), w) \wedge \exists x [P(x, w) \wedge \text{large}(x, w)]$$
- We proceed with the derivation as usual and get:
- $$\lambda w. \text{bring}(\mathbf{s}, \iota x. (\text{dog}(x, w) \wedge \text{poss}(\mathbf{s}, x, w)), w) \wedge \exists x [\text{large}(x, w) \wedge \text{bring}(\mathbf{s}, x, w)]$$
- \approx ‘Stephanie brings her dog and a large object.’

Are these results good? If the purported cosupposition projects, we get a convoluted inference, which, at the end of the day, assures that all Stephanie’s dogs are large. That’s so, because one of the properties the resulting formula can be checked against is the property of being any specific dog. Let’s say, Stephanie has one dog, named Fang. The result above assures that if Stephanie brings an object that has the property of being Fang, then if Stephanie’s dog has the property of being Fang, there is a large object that has the property of being Fang. Or, much simpler, if Stephanie brings Fang, Fang is large. Once again, given that size is an individual-level predicate, we can conclude that Fang is large tout court.

However, the predicted local accommodation/non-triggering interpretation is not attested. And of course, if gestural cosuppositions can in principle be locally accommodated, they should be able to be locally accommodated regardless of whether the gesture adjoins to an NP or a DP. Similarly, if we need to start out with conjunctive semantics for co-speech gestures in order for cosuppositions to apply to them, and we want cosuppositions to apply to DP-level gestures, we would start out with conjunctive semantics for them as well.

Moreover, given the assumptions above, nothing prevents gestures from adjoining to quan-

tifiers like *no dog*, with some devastating results, as shown in (4.6) (I skip the intermediate steps).⁴¹

(4.6) Stephanie brings $[[_{\text{DP}} \text{no dog}]]_{\text{LARGE}}$.

- a. cosupposition $c' \Rightarrow_{\forall} (S \Rightarrow G)$:

$$\begin{aligned} \forall P \forall w [(\mathbf{C}(w) \wedge \exists x [\mathbf{bring}(\mathbf{s}, x, w) \wedge P(x, w)]) \rightarrow \\ (\neg \exists x [\mathbf{dog}(x, w) \wedge P(x, w)] \rightarrow \exists x [P(x, w) \wedge \mathbf{large}(x, w)])] \end{aligned}$$

\approx ‘For all properties P , for all worlds w in \mathbf{C} : if Stephanie brings an object that has P in w , then if there is no dog that has P in w , there is a large object that has P in w .’

- b. local accommodation or non-triggering $S \& G$ (final result):

$$\begin{aligned} \lambda w. \neg \exists x [\mathbf{dog}(x, w) \wedge \mathbf{bring}(\mathbf{s}, x, w)] \wedge \exists x [\mathbf{bring}(\mathbf{s}, x, w) \wedge \mathbf{large}(x, w)] \\ \approx \text{‘Stephanie brings no dog, and she brings a large object.’} \end{aligned}$$

The issue can be replicated with some other quantifiers. At this point one could say that the assumption that DP-level gestures denote existential quantifiers is incorrect. After all, under this assumption, there is no link between the spoken expression and the gesture. To introduce such a link and to avoid the issue with quantifiers, we could assume that DP-level gestures are anaphorically linked to the DPs they adjoin to, so that a DP-level $LARGE_i$ would mean something like ‘that_i thing and that_i thing is large’. Thus, when we have a gesture that’s trying to adjoin to a DP that doesn’t introduce a discourse referent, we get a failure. In a way, this would be an attempt to give DP-level gestures a semantics that is very similar to that of appositives while still making use of the cosupposition mechanism. This to my mind is a much less intuitive assumption than the one under which the gesture simply represents a large object, and it is unclear why NP-level gestures don’t have such a semantics, but at this point we are trying to see if we can in principle get reasonable

⁴¹ As for gestures adjoining to NPs within DPs like *no dog*, Schlenker’s empirical claim is that such examples are acceptable and give rise to the same projecting inference as with gestures adjoining to NPs within indefinite DPs like *a dog*. For the target example the inference would be ‘If Stephanie brings a dog, it is large’. As far as I can tell, to get this inference via the cosupposition mechanism, the local context in both cases would need to be the property of being brought by Stephanie—regardless of whether the determiner is *a* or *no*. Once again, the issue of whether this assumption is justified is beyond the scope of this dissertation.

results for DP-level gestures via the cosupposition mechanism, so let's allow this.

How would this tweak affect our results? I will skip some unsuccessful attempts to implement it without adjusting the local context of *her dog_{LARGE}* and will immediately suggest that local contexts can also come with anaphoric elements, including to something that is yet to come. This assumption is extremely controversial, but again, we are trying to give the cosuppositional analysis our best shot. In (4.7) I provide the derivations for *Stephanie brings her dog_{LARGE}* with the gesture adjoining to *her dog*, with these new assumptions in place. Now *her dog* bears an index *i*, which gets picked up both by *LARGE* and by the local context of *her dog_{LARGE}*.

(4.7) *Stephanie brings [[_{DP} her dog]_iLARGE_i].* (Attempt 2.)

- a. spoken expression *S*:

$$\llbracket [her \ dog]_i \rrbracket^g = \lambda P \lambda w. P(\iota x. \text{dog}(x, w) \wedge \text{poss}(\mathbf{s}, x, w))$$

- b. gesture *G*:

$$\llbracket LARGE_i \rrbracket^g = \lambda P \lambda i. P(g(i), w) \wedge \text{large}(g(i), w)$$

- c. local context *c'*:

$$\lambda P \lambda w. C(w) \wedge \text{bring}(\mathbf{s}, g(i), w) \wedge P(g(i), w)$$

- d. cosupposition *c' ⇒_{forall} (S ⇒ G)*:

$$\forall P \forall w [(C(w) \wedge \text{bring}(\mathbf{s}, \iota x. (\text{dog}(x, w) \wedge \text{poss}(\mathbf{s}, x, w)), w) \wedge$$

$$P(\iota x. (\text{dog}(x) \wedge \text{poss}(\mathbf{s}, x)), w)) \rightarrow \text{large}(\iota x. \text{dog}(x, w) \wedge \text{poss}(\mathbf{s}, x, w))]$$

≈ ‘For all properties *P*, for all worlds *w* in *C*: if *Stephanie brings her dog in w* and *her dog has P in w*, then *her dog is large in w*.’

- e. local accommodation or non-triggering *S & G* (final result):

$$\lambda w. \text{bring}(\iota x. (\text{dog}(x, w) \wedge \text{poss}(\mathbf{s}, x, w)), w) \wedge \text{large}(\iota x. \text{dog}(x, w) \wedge \text{poss}(\mathbf{s}, x, w))$$

≈ ‘*Stephanie brings her dog, and her dog is large.*’

The result for projection is quite good; once again, one of the properties we could consider is the property of being *Stephanie's dog*, so this result assures that all *Stephanie's dogs* are large. The result for local accommodation/non-triggering, however, is exactly the ‘Non-projecting non-

restricting’ interpretation the experiment in Chapter 3 showed to be unavailable—or at least significantly more degraded than the ‘Restricting’ interpretation, which is unexpected if both obtain via local accommodation/non-triggering of a gestural cosupposition.

We could try positing other $\langle\langle e, st \rangle, t \rangle$ -type denotations for DP-level gestures in the hope of making the cosupposition mechanism yield reasonable results, but we would keep running into the problem of unattested local accommodation/non-triggering interpretations.

I’d like to discuss the importance of this result. If this attempt to give DP-level gestures a cosuppositional semantics with an anaphoric twist had been successful, we could have given co-suppositional semantics to appositives, too. However, we would have run into the same problem: it would be entirely puzzling why cosuppositions triggered by appositives can’t be locally accommodated in the vast majority of cases. This conclusion, thus, echoes the one at the end of section 2.2.2, whereby there are inherent differences in how modifiers and supplements project, and we simply can’t use the same projection mechanism for non-restricting modifier inferences and supplements. It is, thus, imperative that cosuppositions be constrained to modifier construals only.

Furthermore, the failure of this attempt also shows that the suggestion made independently by Rob Pasternak and Manfred Krifka and discussed in Schlenker 2018a to treat appositive-like gestures⁴² as triggering cosuppositions with anaphoric elements is unlikely to work. Just like most spoken appositives, appositive-like gestures can’t be interpreted locally even under pressure, as shown in (4.8) (the gesture *LARGE* has its own time slot and follows the spoken expression *her dog* that it associates with), which would be entirely puzzling if they triggered cosuppositions.⁴³

- (4.8) #Do you know how big Stephanie’s dog is? ’Cause if she’s bringing her dog, **SMALL**,

⁴²Schlenker (2018a) calls such gestures “post-speech gestures”; I have my reservations about this notion, to which I come back in subsection 4.3.4.

⁴³However, I believe appositive-like gestures can be used as clarifications, similar to *one*-appositives previously exemplified in (2.50), for example:

- (iii) a. Bring me a beer, a small one.
- 
- b. Bring me a beer, SMALL.

we'll be fine, but if she's bringing her dog, **LARGE**, we should get a bigger van.

Intended: '...if (Stephanie is bringing her dog and her dog is ⟨small, large⟩)...'

Now, this general issue aside, if we only wanted to focus on co-speech gestures, we could stipulate a ban on DP-level co-speech gestures, but it is unclear how one would motivate such a constraint. First, this constraint would have to distinguish between gestural and non-gestural content in narrow syntax; in the next section I will talk about why this might be undesirable. Second, this narrow syntax constraint would have to be sensitive to linearization, which happens after narrow syntax, so we would be running into a look-ahead problem. The reason why this constraint would have to be sensitive to linearization is that, as Schlenker (2018a) himself argues, gestural supplements do exist. We have already seen an infelicitous example with appositive-like gestures in (4.8); (4.9) is a felicitous one.

(4.9) If Stephanie's bringing her dog, **LARGE**, we should get a bigger van.

→ Stephanie's dog is large.

Schlenker's main argument for gestures such as in (4.9) being supplements rather than presupposition-contributing expressions is that they seem to require discourse referents (see (4.1) for the relevant examples with appositives), as shown in (4.10).⁴⁴

(4.10) ??Stephanie didn't bring a dog, **LARGE**.

Thus, we would have to say that gestures can in principle adjoin to DPs, but when they do, they must be linearized as having their own time slot. Perhaps, one could come up with a principled reason why that would happen, but why would one even try? The only reason to do so would be if one wanted to maintain that co-speech gestures trigger cosuppositions across the board, but there is no obvious reason why this would be a desirable goal.

⁴⁴In my experience, the judgements are somewhat gradient and variable. For example, some people seem to be able to get the kind-targeting reading of the gesture, resulting in an inference that dogs in general are large.

4.3 Proposal: gestures can be modifiers or supplements

4.3.1 The principle of no gesture-specific composition

The problem with applying Schlenker’s cosuppositions to all co-speech gestures across the board is that we are then forced to allow all co-speech gestures, including DP-level gestures, to compose conjunctively with the spoken expressions they associate with either via local accommodation or by default—because that’s what we need to derive restricting interpretations of NP-level gestures. However, no spoken DP-level adjuncts compose conjunctively with the DPs they adjoin to. The only DP-level adjuncts I can think of are spoken adnominal appositives, and they certainly don’t do that. But if spoken adjuncts can’t do that, why should gestural adjuncts be allowed to? I propose that they shouldn’t.

In a series of recent presentations (Esipova 2017, 2018b,c, 2019c) I have been arguing that narrow syntax and semantics proper are modality-blind, i.e., when gestures integrate into a compositional structure, they do so in the same way as spoken content. In particular, gestural constituents can bear all and only syntactic labels spoken constituents can bear (i.e., no mysterious GestPs exist); they can merge at all and only sites where spoken constituents bearing the same syntactic labels can merge; they can semantically compose with other constituents in all and only ways available to spoken constituents, etc. Any modality-specific effects arise in phonology and its interfaces and in pragmatics and, due to their gradient and variable nature, are best captured via violable weighted constraints. Also, since compositional structures are not linearized, no distinctions are made at this level between co-speech gestures and gestures that have their own time slot. Such distinctions, too, only arise in phonology and its interfaces and in pragmatics.

This view is very appealing conceptually. An alternative would be to allow for gesture-specific composition strategies, and it is unclear how they would emerge or be acquired. Instead, my view relies on the intuition that, when faced with content-bearing gestures, speakers of a spoken language apply the general linguistic rules they have internalized mostly from spoken input to

said gestures in a way that best fits their content⁴⁵ and what little morphosyntactic information they carry. Oftentimes, several possibilities for compositional integration will be open for a given gesture, especially when it comes to co-speech gestures, whose level of adjunction in the syntax isn't always obvious from their temporal alignment.

Furthermore, a view that allows for gesture-specific compositionality would be hard to reconcile with any framework that relies on late lexical insertion, since one would need to know in advance if they are going to insert a spoken item from the lexicon or a gestural item, possibly created on the spot, into a given slot. My view is perfectly compatible with—and, in fact, favoring—a view of the architecture of grammar whereby narrow syntax simply generates labeled hierarchical structures and compositional semantics tells you how to interpret these structures, but neither knows or cares about what items will be inserted into the nodes at spell-out, and in particular, whether those items will be spoken or gestural.

4.3.2 Integrated and non-integrated gestures

Now, the view I propose does not imply that no gesture-specific interpretation strategies exist at all, only that compositional integration of gestures happens in the same way as that of spoken content. Similarly, I don't claim that all co-speech gestures integrate compositionally into the utterances they co-occur with. It does look like sometimes gestures co-occurring with some spoken content don't integrate with this content compositionally. In particular, it seems that humans are capable of simultaneously producing two compositionally independent utterances, in the spoken and in the gestural modality. The two can still interact at the level of prosody,⁴⁶ and we might still be able to establish some meaning correspondences between the two non-compositionally.

Some examples of what could be analyzed as compositionally non-integrated gestures that are still linked to the spoken content they co-occur with via some discourse relation are discussed in Hunter 2018. For example, Hunter claims that examples like (4.11a) can be analyzed along the

⁴⁵Conventionalized or not, although some conventionalized gestures can be quite constrained morphosyntactically. See section 4.5 for a brief discussion of the role of convention.

⁴⁶See, e.g., Loehr (2004) on alignment of spoken content and gestures, including non-content bearing ones.

same lines as (4.11b), i.e., as two independent utterances linked by a discourse relation only.

- (4.11) a. Do we have any more paper_{SCRIBBLE}?
- b. I want to write/draw something. Do we have any paper?

While I agree that a priori there is nothing that excludes an analysis along these lines for examples like (4.11a),⁴⁷ there is also nothing that enforces it.⁴⁸ The *SCRIBBLE* gesture in (4.11a) can also be construed of as a property, say, the property of being something one could draw or write on, in which case it would be a compositionally integrated modifier and will (by default) give rise to a projecting inference that paper is used for drawing or writing.

In fact, as mentioned before in subsection 2.1.2.1, non-restricting modifiers need to be licensed by a discourse relation (Leffel 2014), so, under the modifier analysis, in the right extralinguistic context, the addressee might infer that the speaker intends to write or draw on the paper, which is why they are using the non-restricting modifier. Thus, the two analyses, in a way, end up being mirror images of one another. On one analysis, the gesture is not compositionally integrated (at least not as a modifier) and directly communicates the speaker's intentions to use the paper for drawing or writing; the inference akin to the non-restricting modifier one emerges via a salient discourse relation. On the other analysis, the gesture is integrated as a modifier and (by default) gives rise to a non-restricting modifier inference; it is the inference about the speaker intending to use the paper for drawing or writing that arises via a discourse relation.

Most importantly, the two analyses correspond to two different construals of the example in (4.11a), both of which I believe to be in principle possible (it's a separate question if both are equally likely), rather than two different theories of co-speech gestures.

A less ambiguous example was given in Hunter 2019, cited from memory in (4.12), where the example with the gesture in (4.12a) is supposed to mean something like (4.12b).

⁴⁷ Although I'm a bit skeptical about how far the parallels between the two specific examples in (4.11) go, given that there is nothing obvious about the speaker's desires encoded in the *SCRIBBLE* gesture itself.

⁴⁸ Granted, Hunter doesn't use examples like (4.11) to argue that co-speech gestures are never compositionally integrated either, only that we don't need to assume that they are.

(4.12) *Context: The speaker is observing someone trying to fit a sheet of paper into an envelope.*

- a. It won't fit_{FOLD}.
- b. It won't fit. (You have to) fold it.

To the extent that (4.12a) is acceptable,⁴⁹ it does indeed seem to consist of two compositionally independent utterances. One piece of evidence in favor of this construal as the preferred one is that if we try to embed (4.12a) in a way that blocks this construal, the result is hard to make sense of, as shown in (4.13).

(4.13) a. ??If it won't fit_{FOLD}, let me know.
b. ??I doubt that it won't fit_{FOLD}.

Even if one were to argue for the existence of a conditional inference in (4.12a) (along the lines of ‘If it won’t fit, you have to fold it’), it doesn’t arise as a non-restricting modifier inference, because no restricting construal is possible for it (for instance, (4.13b), if it means anything at all, clearly can’t mean ‘I doubt that (it won’t fit and you have to fold it)'). Under the unconstrained cosuppositional theory of co-speech gestures from Schlenker 2018a, it would be similarly puzzling why local interpretations of this conditional inference are unavailable when they are available in other cases. If one were to maintain any compositional link between the two utterances in (4.12a), the most natural analogy I can think of would be a sentence-level appositive along the lines of (4.14), which makes the discourse relation between the two utterances explicit.

(4.14) It won’t fit, which [= the fact that it won’t fit] is why you have to fold it.

We can also arguably produce two utterances in the two modalities simultaneously without having any link between them whatsoever. The most obvious example of that would be two utterances, a spoken and a gestural one, produced simultaneously but directed at two different addressees. For

⁴⁹Some people find it awkward. I personally find it acceptable, and I myself produce utterances like this quite regularly.

example, I can be talking to someone while simultaneously beckoning someone else via a gesture (or, conversely, gesturing for them to go away). Or I can nod or show a thumbs-up gesture to the chair of a conference session, acknowledging that I've seen them showing the '10 minutes left' sign, without interrupting my talk.

Now, I would imagine that there are cognitive limits on how semantically rich the simultaneous utterances in the two modalities can be (which would likely vary across individual speakers), and exploring these limits would be an interesting research endeavor, but this dissertation will not do so and will continue to focus on compositionally integrated gestures.

Crucially, one can't maintain that gestures never integrate compositionally into otherwise spoken utterances; compositional integration of gestures is required at the very least for many cases of gestures with their own time slot. For the latter, I specifically have in mind what Schlenker (2018a) calls PRO-SPEECH GESTURES. Schlenker defines pro-speech gestures as gestures that "replace" spoken expressions. I will not adopt this definition. If taken literally, this definition seems to imply a very specific procedure of generating utterances with such gestures, whereby a speaker first generates an utterance with spoken lexical items inserted and then "replaces" some of the spoken items with gestures. I would prefer to maintain a more conservative view of the architecture of grammar under which this procedure is unlikely to be legal. Schlenker himself doesn't endorse this interpretation of his definition (Philippe Schlenker, p.c.), but it is unclear to me what a linguistically meaningful non-literal interpretation of his definition would be. I will still use the term PRO-SPEECH GESTURES, however, because I think it is useful to distinguish between prosodically dependent and prosodically independent compositionally integrated gestures, but I will define pro-speech gestures accordingly, as compositionally integrated gestures that are linearized as having their own time slot rather than co-occurring with speech. Five naturally-occurring examples of non-adjunct pro-speech gestures from TV shows are given in (4.15).

- (4.15) a. Well, you're a wonderful chef, you know, aside from being super-talented with the VIOLIN. ('Better Call Saul', S02E05, <https://youtu.be/ZRSHh6aqNGY>)

- b. Anyhow. I grab his knife and STAB-STAB-STAB.
 (‘The Marvelous Mrs. Maisel’, S01E03, <https://youtu.be/NiBwDkL6nF0>)
- c. I need a summary judgement hearing with Michaelis for Frey V. O’Connell, and
 I’m hoping for Thursday at 10, before he uh... DRINK.
 (‘Better Call Saul’, S02E05, <https://youtu.be/HytB0hEGwwI>)
- d. So... did you actually HIT-WITH-HAMMER White Girl?
 (‘iZombie’, S04E12, <https://youtu.be/jZxhZHzhMTI>)
- e. Look, if this option is too expensive you can still always STAB Badger in the chow
 line. (‘Breaking Bad’, S02E08, <https://youtu.be/dZxoPh24VIU>)⁵⁰

There is little doubt that such prosodically independent gestures are compositionally integrated in the same way as spoken expressions⁵¹ (in subsection 4.3.4 I show that such gestures also follow language-specific linearization, displacement, and prosodic grouping rules).

Pro-speech gestures aside, compositional integration of some sort is required at the very least for restricting interpretations of co-speech gestures, too, since those affect the truth conditions of the utterance.

The existence of both compositionally non-integrated and compositionally integrated gestures is only a problem under a view that there has to be one single way in which co-speech gestures contribute to the meaning of spoken utterances. Once one lets go of this assumption, the problem disappears. As said before, however, this dissertation mostly focuses on compositionally integrated gestures and on showing that gestures can integrate compositionally in different ways, too, which means they can also project in different ways.

⁵⁰It would seem that the actor Bob Odenkirk has a special affinity for pro-speech gestures (or perhaps his character Jimmy McGill a.k.a Saul Goodman does).

⁵¹Again, once they have been made into compositionally integratable objects, however that happens. See footnote 40.

4.3.3 Deriving the data at hand

Now, how does the principle of no gesture-specific composition apply to the data from Chapter 3?

Gestural adnominal content, just like spoken content, can compose as modifiers or as supplements. This goes in line with the intuition that, being non-lexicalized and carrying little morphosyntactic information, the gesture *LARGE*, for example, can be taken to iconically represent either the property of being large (thus, being akin to the adjective *large*) or a large object (thus, being akin to the nominal *a large object*). A property-like gesture will compose as a modifier; it will adjoin to the NP and can in principle be restricting or non-restricting. A nominal-like gesture will compose as a supplement; it will adjoin to the DP and will yield a proposition of a special kind about it, without having a compositional potential to be restricting.

In line with the general principle proposed in this dissertation that for non-sublexical content composition determines projection, supplement gestures will necessarily project and they will project as supplements (however their projection is operationalized). Modifier gestures will only give rise to projecting inferences when they are non-restricting, and these inferences will project in the same way as other non-restricting modifier inferences, as defined in (2.29). The result will be ultimately the same as in (4.4d), if we treat *LARGE* as an intersective modifier:

(4.16) Stephanie brings her $[[\text{NP dog}]_{\text{LARGE}}]$.

a. spoken expression S :

$$\llbracket \text{dog} \rrbracket = \lambda x \lambda w. \text{dog}(x, w)$$

b. gesture G :

$$\llbracket \text{LARGE}_{\text{NP}} \rrbracket = \lambda x \lambda w. \text{large}(x, w)$$

c. spoken expression modified by the gesture $[[S]^G]$:

$$\llbracket [[\text{dog}]_{\text{LARGE}}] \rrbracket = \lambda x \lambda w. \text{dog}(x, w) \wedge \text{large}(x, w)$$

d. propositional local context c' :

C

e. non-propositional local context c'' :

- $$\lambda x \lambda w. \llbracket c' \rrbracket(w) \wedge \text{bring}(\mathbf{s}, x, w) \wedge \text{poss}(\mathbf{s}, x, w)$$
- f. non-restricting modifier inference $c'' \Rightarrow_{\forall} (S \Rightarrow [S]^G)$:
- $$\begin{aligned} \forall x \forall w [(\mathbf{C}(w) \wedge \text{bring}(\mathbf{s}, x, w) \wedge \text{poss}(\mathbf{s}, x, w)) \rightarrow \\ (\text{dog}(x, w) \rightarrow (\text{dog}(x, w) \wedge \text{large}(x, w)))] \\ \forall x \forall w [(\mathbf{C}(w) \wedge \text{bring}(\mathbf{s}, x, w) \wedge \text{poss}(\mathbf{s}, x, w)) \rightarrow (\text{dog}(x, w) \rightarrow \text{large}(x, w))] \\ \approx \text{'For all individuals } x, \text{ for all worlds } w \text{ in C: if Stephanie brings } x \text{ in } w \text{ and } x \\ \text{belongs to her in } w, \text{ then if } x \text{ is a dog in } w, x \text{ is large in } w.' \end{aligned}$$

In contexts supporting the ‘Projecting non-restricting’ interpretation, an adnominal gesture can thus be either a non-restricting modifier or a supplement. In contexts supporting the ‘Restricting’ interpretation, a gesture can only be a modifier. Neither strategy can yield the ‘Non-projecting non-restricting’ interpretation, because both supplements and non-restricting modifiers have to project. As I showed in subsection 2.1.2.3, local accommodation is not a meaningful notion for non-restricting modifier inferences.

At this point we predict gestures to pattern exactly like adjectives across the three interpretation types, which is not what we see in Figure 3.2. We still need to explain why restricting interpretations of gestures are degraded for many people. Under the view I propose, this would not be due to any constraints in narrow syntax or semantics; compositional semantics allows for restricting interpretations of co-speech gestures. Instead, the degradedness will have to be due to some bias against restricting interpretations of co-speech gestures due to pragmatic or phonological considerations or both.

One potential explanation is following the original intuition in Schlenker 2018a that there is a general pragmatic preference for co-speech gestures to be truth-conditionally vacuous, the idea being that when you have an expression in a more primary (e.g., spoken) modality co-occurring with an expression in a secondary (e.g., gestural) modality, the latter should not be encoding anything affecting the truth conditions of the utterance. Restricting interpretations are by definition truth conditionally non-vacuous, so such interpretations of gestures should be dispreferred.

A more involved explanation, relevant specifically for the data discussed here and in Esi-pova 2019a, is that it's hard for vocal prosodic prominence to associate with co-speech gestures, once again the idea being that there is some competition between co-occurring spoken and gestural content in which the former has the priority. If it's hard to mark co-speech gestures as prosodically prominent, it will be hard to semantically focus them, which is necessary for the restricting interpretations in (3.5b). The two explanations are not mutually exclusive, and both allow for gradience and variability, so both can be at play.

With a pragmatic and/or prosody-related bias against restricting co-speech gestures in place, we now derive the results for gestures in Figure 3.2.

Now, let me add a quick note about modifier vs. supplement gestures. I take it that by default both construals are possible, and as I said before, temporal alignment of a co-speech gesture is unlikely to be a good predictor of its level of adjunction (NP or DP). It is possible that some gestures (especially, conventionalized ones) are more likely to be construed of as modifiers or supplements based on their content and/or morphosyntax. For size gestures, which are the only kind of gestures I used in the experiment from Chapter 3, I don't see a reason to prefer one construal over the other (although, of course, there might be speakers who do have a relevant preference).

Again, as mentioned before, only modifier gestures can be restricting, so if you have restricting gestures in an utterance, you know they have to be modifiers. Another case in which a co-nominal gesture has to be construed as a modifier is when the nominal expression the gesture co-occurs with does not introduce a discourse referent and thus cannot be an anchor for a supplement. These are the cases previously discussed in (4.1), partially repeated in (4.17).

- (4.17) a. %?Stephanie didn't bring a dog_{LARGE}.
b. %?Stephanie brought no dog_{LARGE}.

As I mentioned before, such examples are somewhat degraded for many speakers I have consulted. Should we take it to mean that modifier construals of size gestures are dispreferred for some reason (this alone could explain the degradedness of restricting gestures in Figure 3.2, since those have to

be modifiers)? I wouldn't jump to this conclusion. I think what's going on in (4.17) is that since the gestures there have to be modifiers, and modifier gestures prefer to be non-restricting, we end up with a violation of the *Minimize Restrictors!* principle (mentioned previously in subsections in subsection 2.1.2.1 and 3.4.2). In the 'Projecting non-restricting' examples in the Chapter 3 experiment these violations were routinely justified in the context (the size always mattered), but there are no similar justifications in the out-of-the-blue examples in (4.17), and in fact, it would be hard to come up with the right context in which the size of the dogs Stephanie typically brings (or even all dogs) would matter if she didn't bring one.

4.3.4 Gestures and linearization

The bias against restricting interpretations introduced in the previous subsection does not apply to pro-speech gestures, as it is competition with co-occurring spoken content that is responsible for it. Furthermore, under the view of how gestures fit into the architecture of grammar developed in section 4.3.1, gestures integrate compositionally before they linearize (just like spoken expressions). Thus, modifier gestures should in principle be able to linearize as pro-speech, and they furthermore should be able to receive restricting interpretations fairly easily. However, various linearization, prosodic grouping, and articulatory considerations make adnominal pro-speech gestures construed of as modifiers unlikely in English. I will now discuss these considerations and how some of them are language-specific. Earlier versions of this discussion can be found in Esipova 2017, 2018a,c.

Schlenker (2018a) identifies a category of "post-speech gestures" and argues that they are interpreted as and project as supplements. He defines "post-speech gestures" as "gestures that come after the expressions they modify".⁵² This definition makes no reference to how these gestures are packaged prosodically, but the tacit assumption (visible in his notation and the way he produces the relevant examples himself) is that these gestures are packaged into their own prosodic phrases.⁵³

⁵²He seems to be using the term "modify" to mean "adjoin to"; under my use of this term, supplemental gestures are not modifiers.

⁵³I will remain non-committal about whether for English these correspond to intermediate or intonational phrases in ToBI (Beckman & Ayers 1997); for the rather broad distinctions I am drawing in this dissertation this doesn't seem to matter. I do think, however, that they are usually intonational phrases, because the preceding phrase typically ends

Such gestures cannot have restricting interpretations, at least not in English, as shown in (4.18).

- (4.18) #Do you know which one of Stephanie's dogs is coming with us? 'Cause if she's bringing her dog, (_{P_{TP}} **SMALL**), we'll be fine, but if she's bringing her dog, (_{P_{TP}} **LARGE**), we should get a bigger van.

Intuitively, it is clear why (4.18) is bad. Configurationally, these gestures resemble English appositives (and nothing much else), which also linearize as following their anchors and are packaged into their own prosodic phrases (see Selkirk 2005 and references therein for this latter point), and are, thus, construed of as such. But, of course, appositives can't have restricting interpretations, so the gestures in (4.18) can't either. However, the mere claim that gestures compositionally integrate into utterances in the same way as spoken expressions doesn't yet capture this intuition. We need to posit a further principle that prosodically independent gestures that compositionally integrate into otherwise spoken utterances follow (or at least tend to follow) language-specific linearization and prosodic grouping rules for spoken content. The same principle doesn't apply to co-speech gestures, naturally, as those linearize differently from any spoken expressions and are prosodically parasitic on the spoken string they co-occur with—a configuration that is simply impossible for spoken expressions.

In English, one-word size modifiers are typically adjectives that linearize on the left of the NP they modify. Furthermore, modifiers in English prefer to be in the same prosodic phrase as the constituent they modify (in Esipova 2018a,c, I formalize this constraint for NP-level modifiers as a narrow version of Truckenbrodt's (1999) WRAPXP constraint on syntax/prosody mapping).⁵⁴ It is thus expected that prosodically independent size gestures composing as modifiers would also prefer to linearize on the left of the NPs they modify and be in the same prosodic phrase as those NPs, as in (4.19)—except (4.19) is very hard to produce.

in a –% boundary tone sequence.

⁵⁴Again, I am using the term “modifier” as defined in this dissertation; appositives are not modifiers, so this doesn't apply to them.

- (4.19) ??Do you know which one of Stephanie's dogs is coming with us? (_{PrP} 'Cause if she's bringing her **SMALL** dog), we'll be fine, (_{PrP} but if she's bringing her **LARGE** dog), we should get a bigger van.

Articulatory integration of prosodically independent gestures is a very interesting issue in and of itself, and its effect on prosodic grouping and, consequently, on what syntactic structures are compatible with the resulting string is crucial for how these gestures can be interpreted. Intuitively, the articulatory awkwardness in (4.19) is at least partially due to the necessity to switch back and forth from spoken to gestural material within one prosodic phrase.⁵⁵ In Esipova 2018c, I proposed an OT-style constraint that punishes any instance of a boundary of a prosodically independent gesture that is not aligned with a prosodic phrase boundary.⁵⁶ I will not reproduce the entire analysis here; the general consequence of this constraint is that mixing spoken expressions and prosodically independent gestures within one prosodic phrase is suboptimal.

However, adjunct gestures like in (4.19) can also be linearized as co-occurring with the spoken expressions they adjoin to, as in (4.20), which is what we have seen before.

- (4.20) ?Do you know which one of Stephanie's dogs is coming with us? (_{PrP} 'Cause if she's bringing her dog_{SMALL}), we'll be fine, (_{PrP} but if she's bringing her dog_{LARGE}), we should get a bigger van.

(4.20) is easy to produce, but, of course, we now have to deal with the fact that co-speech gestural modifiers don't like to be restricting.

Now, the non-adjunct pro-speech gestures in (4.15) all want to be in the same prosodic phrase as the adjacent spoken material, but linearizing as co-speech doesn't seem to be an option

⁵⁵Relatedly, Schlenker (2018b) says that pro-speech gestures "are happiest at the end of a clause" and claims it to be a constraint "on the syntax of pro-speech gestures". While his empirical observation goes in the right direction, I believe it's too coarse-grained and is made at the wrong level of representation (syntax rather than prosody).

⁵⁶There I also posited asymmetries between left and right boundaries; I am not entirely sure this was justified, however, and I now think the more subtle differences I was trying to get at have more to do with how prominence is assigned within a prosodic phrase. I will not pursue this issue here, though.

for them.⁵⁷ Intuitively, this is not surprising—they are not adjuncts, so what would they be co-speech with? This linearization constraint is, of course, gesture-specific,⁵⁸ and it has to be, because, once again, spoken expressions can't linearize as co-occurring with other spoken expressions.

Now, if you watch the videos of the examples in (4.15), you'll notice that the articulatory integration strategy that is used in all of them is pausing and lengthening of the spoken syllable preceding the gesture,⁵⁹ but, as far as I can tell, in none of the examples there are actual prosodic boundaries around the gestures (there are no boundary tones, no pitch reset, etc.). This is an interesting observation, because, if robust, this would suggest that speakers prefer not to violate prosodic grouping rules even if that means they have to resort to various repair strategies to integrate gestural material articulatorily.

Examples (4.15b)–(4.15e), in which the pro-speech gestures are verbal (VPs or V's), all contain another strategy that facilitates articulatory integration of pro-speech gestures into an otherwise vocal stream, namely, iconic vocalizations accompanying the gestures (grunting for both stabbing gestures, gulping for the drinking gesture, and clicking for the hammering gesture). Schlenker (2018b) notices that pro-speech gestures are often accompanied by such vocalizations and attributes that to a general tendency to maximize iconicity and perhaps a need to justify the use of pro-speech gestures rather than words in the first place. While I agree that the iconic nature of the vocalizations is important, I also think these vocalizations help articulatory integration and can be used even when there is nothing iconic about them. For example, if one were to try and make (4.19) slightly easier to pronounce, they could produce some generic vocalizations (I use dental clicks) accompanying the size gestures, for which there are no obvious iconic vocalizations; this is especially crucial for (4.19) as the gestures are supposed to be focused.

Now, while the articulatory considerations above should not be language-specific (although

⁵⁷Again, here I don't take into account the preparatory phase of the gesture, which can start very early; I assume that what matters for the co- vs. pro-speech distinction is the alignment of the stroke of the gesture.

⁵⁸Or, perhaps, specific to any content that can in principle be linearized as co-occurring with something in a more primary modality within a given string. How we determine the hierarchy of modalities in cases when it is not spoken vs. non-spoken content is an interesting question, which will come up again in the next section when I talk about facial expressions, but I will not address it here.

⁵⁹This is especially obvious in (4.15c), where there is a bona fide hesitation pause before the gesture, but is present to some extent in all the examples.

individual speakers might vary at how apt they are at articulatory integration of prosodically independent gestures into spoken utterances and how willing they are to make the effort in the first place), the linearization and prosodic grouping considerations are. There might be some universal tendencies when it comes to prosodic grouping, but the mere fact that different languages have different inventories of prosodic units makes prosodic grouping language-specific. Under the view that I propose here, this suggests that other languages might be more likely to allow for modifier pro-speech gestures. This is indeed corroborated by some data from French and Russian I adduce in Esipova 2018c and repeat here.

In French, many adjectival modifiers are linearized on the right of the NPs they modify. This is actually not the case for all size adjectives, but in (4.21), I use adjectives for which this is the case.

- (4.21) (PrP Si Mélanie amène son chien **minuscule**), ça ira, (PrP mais si elle amène son if Mélanie brings her dog tiny it will-go but if she brings her chien **géant**), ce sera un problème.
dog gigantic it will-be a problem
'If Mélanie brings her tiny dog, that will be OK, but if she brings her gigantic dog, that will be a problem.' (French)

Based on the judgements of three native French speakers (one of whom is also a native speaker of English), the version of (4.21) with gestures instead of adjectives, as in (4.22), is not that bad and is certainly better than the English counterpart. The gestures in (4.22), of course, have to be modifiers, because they have to be restricting.

- (4.22) ?(PrP Si Mélanie amène son chien **SMALL**), ça ira, (PrP mais si elle amène son chien if Mélanie brings her dog SMALL it will-go but if she brings her dog **LARGE**), ce sera un problème.
LARGE it will-be a problem
'If Mélanie brings her small dog, that will be OK, but if she brings her large dog, that will be a problem.' (French)

The French example in (4.22) still has the issue of articulatory integration. It is not as severe as in English because the gestures are aligned with prosodic boundaries on the right. It's also true that in French the postnominal modifiers will be in their own Accent Phrases (see Delais-Roussarie et al. 2015 for an overview of French prosody), but this does not seem to be enough to fully mitigate the issue (at least for two of the three speakers that I consulted).

In Russian, size adjectives are by default linearized on the left of the NPs they modify, but Russian allows for much more syntactic displacement than either English or French. In particular, in Russian you can move out an NP out of a DP. For example, (4.23a) illustrates the default word order, and (4.23b) illustrates a ‘contrastive topic + focus’ configuration whereby the NPs are contrastive topics and have been fronted, and the AdjPs are left behind and are focused (and interpreted as restricting).⁶⁰

- (4.23) a. (_{PrP} Ja zakazala malen'kij kofe) (_{PrP} i bol'soj čaj).
 I ordered small coffee and big tea
 ‘I ordered a small coffee and a large tea.’ (Russian)
- b. (_{PrP} [Kofe]_{CT} ja zakazala) (_{PrP} [malen'kij]_F), (_{PrP} a [čaj]_{CT})—(_{PrP} [bol'soj]_F).
 coffee I ordered small and/but tea big
 ‘As for coffee, I ordered a small one, and as for tea, I ordered a big one.’ (Russian)

My prosodic phrase boundaries in (4.23) should be taken with a grain of salt, as no good ToBI-like system exists for Russian, but the focused constituents are packaged into some sort of an ostensible prosodic unit (I would be more inclined to equate it with the English intermediate phrase). This is fortunate, because now, if we have a counterpart of (4.23b) with gestures instead of adjectives, we can articulatorily integrate the gestures relatively easily. Indeed, I find (4.24) quite natural (to the extent that examples with pro-speech gestures are natural to begin with) and easy to produce (although I still prefer to have clicks accompanying the gestures, especially because they are focused).

⁶⁰The English translation in (4.23b) might mislead one to believe that in such examples the contrastive topics are base-generated, and the focused constituents are nominals with silent heads; the case assignment facts show clearly that this is not the case, and we do in fact have fronted NPs with AdjPs left behind.

- (4.24) (_{PrP} [Kofe]_{CT} ja zakazala) (_{PrP} []_F **SMALL**)_F, (_{PrP} a [čaj]_{CT})—(_{PrP}
 coffee I ordered **SMALL** and/but tea



[]_F **LARGE**).
LARGE

‘As for coffee, I ordered a small one, and as for tea, I ordered a large one.’ (Russian)

Note that the gestures in (4.24) satisfy Schlenker’s definition of “post-speech gestures”—they are adjuncts, and they do linearly follow the expressions they adjoin to—but they are certainly not supplements (just like the adjectives in (4.23b), they are interpreted as restricting modifiers).

Of course, Schlenker doesn’t mean to extend his notion of “post-speech gestures” to such cases. The “post-speech gestures” he has in mind are just pro-speech gestures under the refined definition proposed here (i.e., compositionally integrated gestures with their own time slot) whose configurational properties (in particular, ability to be construed of as adjuncts to something preceding them and prosodic packaging into their own phrases) resemble those of English appositives (and nothing much else) and they are thus likely to be construed of as such in English.⁶¹ The definition of “post-speech gestures” given in Schlenker’s work (“gestures that come after the expressions they modify”), however, does not—and cannot possibly—capture that. In fact, it doesn’t pick out a natural class of linguistic objects at all. Once, however, we give the objects that Schlenker has in mind a proper linguistic definition, it becomes painfully obvious that singling out “post-speech gestures” as a separate category makes no more sense than singling out gestural VPs, gestural subjects, gestural objects, gestural modifiers, etc. The three-way distinction put forward in Schlenker’s work among “co-speech”, “pro-speech”, and “post-speech” gestures is, thus, not a meaningful one, as the underlying distinctions among these categories are not all of the same nature, and “post-speech gestures” should be subsumed under the refined definition of pro-speech gestures.

⁶¹This follows naturally from the principles I have been advocating for here that call for maximum similarity between how gestural and spoken content integrates into spoken utterances.

All this doesn't undermine the empirical value of Schlenker's observation that gestural supplements exist, but it does expose the theoretical inadequacy of such pre-linguistic notions as "post-speech gestures" or Schlenker's original definition of "pro-speech gestures", which arises as an inevitable consequence of any approach to investigating the semantics of gestures that is not based on a clear, explicit view of how gestures integrate into the linguistic structure at other levels of representation. Pre-theoretical notions might work for preliminary empirical generalizations, but one can't build a theory of projection based on them.

The upshot is that if we want to approach gestures as linguistic objects (which we clearly have to in the case of compositionally integrated gestures), we need to do so at all levels of representation. And for that we need to ask the question of syntactic integration of gestures in proper syntactic terms, with attention to syntax/prosody mapping and articulatory integration of gestures. The linear position of a given piece of content, spoken or gestural, doesn't directly determine its interpretation or projection, but can be indicative of its syntactic integration, which, of course, feeds into its compositional integration, which, in turn, determines how it will project, if it does.

4.4 Beyond hand gestures: facial expressions

So far I have made two crucial points based on the data on hand gestures. First (more narrowly), Schlenker's (2018a) cosuppositions cannot possibly determine how co-speech gestures are interpreted or how they project across the board, unless we are willing to constrain syntactic integration of gestures in a non-trivial way. Second (more broadly), we need to study integration of non-spoken content into utterances in a linguistically meaningful way at all levels of representation. I will now strengthen both points by adducing empirical observations about content-bearing facial expressions (which are often taken to be a type of gestural content).

In his other work, Schlenker (2018b,c) develops a broader view of cosuppositions as assertion-dependent inferences that project like presuppositions and can be triggered by all sorts of iconic content. In Schlenker 2018c, he specifically develops what he calls a typology of "iconic

enrichments”, which is meant to determine how a given piece of content can project based on the following parameters:

- (i) whether this is an internal or external enrichment, i.e., whether it is “syntactically eliminable” (no definition is given for what it means to be syntactically eliminable, nor diagnostics for how to determine whether a given piece of content is syntactically eliminable);
- (ii) whether it has a separate time slot or co-occurs with something else (the tacit assumption seems to be that the enrichment exists in a secondary modality relative to the thing it co-occurs with, however the hierarchy of modalities is established).

This typology puts co-speech facial expressions in the same bin as co-speech hand gestures; according to it, both are external enrichments and don’t have a separate time slot. Schlenker claims that such enrichments trigger cosuppositions across the board.

The data for facial expressions discussed by Schlenker (both in spoken and sign language) all involve a facial expression expressing disgust (or some sort of negative attitude). It is hard for me to engage with those data; I find some of those examples unnatural and don’t get some of the reported judgements, and even when I do, there are often other interpretation possibilities I could envisage. This is not to question the validity of Schlenker’s data, just to point out one of the reasons I personally don’t want to use such facial expressions here.

Furthermore, in line with what I have been doing for hand gestures, I think it would be informative to have a systematic comparison between facial expressions and spoken expressions with similar semantics. In the case of negative attitude facial expressions, a natural point of comparison would be spoken expressives (keeping in mind that spoken expressives will likely be more constrained than facial expressions in where they can adjoin by their morphosyntax). As I mentioned before in footnote 22 in Chapter 2, true expressives have a distinctive property of being able to convey an attitude about things other than what they adjoin to syntactically. Thus, the example in (4.25), due to Alicia Parrish (p.c.), does not necessarily give rise to an inference that the speaker hates pens in general, or this specific pen, or pen-giving events, etc.; it can be uttered in a context

in which the speaker is simply angry (for example, at the addressee for taking their pen without permission).

(4.25) Give me my fucking pen!

This property of expressives really suggests that they either don't compose with the things they merge with in the standard way (or at all, if they are late merged), or that there is something special about their lexical semantics, which makes this composition vacuous at least sometimes.⁶² I believe we first need to check if facial expressions expressing general positive or negative attitudes have this property, because if they do, we better set them aside before we understand composition of spoken expressives better.

In what follows I will focus instead on what I will refer to as MIRATIVE FACIAL EXPRESSIONS, i.e., facial expressions conveying surprisal (broadly speaking), and I will be systematically comparing them to spoken adverbs with similar semantics. In particular, I have in mind the facial expression whose most prominent characteristic is one's eyes open wide (which also results in raised eyebrows) and variations thereof.⁶³ In what follows I will gloss this facial expression as *O_O*.

Spoken adverbs like *surprisingly*, *impressively*, *unfortunately*, etc. can target propositional content, in which case they merge as sentence-level supplements (and are sometimes referred to as "high adverbs"), as shown in (4.26a). Besides, many of such adverbs can be used as degree modifiers of properties, akin to *very* or *extremely*, in which case they merge locally with those properties, as shown in (4.26b).

⁶²The analysis of expressives as compositionally flexible expressions that can attach to any constituent at LF in Potts 2005 is still too compositional to account for examples like (4.25), as under this analysis, the expressive still needs to compose with something within the sentence, which doesn't seem to be the case in (4.25). Potts's later analysis (Potts 2007), which treats expressives as operators that affect the context of utterance in a certain way (namely, by shifting some expressive index, formalized as a triple of two salient objects *a* and *b* and a numeric interval within $[-1, 1]$ encoding the attitude of *a* towards *b*), or some variation thereof, seems to be a more viable alternative.

⁶³For example, one can add an evaluative component—for instance, by pursing their lips and/or shaking their head to express disapproval, by protruding their lips to express being positively impressed, by adding a smirk to express being amused, etc. I do some of these in the examples below, especially in facial expressions co-occurring with prospeech gestures, since in these cases I am not constrained by vocal articulatory considerations; these additions help make the utterances more natural and give rise to further inferences, but they don't affect my main points at all.

- (4.26) a. Yesterday there was a party, and, ⟨surprisingly, impressively, unfortunately, *very, *extremely⟩, Mia got drunk.
- ✓ It is ⟨surprising, impressive, unfortunate⟩ that Mia got drunk.
- ✗ Mia got drunk to a(n) ⟨surprising, impressive, unfortunate, high, extreme⟩ extent.
- b. Yesterday there was a party, and Mia got ⟨surprisingly, impressively, unfortunately, very, extremely⟩ drunk.
- ✗ It is ⟨surprising, impressive, unfortunate⟩ that Mia got drunk.
- ✓ Mia got drunk to a(n) ⟨surprising, impressive, unfortunate, high, extreme⟩ extent.

I observe that co-speech mirative facial expressions can have both interpretations as well, although their syntactic attachment is harder to pinpoint, which causes more ambiguity. In particular, it would seem that there is a general tendency to interpret *O_O* as a sentence-level supplement when it spans the entire sentence (or rather when its onset is at the beginning of the sentence) and as a degree modifier when it only co-occurs with the targeted property. Both preferences can be overridden (I think), but crucially, both interpretations are in principle possible, even though the facial expression remains co-speech. The benefit of focusing on this facial expression is that the two readings due to two different adjunction sites are relatively easy to disentangle (which is not always the case in expressive-like facial expressions).



- (4.27) a. Yesterday there was a party, _____ and Mia got drunk^{O_O}.
- ✓ It is {surprising, impressive} that Mia got drunk.
- ? Mia got drunk to a(n) {surprising, impressive, high} extent.
- b. Yesterday, there was a party, and Mia got drunk^{O_O}.
- ? It is {surprising, impressive} that Mia got drunk.
- ✓ Mia got drunk to a(n) {surprising, impressive, high} extent.

Now, I am well aware of the fact that (4.27b) is likely to contain additional iconic prosodic modu-

lations, such as lengthening of the vowel and/or lower F0 on *drunk*, which often have interpretative effects, as discussed, in particular, in Schlenker 2018b. One could then argue that it is those modulations that are driving the interpretation, and the facial expression is just coming along for the ride. While I think that's not the case and the effect of iconic prosodic modulations combined with co-speech facial expressions in such cases is cumulative,⁶⁴ to keep the contributions of the facial expressions separate from those of prosodic modulations, I will replace the spoken predicate *drunk* with a Russian conventionalized gesture that means ‘drink (alcohol)’ or ‘drunk’.⁶⁵ The data, of course, can be replicated with co-speech facial expressions only, if necessary.

The gesture consists of flicking one's finger on one's neck; a variation exists whereby one taps their neck with the back of their hand. I choose the first version of the gesture in my examples, first, because it is the one I use myself, and second, because its articulatory characteristics make it hard to accidentally modulate it iconically (although, ideally, one still needs to control for how long the onset and the holding end phase of the gesture is). The judgements reported for examples with this gesture are mine; I am keeping the examples themselves in English for simplicity—my judgements for such examples are very robust for both English and Russian. (4.28) repeats (4.27), but with the *DRUNK* pro-speech gesture as the target predicate.



- (4.28) a. Yesterday, there was a party,  and Mia got DRUNK^{O.O}.

video: <https://youtu.be/pw2nXfIU8Ro>

⁶⁴My (currently) favorite naturally occurring example of at-issue multi-modal modification is given below. I will not attempt to disentangle the entire compositional structure there (as in what modifies what), but it is clear that the gesture, the lower F0 and louder voice, and the mirative facial expression all contribute to the interpretative effect, creating the image of a bigger, fuller chest.

(iv) *Context: the speaker is giving advice on how to build pec muscles.*
 You won't have a chest that looks like this_{POINT-AT-OTHER-PERSON'S-CHEST}, you'll actually have
a chest_{lower-F0+louder-voice+O.O}
LARGE-ROUND.
 (Jeff Cavaliere, ‘Athlean-X’ Youtube channel, <https://youtu.be/6Qrr0GLeGp0?t=821>)

⁶⁵If one is to investigate this issue experimentally in the future, they could also control for that by using the same audio track with different videos (this is what I did in Esipova 2019a). For small-scale introspective judgements I don't find this technique helpful, though.

✓ It is {surprising, impressive} that Mia got drunk.

? Mia got drunk to a(n) {surprising, impressive, high} extent.



- b. Yesterday there was a party, and Mia got $\overline{\text{DRUNK}}^{\text{O-O}}$.

video: <https://youtu.be/1X45ztoZ6OQ>

? It is {surprising, impressive} that Mia got drunk.

✓ Mia got drunk to a(n) {surprising, impressive, high} extent.⁶⁶

Schlenker's 2018b typology predicts that O_O should give rise to a cosupposition under both interpretations. Now, there is technically a question about which modality is primary in (4.28b), since in this case the facial expression is co-gesture, not co-speech. I believe that fully integrated pro-speech gestures have the same primary status as spoken expressions, and Schlenker himself seems to assume that, given that he grants them AT-ISSUE (i.e., non-projecting) status by default. Furthermore, once again, the data below can be replicated without any pro-speech gestures, using co-speech facial expressions only and controlling for vocal prosodic modulations.

Now let us test this prediction. Embedding spoken sentence-level adverbs like *surprisingly* gives rise to conditional inferences, as shown in (4.29a). The same seems to be true for sentence-level O_O , as shown in (4.29b).

- (4.29) a. When, ⟨surprisingly, impressively⟩, a friend of mine gets drunk, I sometimes comment on that.

→ When a friend of mine gets drunk, this fact is ⟨surprising, impressive⟩.

- b. When a friend of mine gets $\overline{\text{DRUNK}}^{\text{O-O}}$, I sometimes comment on that.

(under the sentence-level reading) → When a friend of mine gets drunk, this fact is {surprising, impressive}.

The mere fact that the inference we get in (4.29) is conditional doesn't mean that we should ana-

⁶⁶I also think it's possible to combine the two readings by having two onsets for the mirative facial expression.

lyze it as a cosupposition. It is a well-known fact, discussed, for example, in Schlenker 2013, that supplements can in principle give rise to conditional inferences. In (4.29), there is no anchor for the supplement in the actual world, so the conditional nature of those inferences is... unsurprising.

Degree modifier adverbs and facial expressions, however, don't seem to project by default (i.e., by default they seem to be part of the at-issue content), as shown in (4.30). In this respect they are no different from other degree modifiers like *very* or *extremely*. This already goes against the prediction that the facial expression in (4.30b) should trigger a cosupposition.

- (4.30) a. When a friend of mine gets ⟨surprisingly, very, extremely⟩ drunk, I sometimes comment on that.
 - ↗ When a friend of mine gets drunk, they do so to a ⟨surprising, high, extreme⟩ extent.
- b. When, a friend of mine gets $\overline{\text{DRUNK}}^{0\text{-}0}$, I sometimes comment on that.
 - ↗ When a friend of mine gets drunk, they do so to a ⟨surprising, high, extreme⟩ extent.

Not only is it the case that the degree modifier *O_O* doesn't seem to project by default, it is also the case that the sentence-level *O_O* cannot be interpreted locally, as shown in (4.32)—just like spoken sentence-level adverbs, as shown in (4.31).

- (4.31) a. When a friend of mine gets drunk, I usually don't say anything, but when a friend of mine gets {surprisingly, very} drunk, I sometimes comment on that.
- b. #When a friend of mine gets drunk, I usually don't say anything, but when, ⟨surprisingly, impressively⟩, a friend of mine gets drunk, I sometimes comment on that.

Intended interpretation: ‘When (a friend of mine gets drunk and I am ⟨surprised, impressed⟩ by this fact)...’⁶⁷

⁶⁷One has to be very careful about the prosody here. If one fails to package the adverb in its own intonational phrase (and not just intermediate phrase) here, one might be able to accept this sentence. Cf. a similar sentence with *unexpectedly* as an actual restricting modifier:

- (4.32) a. When a friend of mine gets DRUNK, I don't say anything, but when a friend of mine gets $\overline{\text{DRUNK}}^{O_O}$, I sometimes comment on that.
- ✓ '...when a friend of mine gets very drunk...'.
- ✗ '...when (a friend of mine gets drunk and I am {surprised, impressed} by this fact)...'.
- b. ?When a friend of mine gets DRUNK, I don't say anything,
but when a friend of mine gets DRUNK^{O_O}, I sometimes comment on that.
- ✓ '...when a friend of mine gets very drunk...'.
- ✗ '...when (a friend of mine gets drunk and I am {surprised, impressed} by this fact)...'.

Note that it is not just that the second reading is odd for independent reasons. I myself used to think that this might be the reason why some facial expressions cannot be interpreted locally (as noted in Schlenker 2018b), the idea being that one should normally know what they find surprising, disgusting, etc., so it would be weird to question it or target it by question operators, epistemic modals, or *if*.⁶⁸ But, of course, this is not always the case, and, in particular, this is not the case in the examples above. Whether or not I'm surprised by the fact that a given friend of mine gets drunk depends on who that friend is, and it is conceivable that I only comment on my friends getting drunk when I find it surprising. In particular, there is nothing wrong with the paraphrases of this interpretation above. It is thus a plausible interpretation, yet it is unavailable.

Under Schlenker's typology, these differences in projection behavior between the two uses of *O_O* are entirely puzzling, as are the parallels with spoken adverbs. Under the composition-driven view that I have been pursuing here, they are not. The explanation is the same as for spoken

(v) When a friend of mine gets drunk, I usually don't say anything, but when a friend of mine gets drunk **unexpectedly**, I sometimes comment on that.

The effect doesn't carry over to facial expressions, however.

⁶⁸The general point that some facial expressions might not be able to be interpreted locally at all because of their semantics still stands; for example, it might be true about negative attitude facial expressions, if it turns out that they are indeed similar to spoken expressives.

adverbs. The sentence-level *O_O* is a supplement and cannot be interpreted locally. The degree modifier *O_O* is a modifier and can be either restricting or non-restricting. The secondary modality nature of co-speech or co-gesture modifier facial expressions can make them more likely to be non-restricting, but there can be other considerations that would override this preference. What would these considerations be? I am not sure at this point, but I think it might have something to do specifically with their *degree* modifier nature.

When developing his typology, Schlenker argues that iconic vowel lengthening, which is interpreted as a degree modifier, is by default at-issue, as shown in (4.33).

(4.33) If the talk is loooong, I'll leave before the end.

↗ If the talk is long, the speaker will leave before the end.

→ If the talk is very long, the speaker will leave before the end.⁶⁹

However, Schlenker argues that there is no default projection in (4.33) because vowel lengthening is an “internal enrichment” without a separate time slot. Under Schlenker’s typology, such “enrichments” can but don’t have to be at-issue by default; for example, some F0 modulations seem to be not-at-issue by default. Schlenker concludes that further research is needed to explain the differences between different types of vocal prosodic modulations.

Now, it is unclear to me in what sense vowel lengthening is “internal”. Schlenker’s point seems to be that vowels always have some length, therefore, changing this length is an “internal enrichment”. I am skeptical about this claim. It’s not like vowel length is always interpreted iconically. We could posit an abstract morpheme in the compositional structure that tells us to interpret vowel length iconically as a degree modifier. When the vowel length is different from the default, the addressee will deduce that the morpheme is present; when it’s not, there is no reason to do so, unless there is a contrastive alternative in the context with a non-default vowel length.

⁶⁹Schlenker’s original example (Schlenker 2018b, (13)) contained an additional context preceding the conditional, *I am normally rather patient. But....* I think the point can and should be made without this context, as it might be introducing the alternative with *long* (without lengthening) implicitly, suggesting that the speaker would be able to survive a long talk that is not very long.

In other words, it's not obvious to me that vowel lengthening should be treated differently from, say, repetition of a degree modifier, as in (4.34).

- (4.34) a. If the talk is very long, I'll leave before the end.
 ↗ If the talk is long, the speaker will leave before the end.
b. If the talk is very-very-very long, I'll leave before the end.
 ↗ If the talk is very long, the speaker will leave before the end.

It's unclear what Schlenker would say about cases like in (4.34). Degree modifiers are clearly optional (i.e., “syntactically eliminable”, or “external”), and so is their repetition. Yet, degree modifier repetition has a very similar iconic effect to vowel lengthening. Putting the two into different categories would be missing a generalization.

The data on degree modifier uses of *O_O* above suggest that there might be something about degree modifiers that makes them unlikely to be non-restricting by default, regardless of whether they are “internal” or “external”. Of course, an investigation is in order to see if that's actually the case, and if it is, an explanation would be in order why that is so. Furthermore, a question arises as to why the degree modifiers in LIS and Italian discussed in Aristodemo 2017 and mentioned previously in subsection 2.1.2.4 are different in that they do prefer to be non-restricting and thus do project by default. One obvious difference between the two, which might be relevant, is that *completely*-type modifiers (which Aristodemo discusses) are maximal-degree modifiers and apply to closed-scale predicates while *very*-type modifiers (which I discuss) are non-maximal-degree modifiers and apply to open-scale predicates. Also, specifically in the LIS case, the mere fact that there are no lexical alternatives without the ‘completely’ component might make the restricting interpretation unlikely. Finally, it would be informative to have systematic data on projection from various environments and possibility of local (i.e., restricting) interpretations both for attributive and predicative uses of predicates modified by the two types of degree modifiers. At this point I only have examples with predicative uses (the reason being that articulatory integration of prospeech gestures and facial expressions is easier in this case), and Aristodemo's examples are mixed.

Either way, this is an interesting question for future research.

To sum up, the differences in projection behavior between the two uses of *O_O* seem to be due to how it composes in the two cases, which is determined by where it merges in the syntactic structure. Furthermore, the way composition determines projection for mirative facial expressions is the same as for mirative spoken adverbs. This, once again, supports the general idea that we should treat content-bearing non-spoken content in the same way as spoken content when it comes to its compositional integration into utterances.

4.5 The role of convention

In reaction to the data on the *O_O* facial expression discussed in the previous section, Philippe Schlenker (p.c.) made the point that this facial expression might not be “iconic”, but rather “grammatical”. Schlenker’s (2018b) typology was originally meant to be about “iconic” expressions only, so one could argue, it does not apply to “grammatical” expressions.

Now, there is, of course, no direct a priori opposition of iconicity (i.e., similarity between the form of an expression and its meaning) and level of grammaticalization, or conventionalization.⁷⁰ For example, iconicity abounds in highly grammaticalized processes in sign language, such as anaphora or role shift (see, e.g., Schlenker 2018e for an overview). However, I take the spirit of Schlenker’s comment to be that his original typology can be taken to only apply to phenomena that are not conventionalized, but rely exclusively on iconic mechanisms for interpretation—as opposed to phenomena that might or might not involve iconicity, but are conventionalized. In this section I will argue against both the practicality of such a typology and the need for it.

The main practical argument against creating such a typology is that it will require one to

⁷⁰In fact, it’s not entirely clear to me what Schlenker means by “grammatical”. In his comment, he likens the *O_O* expression to raised eyebrows, which, for example, often mark focus in sign and spoken language as well as questions in at least some sign languages (e.g., polar questions in American Sign language, both polar and *wh*-questions in Russian Sign Language, etc.). First, the analogy escapes me. The two uses of the *O_O* expression parallel the two types of content identified by me, i.e., sentence-level mirative adverbs and degree modifiers, much more closely than the focus or question operators. Second, the difference between, say, a question operator and a modifier is precisely that they compose differently, not that one is “grammatical” and the other one isn’t.

take a principled stance on what's conventionalized and what's not. What would be the criteria for that? For example, is Schlenker's comment meant to extend to other facial expressions, such as negative attitude expressions in spoken and sign language, which Schlenker does integrate into his typology? The same question can be raised about some sign language processes discussed in Schlenker 2018b. For example, Schlenker's typology includes iconic modulations of verb signs, such as *GROW*, that indicate, for example, the magnitude or the speed of growing. He maintains that the typology of "iconic enrichments" does apply to such modulations and that they are by default at-issue, because they don't have their own time slot and are "internal". It's unclear, however, why treat these modulations as "enrichments" rather than gradient internal morphemes in the first place that contribute at-issue content because semantically they are ordinary manner modifiers, and there is no reason for them to prefer to be non-restricting, because they don't exist in a secondary modality with respect to the verb sign. Non-linear morphology exists even in spoken language and is a trademark property of sign language. The gradient nature of these modulations shouldn't prevent them from being treated as morphemes either; for example, many phenomena in sign language are gradient (e.g., various geometric properties of sign language pronouns), and yet have been treated as grammatical by various people including Schlenker himself (again, see Schlenker 2018e for an overview). In fact, the general view in Schlenker 2018e that iconicity is an integral part of many grammatical processes in both sign language and gesture seems at odds with the assumption that modulations of sign language verbs discussed in Schlenker 2018b are not grammatical. If they are, however, and if Schlenker's typology is indeed meant to only predict projection of non-conventionalized phenomena, these modulations shouldn't be part of it.

Besides, using conventionalization as a criterion for including a given phenomenon into the typology of projective content seems to rely on a tacit assumption that whether something is conventionalized or not is a categorical and rigid distinction. It's unclear to me that this assumption is justified, and it's also unclear to me how such a typology is meant to be applied if this assumption doesn't hold.

Now, all these issues (having to come up with criteria for what's conventionalized and

what's not, having to take a definitive stance on some of the less obvious cases, having to deal with the possibility that conventionalization is a gradient process) only arise if one believes that conventionalization status delineates a natural class of content with respect to projection and, thus, warrants creating a classification of non-conventionalized phenomena in regard to projection patterns to the exclusion of anything conventionalized. I don't think such a belief is justified.

For example, the broad intuition pursued in Schlenker's work that secondary modality content co-occurring with something in the primary modality ("co-something" content) tends to be truth-conditionally vacuous applies to clearly conventionalized co-speech gestures. Thus, in all examples in (4.35), the contribution of the gesture does seem to project by default, even though a non-projecting interpretation of the gesture would be perfectly reasonable (and can be obtained under pressure). If one were to exclude such gestures from consideration when developing a classification of projective content that includes non-conventionalized co-speech gestures, they would be missing a generalization.⁷¹



- (4.35) a. If we wanna  celebrate_{DRINK} my defense, we better go to a store now.

→ If we wanna celebrate my defense, we better go to a store now.

→ If we celebrate my defense, we'll do so by drinking alcohol.

(Cf. *If we wanna celebrate my defense by drinking...*)



- b. If you bring  a semanticist_{CRAZY} to my talk, I'll likely fight with them.

→ If you bring a semanticist to my talk, I'll likely fight with them.

→ ≈All semanticists are crazy.

(Cf. *If you bring a crazy semanticist to my talk...*)

⁷¹In fact, as mentioned in the previous section, Schlenker himself maintains that his typology is meant to apply to the conventionalized Italian gesture *COMPLETELY* discussed in Aristodemo 2017. I believe it would be hard to come up with a principled explanation of why one and the same typology would apply to conventionalized gestures but not conventionalized facial expressions.



- c. If I buy  a smartphone_{MONEY}, I won't be able to pay rent next month.

→ If I buy a smartphone, I won't be able to pay rent next month.

→ ≈All smartphones are expensive.

(Cf. *If I buy an expensive smartphone...*)

Now, in footnote 45, I alluded to the possibility that conventionalized gestures can have conventionalized morphosyntactic constraints, which makes sense, if they are stored in the lexicon just like spoken morphemes. For example, while the gesture *DRINK* above is fairly flexible in this respect (it seems to be a root that can combine with different category heads in different syntactic contexts, so it can become a verb, an adjective, or a noun), there are conventionalized gestures that have much more narrow semantics. For example, one such gesture is *FINGERS-CROSSED*, which seems to only function as an optative (i.e., indicating a wish or a hope) sentence-level adverb, akin to *hopefully*, but can't be used as an adjective akin to *hopeful*, as shown in (4.36).⁷² In this respect, it behaves exactly like its verbatim spoken counterpart, *fingers crossed*, and, unsurprisingly, projects in the exact same way.⁷³



- (4.36) a. If  Zoe wins the race_{FINGERS-CROSSED}, her mother will be proud.

→ The speaker wants Zoe to win the race.

- b. If Zoe, hopefully, wins the race, her mother will be proud.

→ The speaker wants Zoe to win the race.

- c. If Zoe wins the race—fingers crossed—her mother will be proud.

⁷²I thank Patrick Grosz (p.c.) for discussing this gesture with me.

⁷³Curiously enough, *FINGERS-CROSSED* seems to work both as a co-speech gesture and a prosodically independent gesture, following the clause it combines with—without any obvious consequences for projection. A similar gesture, *KNOCK-ON-WOOD*, seems to prefer being postposed. I suspect this might be partially due to the fact that the knocking gesture is harder to integrate prosodically with speech, but it might also be due to the specific lexical semantics of *KNOCK-ON-WOOD*, which is performed to prevent jinxing post factum—ironically or not—and, thus, requires an antecedent that could potentially cause said jinxing.

- The speaker wants Zoe to win the race.
- d. *A stuntwoman_{FINGERS-CROSSED} entered the room.
 - e. A hopeful stuntwoman entered the room.
 - f. *A—fingers crossed—stuntwoman entered the room.

Similarly, conventionalized gestures used in polar responses (e.g., head nods and head shakes in many cultures) also have very narrow semantics and are, thus, very constrained morphosyntactically (see Esipova 2019d for a recent in-depth discussion of such gestures). It's unclear how one would even ask the question of projection for such gestures, and the exact same unclarity arises for spoken polarity marking particles like *yes* and *no*. It's quite certain, however, that gestures in polar responses do not come with any assertion-dependent inferences akin to cosuppositions.

Thus, one would not want to maintain that co-speech *FINGERS-CROSSED* or head gestures used in polar responses generate cosuppositions, which is what applying Schlenker's typology to these cases would predict. If, however, we still wanted to generate cosuppositions for cases like in (4.35), we would have to find a principled way in which the two types of conventionalized gestures are different—a goal that parameters like having its own time slot, being “internal” vs. “external”, or being conventionalized vs. non-conventionalized can't possibly accomplish (in all these cases we are dealing with co-speech, “external”, conventionalized expressions).

Under the composition-based, modality-neutral view of projection developed in this dissertation, there is no dilemma. The conventionalized gestures in (4.35) can give rise to cosuppositions a.k.a non-restricting modifier inferences because they can be construed of as modifiers, but being co-speech, they will then prefer to be non-restricting modifiers.⁷⁴ Modifier construals are not possible for *FINGERS-CROSSED* or head nods/shakes due to their narrow conventionalized semantics, so no cosuppositions a.k.a non-restricting modifier inferences will ever be generated.

Level of conventionalization is, thus, not a relevant parameter for projection under the composition-driven view. What matters is how a given gesture composes, and gestures will be constrained in how they can compose (to a varied degree). Whether these constraints arise due to

⁷⁴Supplement-like construals are likely possible for all the cases in (4.35), too.

convention or due to iconic considerations is immaterial to how they will project, if at all.

If one discovers that, in fact, level of conventionalization does play a role in whether and how a given piece of content projects, further independent principles can be added to the system. For example, we might find out that level of conventionalization interacts with how we determine the hierarchy of modalities, which is relevant for whether a given piece of content will prefer to be truth-conditionally vacuous in the first place and which we still need to investigate in greater detail. Pending such discoveries, however, I do not see any need for integrating level of conventionalization into the system.

To sum up, the response to Schlenker's comment is that indeed, mirative facial expressions can be (and probably are) highly conventionalized, which is probably also the case for other facial expressions. That might be constraining their composition, which, in turn, will constrain their projection. However, there doesn't seem to be a direct link between level of conventionalization and projection, and a typology of projective content that is built on the assumption that such a link exists is both impractical and unnecessary.

4.6 Concluding remarks

The discussion throughout this chapter leads to the following conclusion: if we want to approach content-bearing secondary modality content as linguists, we should do so at all levels of representation. Most importantly, we cannot ask the question of how a given piece of content projects before we determine how it composes in the syntax/semantics.

We have seen throughout this chapter, based on data on hand gestures and facial expressions, that formulating the rules determining the exact projection pattern—in particular, cosupposition vs. supplement—in terms of linearization or being “internal” vs. “external” doesn’t work. Associating the two projection patterns with two composition strategies (modifiers and supplements), however, yields precisely the right result.

Now, the “co-something” status of a given piece of content (i.e., the status of being a sec-

ondary modality piece of content co-occurring with something in a more primary modality), whose role is emphasized in Schlenker's work, is still important for projection in that it can result in a preference for a truth-conditionally vacuous interpretation. However, how such an interpretation is assured—by generating a non-restricting modifier inference/cosupposition or via the supplement projection mechanism—is not determined by the “co-something” status of a given piece of content, but rather by how this piece of content composes.

I would like to finish with two methodological points that echo the main theoretical points made throughout this chapter. The empirical discussion in this chapter clearly shows that when checking if a certain content-bearing secondary modality expression (gestural, musical, pictorial, etc.) integrated into an utterance projects as a cosupposition or as a supplement, one should entertain various attachment possibilities and corresponding interpretations. Otherwise, if one's explicit or implicit goal is to support a typology based on temporal alignment for such content, they might find themselves subject to a confirmation bias whereby they only check the projection behavior of modifier construals for “co-something” expressions and supplement construals for expressions with their own time slot.

The second point is that when checking how something projects, one should not only check whether the resulting inference is conditional or not (supplements can give rise to conditional inferences), but also whether it can in principle be interpreted locally under semantic operators. We have seen over and over again in this chapter that inferences contributed by modifiers can in principle be interpreted locally (when those modifiers are restricting), but inferences contributed by supplements normally cannot. If a given piece of content gives rise to inferences that cannot be interpreted locally under one interpretation, but can under another, one should consider that this piece of content composes differently in the two cases, and, consequently, two different projection mechanisms are at play.

Phi-features on pronouns

Abstract

In this chapter I observe that *phi*-features on pronouns project similarly to non-restricting modifier inferences. I propose an analysis of internal composition of pronouns that assures that *phi*-features are obligatorily non-restricting modifiers, which means they always give rise to non-restricting modifier inferences. The proposed analysis is both more explanatory and morphosyntactically plausible than the standard view whereby *phi*-features on pronouns trigger presuppositions lexically. I also discuss the limited local context sensitivity of gender on pronouns. In particular, I adduce novel data from a mini-poll showing that this phenomenon is subject to more variation across and within speakers than previously believed, patterning with T–V features on pronouns in Russian, and sketch the direction for analyzing both gender and T–V on pronouns as FORM INDEXICALS, i.e., indexicals that are interpreted relative to indices that keep track of appropriate forms of expressions (thus, unifying and enriching some of the previous insights in the literature, such as Yanovich 2010 and Schlenker 2007). This treatment of gender and T–V on pronouns is independent of their non-restricting modifier status. In the final section of this chapter I propose to treat height specifications on verbal gestures as modifiers of incorporated nominal arguments, which give rise to non-restricting modifier inferences when they are non-restricting. This analysis explains the empirical parallels between height specifications on gestures and *phi*-features on pronouns observed in Schlenker & Chemla 2018.

5.1 Projection of *phi*-features

We have seen in (1.1b), repeated below as (5.1), that *phi*-features on pronouns by default contribute projecting inferences:

- (5.1) If Skyler_i brings her_i dog, I'll give you \$10.
→ Skyler is female.

The inferences contributed by *phi*-features on pronouns project very strongly, i.e., they cannot be interpreted locally, not even under severe pressure, as shown in (5.2) below. In (5.2a), the inference contributed by the gender feature on *her* is allowed to project, so the sentence is felicitous. In (5.2b), the gender inference can't project because of the context, so it has to be interpreted locally, but for some reason this option is not available for *phi*-features, so the sentence is infelicitous.

- (5.2) *Context: In a magical universe, the speaker is inspecting a victim of a curse that is especially powerful when a woman casts it. ...*

- a. *...The speaker thinks that either Lucius or Bellatrix cast the curse.*

I don't know if it was Lucius or Bellatrix_i who cast the curse, but if **she**_i cast it, the victim is unlikely to recover.

→ The referent of *she* is female.

- b. *...The speaker has no suspects.*

#I don't know the gender of [the person who cast the curse]_i, but if **she**_i cast it, the victim is unlikely to recover.

↗ The referent of *she* is female.

Intended: ‘...if (x_i is female and she_i cast it)...’

The contrast in (5.2) is intuitively similar to the contrast we have seen for non-restricting modifiers. In (5.2a), there are two potential referents, and the QUD is about which of the referents is the culprit. The two referents differ along a certain dimension, namely, gender, but the gender on the

pronoun only helps you identify the referent of *she* (in the absence of overt indices in spoken language); the content of the gender feature does not address the QUD and is thus free to project. This is similar to ‘Projecting non-restricting’ scenarios from Chapter 3, where we would also have two potential referents (e.g., Stephanie’s only cat and her only dog) that differ along some dimension (e.g., size), which is in some way relevant in the context, but does not address the QUD and is thus free to project.

In (5.2b), however, we only have one referent with an unknown characteristic along the gender dimension, and we are trying to use the gender feature on *her* to locally establish this characteristic, but fail. Something similar is happening in ‘Non-projecting non-restricting’ scenarios from Chapter 3, where we would also have one referent (e.g., Stephanie’s only dog) with an unknown characteristic (e.g., size), which cannot be established locally via a modifier.

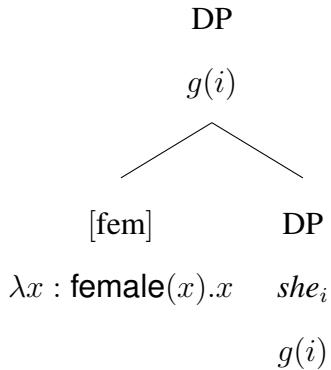
In section 5.3, I am going to argue that this similarity obtains because *phi*-features on pronouns are in fact non-restricting modifiers, but the non-restricting interpretation is obligatory for them due to how denotations of pronouns are built. But before that, let’s take a look at the standard analysis of *phi*-features on pronouns as lexical presuppositions.

5.2 Lexical presupposition analyses of *phi*-features

Projecting inferences of *phi*-features on pronouns have been traditionally analyzed as lexical presuppositions. Below I review two representative analyses within this approach.

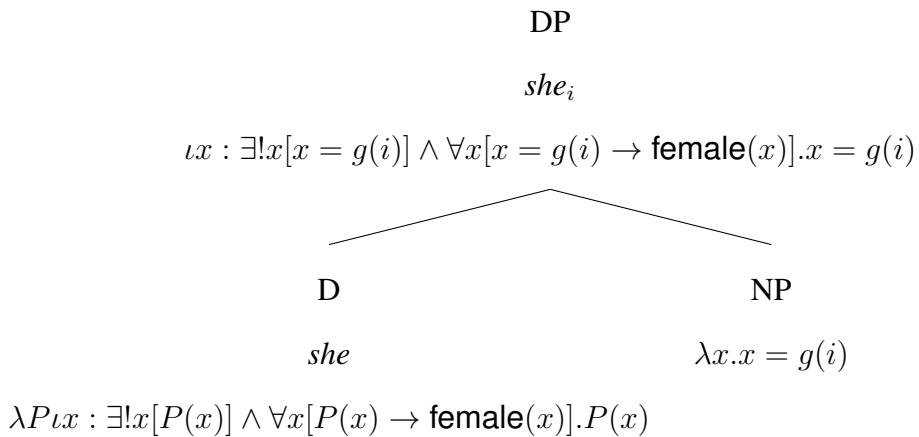
In Heim & Kratzer 1998 (also, e.g., Cooper 1983; Sudo 2012), the pronoun starts out as an index variable of type *e* whose value is supplied by the assignment function *g*. *Phi*-features then merge one by one and check their respective presuppositions about the value of this variable. If its presupposition is satisfied, the feature passes this value on; if it isn’t, the feature returns a failure. In (5.3) I provide Heim & Kratzer’s structure for *she* (for simplicity, I only include the gender feature; their treatment of number and person is analogous).

(5.3) **Pronouns in Heim & Kratzer 1998**



Elbourne (2005) treats pronouns as definite descriptions, which have an *et*-type NP layer. Not all pronouns bear indices in Elbourne's analysis, but in those that do, the NP denotes the property of being the value of the pronoun's index.⁷⁵ This NP combines with a determiner that is similar to the definite article, but comes with additional presuppositions contributed by the *phi*-features. Elbourne further assumes that in English pronouns, the NP doesn't have an overt exponent, and morphemes like *she* spell out the determiner. Elbourne's structure for an indexed *she* is given in (5.4) (again, I only include gender).

(5.4) **Indexed pronouns in Elbourne 2005**



There are several issues with the two analyses of *phi*-features on pronouns above.

First, if *phi*-features are lexical presupposition triggers, they have to be strong, in the sense

⁷⁵I am simplifying Elbourne's treatment of indices in a way that is not crucial for anything here.

that the inferences they trigger cannot be locally accommodated, even under maximal pressure, as shown in (5.2b). As mentioned before in subsection 2.1.2.3, the weak–strong distinction is a descriptive, not an explanatory one. Much of the presupposition literature explicitly or implicitly assumes that why some expressions trigger presuppositions in the first place, as well as how strong those presuppositions are, is determined lexically. The literature that does try to come up with a principled triggering algorithm (see Abrusán 2011 for an overview as well as one of the proposals) typically targets weak triggers (usually by design). Strong triggers thus remain a mystery. Either one needs to explain why a given triggering algorithm cannot be undone for such triggers, or we should just give up on non-lexical triggering algorithms for strong triggers and assume that their presuppositions are purely lexical. The latter option is particularly vexing in the case of *phi*-features, since it is entirely unclear then why languages systematically divide the lexical content of pronouns into at-issue and presuppositional in the same way, i.e., with the index being at-issue but the featural content being presupposed.

Second, the composition strategy used by the feature in (5.3) is an unusual one. Do we have adnominal content other than *phi*-features that only checks presuppositions? If we don't, then, all other things being equal, an analysis that doesn't have to posit a new composition strategy to handle *phi*-features will have an edge over an analysis that does.

Elbourne's analysis doesn't run into this issue, but only because he doesn't put *phi*-features into separate nodes in the syntax in the first place, instead lumping them all in D. This is incompatible with any framework that builds words in the syntax, such as Distributed Morphology (Halle & Marantz 1993 et seq.) or Nanosyntax (see Starke 2010 for a brief overview), given that in many languages pronouns are morphologically complex. Heim & Kratzer's analysis does posit separate nodes for *phi*-features, but it is morphosyntactically implausible for a different reason.⁷⁶ There have been arguments in the literature that pronouns do need an $\langle e, (s)t \rangle$ -type layer going back as early as Postal 1966. Postal adduces examples like *we linguists* to argue that English pronouns can be modified. However, one could argue that in examples like this *linguists* is actually a prosodically

⁷⁶It's also unlikely that *phi*-features are DP-level adjuncts, which is also why I am not proposing to treat them as supplements. But the labels in (5.3) are not a crucial part of Heim & Kratzer's analysis.

integrated appositive, as it doesn't interrupt the pronoun.

Some recent striking evidence that pronouns can be modified internally comes from Lee’s (2019) work on Khoekhoe pronouns, which can be modified by $\langle e, (s)t \rangle$ -type properties, including compositionally complex ones, wedging between the morpheme realizing person and those realizing gender and number, as shown in (5.5). Note also that the data like in (5.5) would be hard to reconcile with an Elbourne-style view whereby all featural content of a pronoun is in D.

There is, of course, an important question of whether these prounoun-internal modifiers in Khoekhoe also behave as non-restricting modifiers, and in ongoing joint work Lee and myself have confirmed that indeed they do. First, these modifiers can never be used to restrict a set of potential referents. For example, (5.6) is OK if uttered while pointing at ten women all of whom are doctors, but bad if only some of these women are doctors and the speaker is trying to use *doctor* to restrict the set.

Furthermore, the inferences contributed by these pronoun-internal modifiers have to project, just like the inferences of *phi*-features, as shown in (5.7), which is OK in a context in which the speaker knows that the woman they are talking about is a doctor, but bad in a context in which they don't know if she is.

- (5.7) ||?í-||ää-tí-?áo-s ká sääó té ?ós(ä) ké ||?í-s-ää ní hüí
 3-heal-?-person-F.SG follow if/SEMLEFACT me then DECL 3-F.SG-OBL FUT help

⁷⁷Namas are the largest Khoekhoe-speaking ethnic group.

- ‘If she doctor follows me, she will help.’ (Khoekhoe)
- ✓ ‘She is a doctor, and if she doctor follows me, she’ll help.’
- ✗ ‘If (she is a doctor and she follows me), she’ll help.’

To sum up, a morphosyntactically plausible structure for pronouns is one where (i) pronouns have an $\langle e, (s)t \rangle$ -type layer, and (ii) *phi*-features have their own nodes within this layer.

5.3 Proposal: *phi*-features as obligatorily non-restricting modifiers

I propose that we can avoid all the issues above and have a conceptually appealing and morphosyntactically plausible analysis of *phi*-features on pronouns by treating them as modifiers that are, furthermore, obligatorily non-restricting. The reason why *phi*-features are always non-restricting is that they always modify the Elbourne-style property of being the value of the pronoun’s index, whose extension is always a singleton set (containing one individual, atomic or plural).⁷⁸ As established before in subsection 2.1.2.3, if a modifier instance modifies an expression whose extension is a singleton set, it can’t be restricting and, thus, has to be non-restricting.

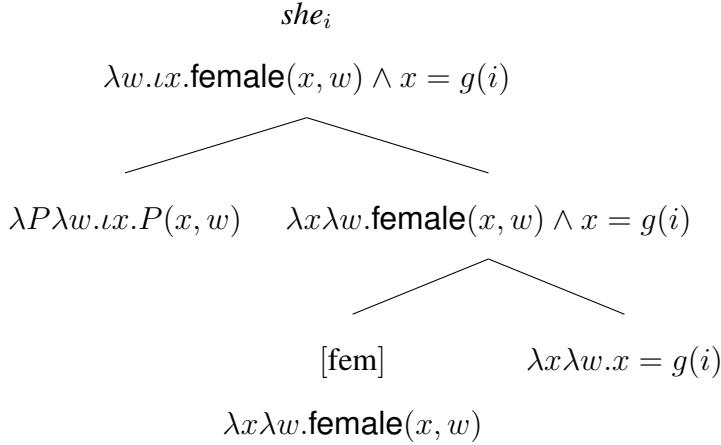
After all the *phi*-features have merged, a determiner applies to the resulting set and returns the individual in it, which is the value of the index. Thus, the internal composition of pronouns resembles that of other nominals much more closely: they start out as $\langle e, (s)t \rangle$ -type properties, can be modified as such,⁷⁹ and finally combine with a determiner yielding an individual.

This composition is exemplified in (5.8) for the English pronoun *she* (once again, for now I am only showing how things work for gender).

⁷⁸Once again, for Elbourne (2005), not all pronouns bear indices. His motivation for having pronouns that are indexless definite descriptions is to handle donkey anaphora. My motivation for treating pronouns as definite descriptions is completely independent of donkey anaphora-related considerations, and I assume that all pronouns bear indices.

⁷⁹Language-specific restrictions on how pronouns can be modified could come from which lexical items a given language has to spell out which chunks of structure. For example, if English *she* can only spell out an uninterrupted structure consisting of the property of being the index value and all the relevant *phi*-features, no additional modifiers will be allowed, since there will be no way to spell out the resulting structure.

(5.8) **Pronouns in the proposed system (incomplete)**



In (5.8), I am assuming intersective composition of the feature and the property it merges with for simplicity. I'm also agnostic about the labels and the morphosemantic mapping within a pronoun, i.e., which morpheme spells out which part of meaning. Finally, the denotation I give to the [fem] feature in (5.8) is likely a simplification and would need to be refined for other cases. For example, in many languages the [masc] feature would need to impose a weaker requirement that at least one of the atoms of x is male. That would be needed to capture the uses of masculine plural pronouns to refer to a mixed-gender group (e.g., French *ils* ‘they.MASC’). Similarly, adjustments would need to be made for various gender-neutral uses of feminine and masculine pronouns. Those adjustments, however, are needed under any analysis of gender features, and they are relatively easy to make without affecting the main point of my proposal, so I won't explore this issue further. I will, however, discuss the more general issue of the limited local context sensitivity often exhibited by gender on pronouns and how it can be captured by changing the lexical semantics of gender features in the next section.

As any other non-restricting modifiers, *phi*-features contribute non-restricting modifier inferences; in (5.8), the inference would be $\forall x \forall w [\llbracket c'' \rrbracket(w) \rightarrow (\lambda x. x = g(i) \rightarrow \mathbf{female}(x, w))]$. As things stand, this inference projects relative to a local context c'' . It is unclear, though, how sensitive gender inferences actually are to local contexts, and I come back to this issue in the next section.

Note that there is an additional difference between my analysis and Elbourne's. Elbourne

explicitly adds presuppositions of existence and uniqueness to the denotation of the pronoun determiner. He does that, because for him not all pronouns bear indices, and the property the determiner combines with can be some contextually salient property. In those cases, he has to rely on the existence presupposition of the determiner to assure projection of the inferences contributed by *phi*-features. For me, all pronouns bear indices and, thus, all *phi*-features are non-restricting modifiers and project as such.⁸⁰ Making the projection of *phi*-feature inferences an epiphenomenon of the projection of a lexical presupposition of existence would be empirically problematic, since existence inferences of definites are not as strong as *phi*-feature inferences—or other non-restricting modifier inferences, for that matter, as shown before in (2.32). If *phi*-feature inferences projected as part of lexical existence presuppositions, it's unclear why local accommodation of said existence presuppositions couldn't save the day in cases like (5.2b). The existence and/or uniqueness presuppositions can still be part of the *l* operator; they just aren't doing any work for me.

Now, what about number and person? The nature of the content contributed by person features is such that the empirical parallels with non-restricting modifiers are not as easy to make as for gender (i.e., it is hard to imagine in context in which the QUD would be about who is the speaker or addressee of the utterance being produced). The contrast in (5.2), however, can be replicated for number, which is what I do in (5.9).⁸¹

- (5.9) *Context: In a magical universe, the speaker is inspecting a victim of a curse that is powerful when cast by a single person, but its effect is weakened if two or more people jointly cast it. ...*

⁸⁰Naturally, I don't need a separate uniqueness requirement either, since the uniqueness of the output individual is assured by it being identical to the value of the index. I, furthermore, assume that the (local) existence of the individual identical to the value of the index is independently assured by the mechanics of the assignment function.

⁸¹The point is a bit harder to make for the plural feature, since if the plural is inclusive (i.e., it is compatible with the referent being singular), the inference about the referent being plural arises, when it does, via competition with the singular. While scalar implicatures can be at-issue, especially under contrastive focus, as shown below, it is unclear how systematic it is. I thank Philippe Schlenker (p.c.) for urging me to think about this issue.

- (vi) If you read **some** of the assigned books, this won't be enough to pass.
 ≈ 'If you read some but not all of the assigned books...'

- a. ...*The speaker thinks that either Bellatrix (standing on the left of the speaker) cast the curse alone or the curse was cast jointly by the siblings Alecto and Amycus (standing on the right of the speaker).*

I don't know who cast the curse, but if **she**_{POINT-LEFT} cast it, the victim is unlikely to recover.

→ The referent of *she* is singular.

- b. ...*The speaker has no suspects, but they know that only women can cast this curse.*

#I don't know how many women cast the curse, but if **she** cast it, the victim is unlikely to recover.

↗ The referent of *she* is singular.

Intended interpretation: ‘...if (x_i is an atomic individual and she_i cast it)...’⁸²

In the proposed system, person and number are meant to work in the same way as gender, as shown in (5.10) (\leq stands for ‘is a mereological part of’; I omit worlds for simplicity); the non-restricting modifier inferences are computed for each modifier ($X_1 \Rightarrow X_2$, $X_2 \Rightarrow X_3$, and $X_3 \Rightarrow X_4$; each of these expressions would need to be entailed by the local context c''). The details might need to be refined (for example, one might want to split person into [\pm speaker] and [\pm addressee] or alter the lexical entries for some features), but I believe (5.10) to be a viable rough sketch. The flexible nature of the proposed compositional structure is compatible with a wide variety of specific morphosyntactic analyses of pronouns, as long as the semantic types work out.

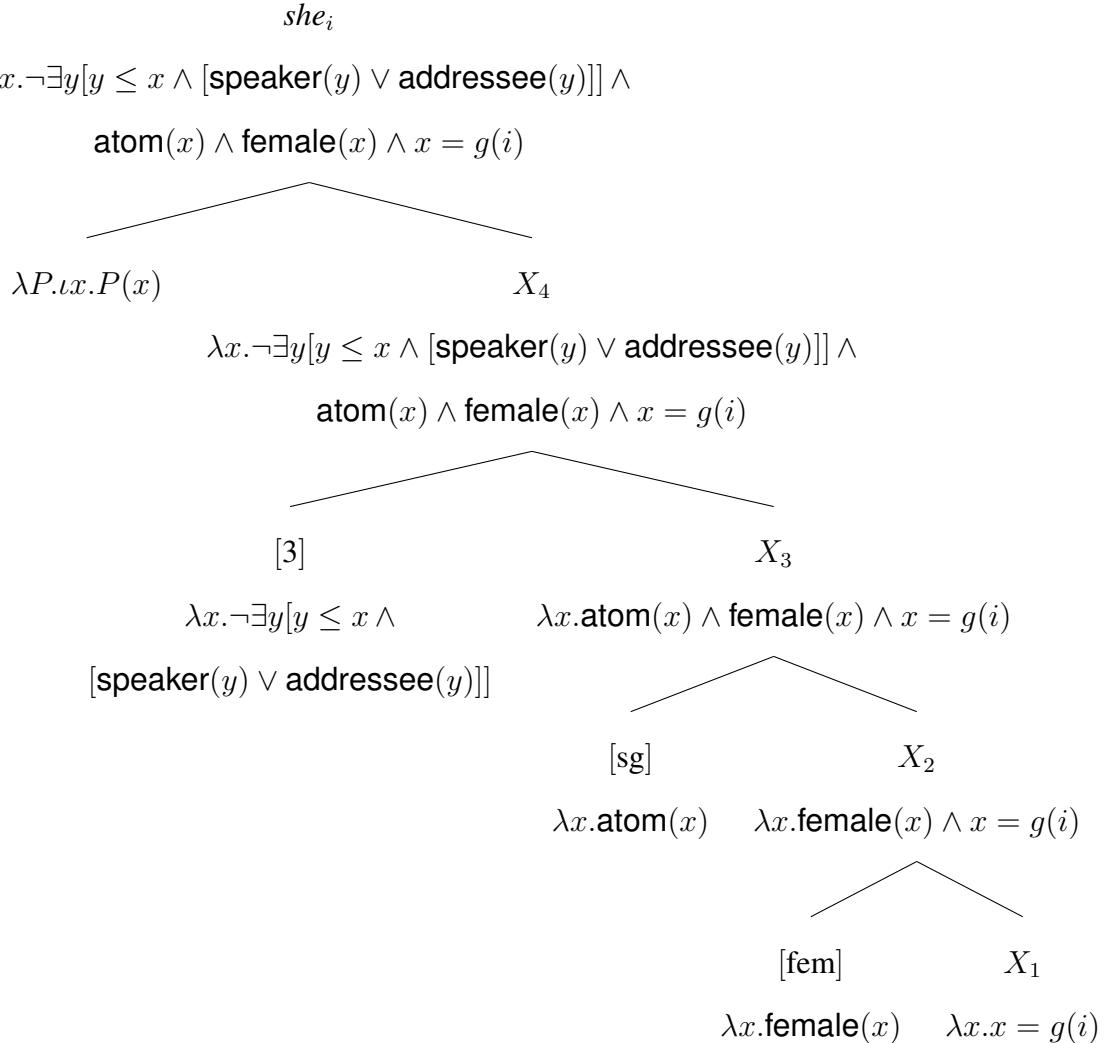
5.4 Gender, local contexts, and indexicals

If the inferences that *phi*-features on pronouns give rise to are non-restricting modifier inferences, they should be sensitive to local contexts. However, as I mentioned in the previous section, things are not as clear-cut for gender inferences, in a way that goes beyond the proviso problem.

In (5.11), the presupposition of *too* is not satisfied globally, but because it is entailed by the

⁸²Cf. *I don't know how many people cast the curse, but if there was only one of them...*

(5.10) **Pronouns in the proposed system (complete)**



local context created by the antecedent of the conditional, the use of *too* is felicitous. This is true both for counterfactual conditionals (the “counterfactual scenario”) and for indicative ones (the “ignorance scenario”).

- (5.11) a. If Mia was in the library, Lea would be there, too. *counterfactual scenario*
 b. If Mia is in the library, Lea is there, too. *ignorance scenario*

Yanovich (2010) claimed that gender on pronouns is insensitive to counterfactual local contexts, i.e., the form corresponding to the actual gender of the referent should be used in examples like

(5.12a), but in ignorance scenarios, as in (5.12b), the form justified by the local context can—and, in fact, has to—be used.

- (5.12) a. *Context: There is a kid named Sasha, and I know that Sasha is a girl.*

I will buy Sasha a toy train. If Sasha were a boy, I would buy {#him, her} a doll.

(Yanovich 2010, (13))

- b. I am at the end of my wits. If Sasha is a boy, I should buy him a doll. But if Sasha is a girl, I'd rather buy her a toy car. (Yanovich 2010, (11))

Sudo (2012) claimed that examples like (5.13b) are better than (5.13a) (the latter is meant to replicate Yanovich's example in (5.12a)).

- (5.13) a. #If John were female, she would be popular among the boys. (Sudo 2012, (49))
b. ?If John were a beautiful girl, all the boys would ask her out. (Sudo 2012, (50))

Yanovich proposed to treat gender on pronouns as indexical presuppositions, refining the original proposal in Cooper 1983 and outlining the directions for technical implementation of this idea.

I believe that Yanovich's insight is ultimately correct, but its implementation can benefit from (i) exploring variation across and within speakers, and (ii) comparing gender on pronouns to ordinary indexicals as well as other expressions that can be viewed as indexical, but also come with a socially relevant dimension. In this section, I do both and observe that while at least some ordinary indexicals categorically follow the pattern in (5.12), gender on pronouns exhibits much more variation, especially in ignorance scenarios and especially if the gender-neutral *they* is added as an option, patterning with T–V pronoun features in Russian. I propose to analyze gender and T–V as a special kind of indexicals (FORM INDEXICALS), and, while I don't provide a detailed technical analysis, I outline a way to capture this insight formally that leaves room for variation.

Now, why would treating an expression as indexical allow us to limit its local context sensitivity? Indexicals are evaluated with respect to a context parameter c_p ⁸³ on the evaluation

⁸³I use the p subscript to distinguish between c as a local context computed for presupposition projection and c_p as

function that stores the speaker, the addressee, the time and place of the utterance, etc. (see, e.g., Schlenker 2018d for an overview). Certain environments can introduce a context parameter c'_p different from the actual one c_p^* , and some indexicals can shift and be evaluated with respect to c'_p . This process can be kept independent of local context computation for presupposition projection, i.e., local contexts don't have to introduce c'_p , and even if they do, indexicals don't have to shift.

How do ordinary indexicals fare empirically in counterfactual and ignorance scenarios from above? In English, if the actual value of an ordinary indexical such as *I* or *here* is known, you can't use this indexical to refer to a counterfactual speaker or location. Thus, in (5.14a), *I* in the consequent cannot refer to Lea, and in (5.14b), *here* in the consequent cannot refer to St. Petersburg.

- (5.14) a. If Lea was the speaker of this utterance, {#I, she} would be speaking in a louder voice.
 b. If we were in St. Petersburg now, we would be having fun {#here, there}.

It would be hard, if not impossible, to come up with relevant ignorance scenarios for *I* (we typically expect the speaker of the utterance to know they are uttering it), and *you* comes with additional complications,⁸⁴ so I will not attempt constructing relevant examples for these indexicals. However, we can try doing it for *here*. In (5.15), *here* can be used to refer to St. Petersburg, and it would be weird to use *there* in this case.

- (5.15) I don't know if we're in Moscow or in St. Petersburg now, but if we're in St. Petersburg, we'll have a lot of fun {here, #there}.

In other words, (5.14b) matches the pattern reported for gender in Yanovich 2010. However, loca-

a context parameter for indexical interpretation.

⁸⁴Such as biscuit conditional interpretations of examples like the one below (said, for instance, over the phone), in which case the antecedent can be viewed as a condition on producing a certain speech act, which can behave more like a quotation.

- (vii) If I am talking to Jackie right now, you owe me \$10.

tive indexicals don't signal a (relatively) stable and socially relevant property such as gender, and, furthermore, English has no neutral or default locative pro-form (the latter fact might be a consequence of the former).

I find that the empirical picture for gender is actually more complicated, however, and people vary a lot in what forms they allow in examples like (5.12), especially in ignorance scenarios. I believe that a more controlled experimental study is needed to establish which judgement patterns are robust and what the magnitude of variation is. In the meantime, in Table 5.1, I report the results of an informal mini-poll for the list of examples in (5.16). I added (5.16e) as (i) a non-contrastive ignorance example, (ii) an ignorance example where the individual in question is presumably an adult. The latter was done, because I had an a priori hunch that some speakers might disprefer using gender-neutral *they* to refer to children; as can be seen from Sp3's comment, they shared this intuition. There is a further difference between (5.16b) and (5.16e) (as well as between (5.16d) and (5.16c)), to which I will come back, in that the former has gendered indefinites while the latter doesn't. Ten native speakers of English were asked to list all the forms they in principle accept for each of the examples, indicating any preferences; some of them also provided additional comments, which I report verbatim (with occasional notational changes).

- (5.16) a. *Context: Sasha is a girl.*

I will buy Sasha a toy train. If Sasha were a boy, I would buy {him, her, them} a doll.

- b. *Context: Sasha's gender is unknown.*

I am at the end of my wits. If Sasha is a boy, I should buy {him, her, them} a doll.
But if Sasha is a girl, I'd rather buy {her, him, them} a toy car.

- c. *Context: John is male.*

If John were female, {she, he, they} would be popular among the boys.

- d. *Context: John is male.*

If John were a beautiful girl, boys would ask {her, him, them} out.

e. *Context: Skyler's gender is unknown.*

If Skyler is female, we should invite {her, him, them} to join our ‘Women in Formal Semantics’ discussion group.

Table 5.1: Results of a mini-poll on local context sensitivity of gender on pronouns. Accepted forms and preferences for the list of examples in (5.16).

	(5.16a)	(5.16b)	(5.16c)	(5.16d)	(5.16e)
Sp1	her	$\langle \text{them, them} \rangle > \langle \text{him, her} \rangle$	he	him	them > her
Sp2	her > them	$\langle \text{him, her} \rangle = \langle \text{them, them} \rangle$	he > they	him > them	her = them
Sp3	her	$\langle \text{him, her} \rangle = \langle \text{them, them} \rangle$	he	him	her = them
Re (5.16b) vs. (5.16e): ‘...they [is] more natural in (5.16e) than in (5.16b). The use of <i>they</i> when you’re ignorant of someone’s gender is most natural for me with adults. Maybe because children are insufficiently animate to me.					
<i>The {??baby, person} was carrying something, but I couldn’t see what was in their hands.</i>					
Sp4	her	$\langle \text{them, them} \rangle$	he	him	her = them
Sp5	her	$\langle \text{them, them} \rangle$	he	him	them
Sp6	him = her = them	$\langle \text{him, her} \rangle = \langle \text{them, them} \rangle$	he > (she = they)	him > her > them	her = them
Re <i>she</i> in (5.16c): ‘It’s not impossible, but it feels playful/degraded.’					
Re <i>they</i> in (5.16c): ‘I’m maybe feeling the combined markedness of <i>they</i> for someone whose gender is known, and the fact that gender is being discussed, confusing the intention of using <i>they</i> .’					
Re <i>her</i> in (5.16d): ‘Better than <i>she</i> in (5.16c).’					
Sp7	her	$\langle \text{them, them} \rangle$	he	him	them
Sp8	her	$\langle \text{them, them} \rangle$	he	him	them
‘I would expect to hear (though not necessarily personally accept) a wider range than this if I were back home; (5.16b), for instance, could be $\langle \text{him, her} \rangle$, and (5.16c) and (5.16d) could be <i>her</i> (though slightly less easily).’					
Sp9	her	$\langle \text{him, her} \rangle > \langle \text{them, them} \rangle$	he	him	her
Sp10	her	$\langle \text{him, her} \rangle > \langle \text{them, them} \rangle$	he	him	her

Overall, there seems to be a robust contrast between the counterfactual examples ((5.16a), (5.16c), and (5.16d)) and the ignorance examples ((5.16b) and (5.16e)) in the amount of variation across and within speakers. In counterfactual examples, eight speakers only allowed the form corresponding to the actual gender of the referent in all the three examples; one speaker also allowed the gender-

neutral *they* (but preferred it less); yet another speaker allowed all the three forms, with further preferences depending on the specific example (the contrast between (5.16c) and (5.16d) reported in Sudo 2012 was only replicated for this speaker). In ignorance examples, the variation was more pronounced: three speakers only allowed *they* in both examples; three speakers equally allowed both the gendered form supplied by the local context and *they* in both examples (but see Sp3's comment on using *they* to refer to children); two speakers preferred the gendered forms to *they* in (5.16b) and only allowed the gendered form in (5.16e); one speaker allowed both, but preferred *they* over the gendered forms; one speaker only allowed *they* in (5.16b), but equally allowed both *they* and the gendered form in (5.16e).

Now, all the polled speakers are PhD students in linguistics and are of roughly the same generation (all born between late 1980s and mid 1990s). They do differ in gender and sexual identity, race/ethnicity, place of origin, what other languages they speak, etc. I do not report any such specific details to better preserve their anonymity; since this is a mini-poll, no further conclusions could be drawn by including such details anyway. I would not be surprised if polling speakers of more varied ages and social backgrounds revealed even more variation (which is also what Sp8 seems to be alluding to in their comment). However, this poll shows that variation emerges even within a small and relatively socially homogeneous group. It is also possible, of course, that this is exactly the generation that is caught in the midst of a language change and, thus, exhibits the most variation.

T–V features on pronouns in Russian are like gender on pronouns in that if the appropriate form is known, counterfactual local contexts typically can't override it. For example, in (5.17), the non-actual forms are inappropriate for the vast majority of speakers.

(5.17) *Context: The speaker is talking to a single person.* (Russian)

- a. Esli by my s vami byli na ty, ja by {#tebja, vas} nazyvala
if SUBJ we with you.PL were on you.SG I SUBJ you.SG you.PL called
Maša.
Masha

'If you and I were on the T form basis, I would be calling you Masha.'

- b. Esli by my s toboj byli na vy, ja by {tebja, #vas} nazyvala
 if SUBJ we with you.SG were on you.PL I SUBJ you.SG you.PL called
 Marija.
 Maria

'If you and I were on the T form basis, I would be calling you Maria.'⁸⁵

However, if the appropriate form is not known, more variation emerges. Thus, in (5.18a), most speakers I have polled so far prefer the—more marked—T form supported by the local context, but usually still allow the—less marked—V form to a varied extent (for one speaker the preference for the T form is categorical). A minority, myself included, prefer the V form.⁸⁶ When the local context supports the V form in an ignorance example, however, it has to be used, as shown in (5.18b).

- (5.18) *Context: The speaker is talking to a single person.* (Russian)

- a. Ja ne pomnju, na ty my ili na vy, no esli na ty, ja budu nazyvat'
 I not remember on you.SG we or on you.PL but if on you.SG I will call
 %{tebja, vas} Maša.
 you.SG you.PL Masha
- 'I don't remember if we are on the T or on the V form basis, but if we are on the T form basis, I will call you Masha.'
- b. Ja ne pomnju, na ty my ili na vy, no esli na vy, ja budu nazyvat'
 I not remember on you.SG we or on you.PL but if on you.PL I will call
 #{tebja, vas} Marija.
 you.SG you.PL Maria
- 'I don't remember if we are on the T or on the V form basis, but if we are on the V form basis, I will call you Maria.'

Of course, it would be interesting to see what happens in other languages with the T–V distinction—

⁸⁵To clarify, lest anyone draw any incorrect conclusions from these examples about how Russians use full names vs. hypocoristics: at least in my experience, it is perfectly fine to call someone Masha while using the V form to address them. Calling someone Maria while using the T form is a bit more weird for me (but there is a song by Boris Grebenshchikov where he does exactly that).

⁸⁶For me the preference is categorical; I cannot use the T form in (5.18a)—unless to performatively establish that we will be on the T form basis (which I am unlikely to do). This judgement about the performative nature of the T form use in (5.18a) is shared by at least one more speaker who prefers the V form in this case.

and more generally, with regularized honorific systems—in particular, in cases when none of the forms is perceived as unmarked/less marked (if such cases exist). So far I have only checked the relevant paradigm with one French speaker, and they pattern with the majority of Russian speakers, i.e., in counterfactual scenarios, they strongly prefer the actual form (although they were able to imagine some scenarios in which the counterfactual form would be marginally OK), and in ignorance scenarios, they allow both forms when the local context supports the T form (with a preference for the T form) and the V form only when the local context supports the V form. Of course, more cross-linguistic work might—and probably will—reveal more variation.

Schlenker (2007) treats T–V features—as well as slurs—as indexical presuppositions, but doesn’t discuss their local context sensitivity, nor does he draw parallels with gender on pronouns. Cepollaro (2017) does draw parallels between gender on pronouns and slurs regarding local context sensitivity, using examples similar to (5.19) to show that a counterfactual local context doesn’t “save” the speaker from signaling their prejudice by using the slur (ignorance scenarios would be hard to construct for slurs, as speakers are usually opinionated on whether or not they are biased against a certain group). However, she doesn’t make a connection to indexicals.⁸⁷

- (5.19) *Context: Mudblood is a term used in the ‘Harry Potter’ universe as a slur to refer to witches and wizards born to muggles, i.e., non-magical folks. A neutral term for magical folks born to non-magical folks is muggleborn.*

If I was prejudiced against muggleborns, I would have cursed that mudblood.

→ The speaker is prejudiced against muggleborns.

⁸⁷Cepollaro assumes that gender features on pronouns are presuppositional and takes the fact that gender features on pronouns are often insensitive to local contexts to be an independent piece of evidence that not all presuppositions are sensitive to local contexts, which she uses to ultimately argue for a presuppositional analysis of slurs. I don’t find this reasoning convincing. The fact that gender features on pronouns and slurs exhibit similar insensitivity to local contexts only shows that there is something we need to explain in both cases (which might be due to a common source), not that both are presuppositional—especially, if one defines presuppositions as inferences projecting relative to local contexts. In fact, the question of whether a certain inference is a presupposition becomes meaningless if we treat “presuppositions” as a heterogeneous class without a single defining feature that includes inferences varying along a range of dimensions (triggering mechanism, availability of local interpretations, sensitivity to local contexts, etc.). Even if there is value in having such a broad class, we still need to explain the varied behavior of the inferences within this class along these dimensions.

As said before, Yanovich (2010) does propose to treat gender on pronouns as indexical, but simplifies the empirical picture. Consequently, his attempts at implementing his theoretical insight rely on the assumption that gender features check one's gender in a world index, which doesn't seem very well-suited for capturing the more complex empirical picture in Table 5.1 and, in particular, the distribution of the gender-neutral *they*. Furthermore, equating gender on pronouns with one's gender does not allow us to capture cases in which there is no direct one-to-one correspondence between one's gender identity and their pronouns, for example, if several pronoun forms can be used to refer to the same person without them having multiple gender identities, if a person has pronoun forms other than $\{he, she, they\}$, etc.

I propose that in order to be able to handle this more complex empirical picture, we need to separate one's gender from the third person pronoun forms that can be used to refer to them—just like we should think of the T–V features as signaling a convention within a pair of individuals about which forms they use to address each other⁸⁸ rather than any independent properties of the individuals themselves or their relationship. While T–V conventions are established based on certain social norms that can involve hierarchical status, age, level of familiarity, etc., there is no direct one-to-one correspondence between a given form and a single property of a pair of individuals or their relationship, and there is ample variation in how these conventions are established and used. I propose that we can think of gender on pronouns in a similar way.

We can thus build on the above mentioned insights in the previous literature, but we need to enrich them. More concretely, I propose that we can treat both gender and T–V on pronouns as FORM INDEXICALS, associated with their own indices within c_p : c_g keeps track of the preferred third person pronoun forms of individuals based on their gender identity (or lack thereof), and c_{TV} —of the appropriate second person pronoun forms for (ordered) pairs of individuals. For example, as a first pass, the denotations for [she-form] and [T-form] could be sketched as in (5.20).

- (5.20) a. $\llbracket [\text{she-form}] \rrbracket^{c_p} = \lambda x \lambda w. x \text{ has } she \text{ as the preferred pronoun form in } c_g \text{ of } c_p$
- b. $\llbracket [\text{T-form}] \rrbracket^{c_p} = \lambda x \lambda w. x \text{ is addressed as } ty \text{ by the speaker } c_s \text{ of } c_p \text{ in } c_{TV} \text{ of } c_p$

⁸⁸This is not to say that said conventions are always established or changed explicitly, of course.

I want to make it clear that the [she-form] feature should not be understood as the denotation for the entire pronoun; the composition of the pronoun remains complex. The new label is there to highlight that the link to the individual's gender is now indirect.

Next, we can let local contexts create a local context parameter c'_p with its own index c'_g , but allow double indexing with both c_p^* and c'_p . We then first check if the form used matches the form stored in c_g^* . This captures the majority pattern whereby in counterfactual scenarios the actual form has to be used. If no form is stored in c_g^* , we check the form used against c'_g . The variation will then be in how speakers handle cases when the actual form is unknown. They can choose to store no form in c_g^* (in which case they will check the form used against c'_g), or they can choose to store *they* in c_g^* (in which case gendered forms will clash with it). Some speakers will have equal command of both options while some will prefer one over the other (perhaps, categorically). It is also possible that some speakers store *they*, along with the more specific forms, for everyone (Sp2 in Table 5.1, for example, could be doing that), possibly with further competition-based preference for using the more specific forms when they are known. Furthermore, some speakers might allow c'_g to “save the day” when the form used clashes with the one stored in c_g^* (there seems to be one such speaker among those polled, Sp6). T–V features are meant to work analogously. Speakers can furthermore vary as to how they update the form indices (for example, whether they can set non-default form values based on perceived gender, what it takes for them to switch from a V to a T form, etc.).

Crucially, treating gender on pronouns as indexical is independent from treating it as an obligatorily non-restricting modifier giving rise to a non-restricting modifier inference. Except instead of something as simple as ‘being $g(i)$ entails being female’, we will be getting something like ‘being $g(i)$ entails having *she* as the preferred pronoun form in c_g of c_p ’. This inference can still project like a presupposition, but as before, it doesn’t have to be hardcoded lexically into any feature entries. This independence of the indexical nature of a given expression and the nature of the inference it contributes (i.e., whether it is presuppositional, and if it is, how this presupposition is triggered) is, of course, equally crucial for person features on pronouns, which also need to be

linked to c_p and are, thus, indexical. Naturally, I would also treat T–V features as obligatorily non-restricting modifiers and the corresponding inferences as non-restricting modifier inferences.⁸⁹

I will remain ignorant as to whether slurs, too, should be analyzed as form indexicals. However, if slurs do trigger presuppositions, I would expect those to be lexical. Once again, these issues—the indexical nature of a given expression and its presuppositional status, as well as the nature of triggering—are independent of each other.

At this point I would like to go back to the purported contrast in (5.13). I suspect that *her* in Sudo’s example in (5.13b) can be understood as anaphoric to *a beautiful girl*, not to John. So, it is possible that for some of the speakers who accept *her* in (5.13b), *a beautiful girl* introduces a new entry in c'_g for the hypothetical girl that John would be; in (5.13a), there is no expression that would do that. This intuition matches the observation in Yanovich 2010, fn. 3 that “[s]ome speakers are consistently more likely to be sympathetic to the use of a pronoun matching a non-actual gender in the presence of an overt obviously gendered DP—even if the DP is non-referential, as in [(5.12b)]”. I am not sure what exactly Yanovich means by “non-referential”, but I do believe that indefinites like the ones in (5.12b) and (5.13b) can introduce discourse referents available for anaphoric pick-up, as shown in (5.21a)—in contrast to (5.21b).

- (5.21) a. If Lea was a man, that man would be very beautiful.
 b. ??If Lea was male, that man would be very beautiful.

In fact, it is likely that we don’t see more speakers from my mini-poll accept counterfactual forms in (5.16a) and (5.16d) precisely because I explicitly specified that the pronouns are meant to refer back to the named individual (I originally did it to avoid construals whereby the pronouns refer to someone else entirely).

⁸⁹Despite the fact that many languages use the same morphemes for the T–V distinction as for the singular–plural distinction, I would not fully equate the two distinctions in the morphosyntax, the main reason being the differences in agreement that we observe in Russian between plural pronouns and V pronouns. For example, in copula constructions with a long form adjective as the predicate, the adjective will be singular if the subject pronoun *vy* is a V pronoun, but plural if it’s a plural pronoun. Things become more complicated for short form adjectives, but there are still discrepancies between the two types of *vy*.

Let me add one final note that even in languages that don't have the T–V distinction or another regularized honorific system, we still need a mechanism to store the appropriate forms used to address specific individuals and refer to them (which can also depend on the utterance situation), for example, which title to use (if any), whether it's OK to use their first name, which form of their name they prefer, etc. I would imagine that the system sketched above can be extended to such cases.

The upshot of this section is that there are reasons to treat gender on pronouns (as well as T–V features) as indexical, but the empirical picture at hand does call for a more expressive indexical system than previously assumed. While the details of the sketch above obviously need to be worked out, and, of course, more empirical work is needed (for one, I haven't even touched upon pronouns in attitude contexts or bound by quantifiers), I do believe that the direction outlined here is a promising one.

5.5 *Phi-features and Minimize Restrictors!*

Finally, if *phi*-features on pronouns are non-restricting modifiers, shouldn't the *Minimize Restrictors!* principle apply to them? For example, shouldn't we only use gendered pronouns when the gender is relevant (for instance, for identifying the referent of the pronoun in the absence of an overt index)?⁹⁰

The answer is likely no, *Minimize Restrictors!* probably doesn't apply to *phi*-features on pronouns. *Minimize Restrictors!* relies on the Gricean (1975) *Be Brief* maxim. If *phi*-features on pronouns are morphosyntactically obligatory, then any alternatives to a gendered pronoun (e.g., singular *they* in English), if at all available, would not be less complex, or “brief”.

This, however, still doesn't explain why *phi*-features feel more obligatory, even when they

⁹⁰Actually, I do know people who routinely use *they* when talking about someone whose gender identity they know if their interlocutor doesn't know this person and the gender isn't relevant (or, alternatively, if they don't want their interlocutor to guess who they are talking about and are trying to give away as little information as possible), and I often do this myself. Similarly, Sp2 in Table 5.1 from the previous section seems to allow *they* across the board, even when the gender of the referent is known.

do not serve the purpose of identifying the referent or appear to be otherwise relevant. The standard lexical presupposition analysis can appeal to *Maximize Presupposition!* (Heim 1991 et seq.), whereby out of two equally informative alternatives one should use the one with the strongest presuppositional content. There is, however, no good reason why a principle like *Maximize Presupposition!* should only apply to lexical presuppositions; see, for example, a recent attempt to come up with a more general principle that builds on the same general intuition but covers a wider range of phenomena in Anvari 2018. It stands to reason that such a principle would apply to *phi*-features on pronouns (but not to non-restricting modifiers that are not morphosyntactically obligatory).

It is, of course, also possible that further morphosyntactic considerations, akin to agreement, come into play and reinforce the obligatoriness of *phi*-features on pronouns, but I will not delve into this issue here.

5.6 Height specifications on gestures

The final type of content mentioned in section 1.2 is height specifications on gestures, illustrated previously in (1.1c) and repeated below in (5.22).

- (5.22) *Context: Zoe is a stuntwoman. The crew just filmed a scene in which she was fighting an extra, while the director of the movie Uma was away. Uma originally wanted Zoe to stab the extra in that scene, but she just learned that Zoe might have punched the extra in the face instead. Uma says:*



If Zoe punched the extra_{PUNCH-HIGH}, we'll have to reshoot the scene.

→ The extra that Zoe was fighting is taller than Zoe.

Similar constraints apply to projection of height specifications on gestures, as in the case of ordinary modifiers, spoken or gestural, and *phi*-features on pronouns. This point is illustrated in (5.23).

(5.23) *Context: Zoe is a stuntwoman. The crew just filmed a scene in which Zoe was punching an extra while the director of the movie Uma was away. ...*

- a. *...Uma knows who the extra was and that he is much taller than Zoe. Uma originally wanted Zoe to punch him in the face, not in the sternum. Uma says:*



If Zoe punched him_{PUNCH-HIGH}, that's OK,



but if she punched him_{PUNCH-LOW}, we'll have to reshoot the scene.

- b. *...Uma doesn't know who the extra was, but she knows that he was punched in the face, and she originally wanted the extra to be much taller than Zoe. Uma says:*

%??If Zoe punched him_{PUNCH-HIGH}, that's OK, but if she punched him_{PUNCH-LOW}, we'll have to reshoot the scene.

Intended: ‘...if (his face is ⟨higher than Zoe’s face, at Zoe’s face level⟩ and Zoe punched him)...’

In (5.23a), we have two potential areas of contact for the punching movement, the extra’s face or his sternum, whose height with respect to Zoe’s face is known and fixed. In (5.23b), we only have one potential point of contact, the extra’s face; its height, however, is not known, and we are trying to locally establish it via a height specification on the co-verbal gesture, but we fail to do so. The second scenario should be reminiscent of the ‘Non-projecting non-restricting’ scenarios from before. The first one could be thought of as reminiscent of either the ‘Projecting non-restricting’ scenarios from before or the ‘Restricting’ ones, depending on how we think of the two potential areas of contact.

Schlenker & Chemla (2018) draw empirical parallels between height specifications on pro-speech gestures and *phi*-features on pronouns (as well as height specifications on directional signs in sign language). The contrast in (5.23) is in line with this comparison, except I am looking at examples with co-speech, not pro-speech gestures whose content is, furthermore, mostly redundant

given the spoken verbs they co-occur with (except for the height specifications). I believe that the same observations can be made for co-speech gestures whose content is not redundant or for pro-speech gestures. The reason I am avoiding the former is to make the inferences of the examples simpler, and the reason I am avoiding the latter is to avoid any articulatory integration issues typical of pro-speech gestures. However, I believe any analysis of cases like in (5.23) should generalize to non-redundant co-speech gestures and to pro-speech gestures.

Schlenker and Chemla assume a presuppositional analysis for both height specifications on gestures and *phi*-features on pronouns. Naturally, treating the inferences of height specifications on gestures as lexical presuppositions raises the same question of how these presuppositions are triggered and why they resist local accommodation (although some people find (5.23b) very marginally acceptable, which I come back to at the end of this section). Moreover, under this view, the nature of the empirical parallels between the two types of content remains a mystery.

I propose analyzing height specifications on gestures as modifiers and further assimilating them to *phi*-features on pronouns or to ordinary modifiers, such as adjectives. The most straightforward way to do so is to say that the gestures in (5.23) contain incorporated nominal arguments specifying the area of contact for the punching movement (akin to *in the face* and *in the sternum*, but without the assumption that there is a PP layer in the syntax). The height specifications on the gestures then modify these arguments. The rest depends on what exactly the nature of these arguments is.

We could say that the incorporated arguments in (5.23) are essentially pronouns, standing for the extra's face (x_i) and sternum (x_j). Then the analysis of height specifications would be the same as for *phi*-features on pronouns. For example, the denotation of the "face" argument in (5.23a) would be $\lambda w. \lambda x. \text{high}(x, w) \wedge x = g(i)$, and we would get a projecting inference that the extra's face is located high(er than Zoe's face). Nothing blocks the projection of these inferences in (5.23a), so the utterance is felicitous. In (5.23b), however, the inferences of the height specifications on the gestures can't project, given the context, so a failure obtains.

Alternatively, we could argue that the incorporated arguments in (5.23) are more like or-

dinary definite descriptions, with a contextually determined descriptor (e.g., $\lambda w.\iota x.\text{high}(x, w) \wedge \text{body-part}(x, w)$). The height specifications in (5.23a) would then be restricting modifiers, just like the adjective *large* in *her large dog*, when there are two contextually relevant dogs of different size. Height specifications in (5.23b) would then be non-restricting modifiers (as there is only one relevant body part), but since they can't project in the context, (5.23b) would fail in the same way as similar examples with adjectives.

Can we distinguish between these two options? A possible direction would be to compare the behavior of height specifications on co-verbal gestures under ellipsis and *only* to that of *phi*-features on pronouns and ordinary modifiers, spoken and gestural. I come back to this issue in section 7.2.2.

Now, why is (5.23b) marginally acceptable for some people? The construal under which the gestures in (5.23) contain incorporated nominal arguments might not be the only one available. It might be possible, at least for some people, to construe the gestures in (5.23) as containing no nominal arguments whatsoever, with the height specifications being adverbials specifying the direction of motion (akin to *upwards* and *forwards*). Under this construal, the height specifications in both examples in (5.23) would be restricting modifiers and should be (marginally) acceptable. Because the contrast in (5.23) still obtains, either this construal is dispreferred for some reason, or there is an independent reason why (5.23b) is still degraded under this construal. It is furthermore possible that the people who marginally accept (5.23b) simply get the metalinguistic focus interpretation, which I discuss briefly in subsection 7.2.1, in which case nothing extra needs to be said. I leave this issue for future research.

Another direction for future research is extending the modifier-based analysis proposed here for *phi*-features on pronouns and height specifications on gestures to various geometric properties of signs in sign language. The extension itself should actually be very straight-forward. The crucial question, however, is, once again, which cases should be analyzed as involving pronouns and which cases should be analyzed as involving ordinary definite descriptions.

Modifiers and supplements in the verbal domain

Abstract

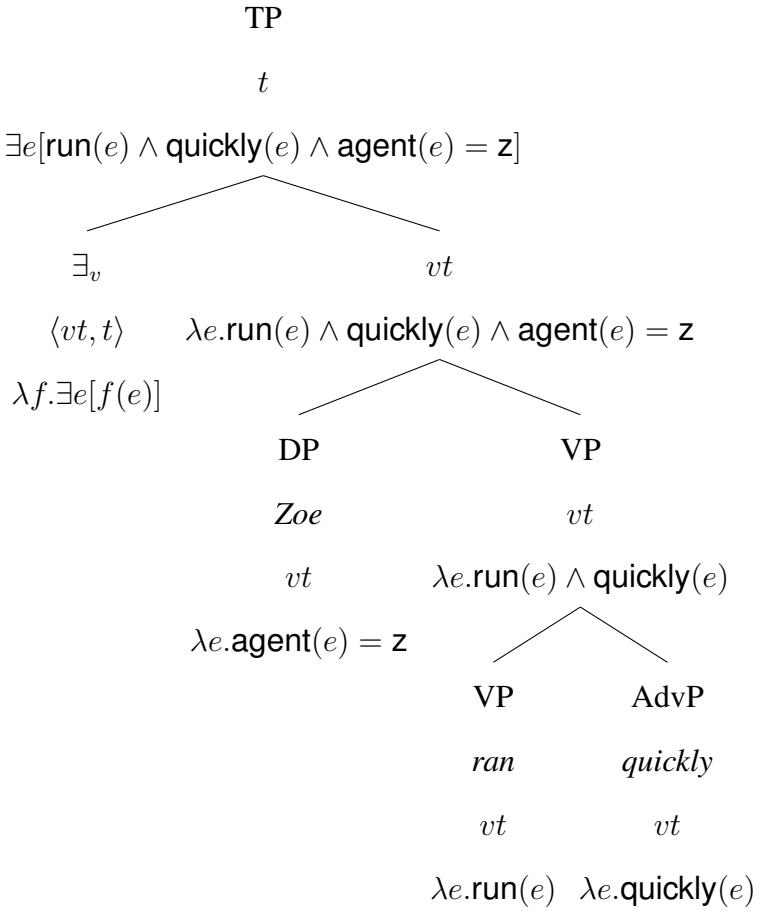
Throughout this dissertation I have mostly focused on adnominal content (with some exceptions). In this chapter I look at the modifier vs. supplement distinction in the verbal domain. First, I show that the distinction exists for spoken content, but supplements in the verbal domain are more flexible regarding what anchors they can combine with than adnominal supplements. Next, I show that co-verbal gestures also come with a ban on ‘Non-projecting non-restricting’ interpretations, which suggests that their projection is constrained by their composition in the same way as for co-nominal gestures. I finish with a more exploratory note on appositive-like gestures in the verbal domain, where I show, among other things, that comparing them to spoken appositive relative clauses only is not the most insightful strategy and cannot, in and of itself, be used as an argument that such gestures are supplements.

6.1 Spoken modifiers and supplements in the verbal domain

The modifier vs. supplement distinction exists in the verbal domain, broadly construed (i.e., anything from V to the entire clause), as well.

If we assume a standard Neo-Davidsonian setup (as in Carlson 1984; Parsons 1990, a.o.), verbs start out as sets of events and are modified by other sets of events (arguments and adjuncts) via predicate modification up until existential closure applies and binds the event variable. A sample derivation, ignoring tense and aspect, is given in (6.1) (for our purposes it doesn't matter how *Zoe* combines with its theta-role; also, I don't include worlds here).

(6.1) *Zoe ran quickly.*



Supplements, i.e., pieces of content that combine with anchors and return a proposition of a special kind about them, exist in the verbal domain, too, however they have more freedom in what they can take as their anchors as compared to the nominal domain. In particular, it is hard or impossible for adnominal appositives to take NPs as their anchors, as shown in (6.2a). As mentioned before, adnominal appositives seem to need a discourse referent as their anchor, which the DP *any wombats*

doesn't introduce. It would make sense for the appositives in (6.2a) to be able to target the NP *wombats* and add some propositional content about wombats in general, but it is impossible (or very hard, as in the case of appositive relative clauses). If appositives contain anaphoric elements (as is assumed in Koev 2013; AnderBois et al. 2013), those anaphoric elements don't seem to be able to target implicitly introduced discourse referents, which is possible for ordinary pronouns, as shown in (6.2b), where *they* can refer to wombats in general.

- (6.2) a. Zoe doesn't have any wombats, {??who are nocturnal animals, *nocturnal animals}.
- b. Zoe doesn't have any wombats. They are nocturnal animals.⁹¹

The same constraint doesn't apply to supplements in the verbal domain, as shown in (6.3), where it looks like the anaphoric element in the appositive relative clause can target just the predicate *run a marathon* without there being any event discourse referent.

- (6.3) Zoe has never run a marathon, which is a hard thing to do.
 ≈ ‘Zoe has never run a marathon, and running a marathon is a hard thing to do.’

Supplements in the verbal domain can also (at the very least) take events and propositions as their anchors, as shown in (6.4a) and (6.4b).

- (6.4) a. The roof collapsed, which injured several people.
 ≈ ‘The roof collapsed, and that event injured several people.’
- b. Kim says that Zoe ran a marathon, which I find hard to believe.
 ≈ ‘Kim says that Zoe ran a marathon, and I find the proposition that Zoe ran a marathon hard to believe.’

Similarly, as we have already seen in section 4.4, sentence-level adverbs like *surprisingly* and *un-*

⁹¹Of course, a discourse link needs to be established between the two sentences for this whole utterance to be coherent. Here it is reasonable to interpret the second sentence as introducing the reason for Zoe not having any wombats.

fortunately, as well as sentence-level mirative facial expressions are supplements that take propositions as their anchors.

This higher flexibility of supplements in the verbal domain makes it especially crucial to check all possible compositional construals for gestures in the verbal domain.

6.2 Gestural modifiers and supplements in the verbal domain

6.2.1 Projection of co-verbal gestures

Schlenker (2018a) observes that co-verbal gestures, too, give rise to projecting inferences by default, as in (6.5).

- (6.5) *Context: Zoe just took part in a multi-sport competition; we don't yet know how she performed. One of the tasks was shooting a weapon once.*



If Zoe hit the bullseye when she shot_{LONGBOW}, I'll give you \$10.

→ When Zoe shot, she shot a longbow.

Schlenker claims co-verbal gestures trigger cosuppositions, just like co-nominal gestures, but doesn't say much about how that relates to their compositional integration.

Under the analysis developed in this paper, co-nominal gestures are either supplements or preferably non-restricting modifiers. Curiously enough, we can replicate the empirical observation I have made previously for co-nominal gestures with co-verbal gestures: co-verbal gestures can (marginally) be restricting, but when they are non-restricting, their contributions have to project, as shown in (6.6). I take *when*-clauses to contain definite descriptions of events. In (6.6a), there is only one relevant shooting event; the type of weapon used is known and is not at-issue, thus, the content of the gesture is allowed to project, and the utterance is felicitous. In (6.6b), there are two relevant shooting events, performed with two different types of weapon; the gesture is acting as a

restricting modifier helping us pick out the right event based on the type of weapon used. In (6.6c), there is only one relevant shooting event with the type of weapon used unknown; we are trying to make the type of weapon at-issue, but fail.

(6.6) *Context: Zoe just took part in a multi-sport competition; we don't yet know how she performed. ...*

- a. *...Participants first shot a weapon of their choice once and then swam a mile.*

If Zoe hit the bullseye when she **shot_{LONGBOW}**, I'll give you \$10. I won't be betting on her **swimming**, though.

→ When Zoe shot, she shot a longbow.

- b. *...For the shooting part of the competition, participants first shot a longbow once and then a gun.*

?If Zoe hit the bullseye when she **shot_{LONGBOW}**, I'll give you \$10. I won't be betting on her **gun** shooting, though.

↗ When Zoe shot, she shot a longbow.

- c. *...For the shooting part of the competition, participants had to choose a longbow or a gun and then shoot it once.*

#If Zoe hit the bullseye when she **shot_{LONGBOW}**, I'll give you \$10. I won't be betting if she was shooting a **gun**, though.

Intended: ‘If ((when Zoe shot, she shot a longbow) and (Zoe hit the bullseye when she shot))...’

Under the view proposed here, the gesture *LONGBOW* in (6.6) can in principle (i.e., from a purely compositional perspective) be a modifier or a supplement. As a modifier, it can in principle be taken to represent either just the longbow as a theme ($\lambda e \lambda w. \text{longbow}(\text{theme}(e, w))$), or shooting a longbow ($\lambda e \lambda w. \text{shoot}(e, w) \wedge \text{longbow}(\text{theme}(e, w))$), or even Zoe shooting a longbow ($\lambda e. \lambda w. \text{shoot}(e, w) \wedge \text{longbow}(\text{theme}(e, w)) \wedge \text{agent}(e, w) = z$).⁹² We could even posit more

⁹²This last construal can emerge, for example, if we think that the speaker is actually representing Zoe within their

silent syntactic structure, construing of the gesture as equivalent to a PP *by shooting a longbow* (roughly, $\lambda e \lambda w. \exists e' [\text{shoot}(e', w) \wedge \text{longbow}(\text{theme}(e', w)) \wedge \text{agent}(e', w) = z \wedge \text{by}(e, e', w)]$), in which case *Zoe* would control the silent subject of the clause within that PP.

As is the case with co-nominal modifier gestures, co-verbal ones will be preferably non-restricting, in which case they will give rise to non-restricting modifier inferences. I will not go through all the possible construals for the example above. In (6.7), I only show a derivation for one such construal. I assume, by stipulation, that the local context of *shot_{LONGBOW}* is the set of all events whose agent is *Zoe*.⁹³

(6.7) *Zoe* [[*shot*]_{LONGBOW}].

- a. spoken expression *S*:

$$[\![\text{shot}]\!] = \lambda e \lambda w. \text{shoot}(e, w)$$

- b. gesture *G*:

$$[\![\text{LONGBOW}]\!] = \lambda e \lambda w. \text{longbow}(\text{theme}(e, w))$$

- c. spoken expression modified by the gesture $[[S]^G]$:

$$[\![[\text{shot}]_{\text{LONGBOW}}]\!] = \lambda e \lambda w. \text{shoot}(e, w) \wedge \text{longbow}(\text{theme}(e, w))$$

- d. propositional local context *c'*:

C

- e. non-propositional local context *c''*:

$$\lambda e \lambda w. C(w) \wedge \text{agent}(e, w) = z$$

- f. non-restricting modifier inference $c'' \Rightarrow_{\forall} (S \Rightarrow [S]^G)$:

$$\forall e \forall w [(C(w) \wedge \text{agent}(e, w) = z) \rightarrow (\text{shoot}(e, w) \rightarrow \text{longbow}(\text{theme}(e, w)))]$$

≈ ‘For all events *e*, for all worlds *w* in C: if *Zoe* is the agent of *e* in *w*, then if *e* is a shooting in *w*, the theme of *e* is a longbow in *w*.’

Of course, this gestural modifier can still be restricting, as in (6.6b), even if this interpretation is iconic demonstration.

⁹³This is not crucial, because by the assumption I have been making throughout this dissertation, the domain has already been restricted only to the relevant shooting events, which are shooting events by *Zoe*.

degraded, in which case no projecting inference arises.

However, the gesture *LONGBOW* in the examples above can also be construed of as representing an existential quantifier over events rather than a predicate of events (akin to *a longbow shooting*), in which case it would compose with the spoken definite description of the event as a supplement, adding some propositional content about that event (along the lines of ‘Zoe shot, which was a longbow shooting event’). This supplement will have to project, just like the other supplements we have seen. However, unlike in the nominal domain, it should also in principle be possible that the gesture is construed of as a supplement with an event predicate anchor (e.g., something along the lines of ‘Zoe shot, and, by the way, shooting events are done with a longbow’ or ‘Zoe shot, and, by the way, shooting events by Zoe are done with a longbow’), because, as we have seen in (6.3) and earlier in subsection 4.2.1, event predicates can in principle serve as anchors for supplements in the verbal domain. However, such a construal would be really hard to distinguish from a non-restricting modifier construal.

As before, under a view whereby co-speech gestures trigger cosuppositions and compose conjunctively (either by default or as a result of local accommodation) across the board, it is unclear why for something like *when Zoe shot_{LONGBOW}* we don’t get interpretations along the lines of ‘during the event of Zoe shooting and an event of (Zoe) shooting a longbow’ or ‘during the event of Zoe shooting, and that event was an event of (Zoe) shooting a longbow’.

The conclusion is, thus, the same as for the nominal domain: how gestures project is determined by how they compose, and they compose in the way maximally similar to spoken content.

6.2.2 Appositive-like gestures

Schlenker (2018a) further argues that “post-speech gestures” (i.e., once again, gestural adjuncts adjoining to spoken expressions preceding them and packaged into their own prosodic phrases) associating with verbal spoken expressions are supplements. As before, his main arguments are that (i) contributions of such gestures seem to project by default, as shown in (6.8), and (ii) they seem to need event discourse referents in a way in which analogous co-speech gestures don’t, as shown

in (6.9). I must say that in my experience judgements for examples like in (6.9) become extremely variable and subtle, but overall, there does seem to be a contrast between the two examples in (6.9).

- (6.8) If Zoe shoots at the target, LONGBOW, I'll give you \$10.

→ If Zoe shoots at the target, she'll shoot a longbow.

- (6.9) a. ?Zoe didn't shoot at the target_{LONGBOW}.

- b. ??Zoe didn't shoot at the target, LONGBOW.

It is quite clear why gestures like the one in (6.8) are unlikely to be construed of as modifiers and, thus, have a potential to be restricting and, consequently, truth-conditionally non-vacuous. Modifiers in the verbal domain, too, prefer to be packaged into the same prosodic phrase as the constituent they modify, and gestures like in (6.8) prefer to be in their own prosodic phrases for articulatory reasons. The same considerations don't apply to co-verbal gestures like in (6.9a), so they can be modifiers, which, however, doesn't exclude the possibility that they can also be supplements.

Now, Schlenker's second argument, which he uses to distinguish between supplements and other projective content, is built on an analogy with appositives that require an event anchor. In particular, he adduces paradigms like in (6.10) to show that the gestures that he calls "post-speech" pattern with appositives.



- (6.10) a. {It's likely, It's unlikely} that John  helped_{LIFT} his son.
- b. {It's likely, #It's unlikely} that John helped his son, which (by the way) he did by lifting him.
- c. {It's likely, #It's unlikely} that John helped his son, LIFT.

(Schlenker 2018a, (76), cited with minor notational changes)

The appositive in (6.10b) is unacceptable because it requires an event anchor and there is no avail-

able event discourse referent. Yet, as we have previously seen in (6.3) and in subsection 4.2.1, spoken supplements in the verbal domain don't necessarily need event discourse referents and can, in particular, take event predicates as their anchors. If *LONGBOW* is indeed a supplement, it is, in fact, a puzzle why (6.9b) can't be construed of along the lines of (6.11).

- (6.11) Zoe didn't shoot at the target, which [= shooting at the target] is done with a longbow.

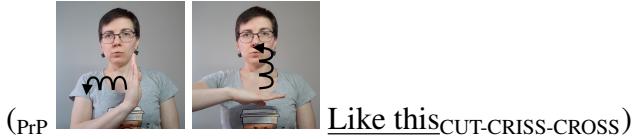
One could argue that supplement construals with predicate anchors like in (6.3) or (6.11) are constrained, and iconic gestures are simply too semantically and/or morphosyntactically poor to have such construals. I am not sure how one can maintain that while still maintaining that gestures in the verbal domain can be supplements in general (after all, no spoken counterpart of nominal appositives exists in the verbal domain, at least not in English; all such supplements are appositive relative clauses with an overt relativizer), but I think this direction should be considered.

Another possibility (which doesn't exclude the previous one) is that it is unclear why one would even bother adding the content contributed by the gesture in both examples in (6.9), given that no event of the relevant kind occurred in the first place (see the discussion of the examples in (4.17) in subsection 4.3.3); the spoken paraphrase of this construal that I provided in (6.11) certainly feels incredibly stilted. This irrelevance would make both gestures in (6.9) degraded to a certain extent. However, a co-verbal gesture likely requires less articulatory effort than a prosodically independent one, and its irrelevance might thus be less offensive.

Either way, the contrast in (6.9) cannot be derived under a simple "co-speech gestures trigger cosuppositions, post-speech gestures trigger supplements" view. Once again, one needs to seriously engage with how such gestures are construed of and integrated into the compositional structure, if at all.

One final note is that it is quite possible that appositive-like gestures like in (6.8) can have elaboration uses, akin to those of nominal appositive-like gestures discussed in footnote 43. For example, both sentences below are naturally interpreted as requests to cut the pie in a specific way, without giving rise to an inference that if the addressee cuts the pie, they'll do so in a certain way:

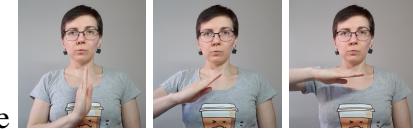
- (6.12) a. (_{PrP} Can you please start cutting the pie?)



- b. (_{PrP} Can you please start cutting the pie?) (_{PrP} CUT-CRISS-CROSS?)

As with *one*-appositives (see (2.51) from section 2.2), such elaborations are clearly not bona fide restricting modifiers, as they cannot be used as planned restrictors, as shown in (6.13).

- (6.13) a. (_{PrP} If Daisy starts cutting the pie like this_{CUT-CRISS-CROSS}),



that'll do, (_{PrP} but if she starts cutting the pie

like this_{CUT-WEDGE}), we won't have enough slices for everyone.

- b. #(_{PrP} If Daisy starts cutting the pie), (_{PrP} like this_{CUT-CRISS-CROSS}), that'll do, (_{PrP} but if she starts cutting the pie), (_{PrP} like this_{CUT-WEDGE}), we won't have enough slices for everyone.
- c. ?(_{PrP} If Daisy starts cutting the pie **CUT-CRISS-CROSS**), that'll do, (_{PrP} but if she starts cutting the pie **CUT-WEDGE**), we won't have enough slices for everyone.⁹⁴
- d. #(_{PrP} If Daisy starts cutting the pie) (_{PrP} **CUT-CRISS-CROSS**), that'll do, (_{PrP} but if she starts cutting the pie) (_{PrP} **CUT-WEDGE**), we won't have enough slices for everyone.

The paradigm in (6.13) is particularly informative, because it shows how much depends on prosodic grouping. Schlenker (2018a) makes systematic comparisons between gestures and modifiers of the *like this...* form. However, he doesn't control at all for prosodic grouping.

This is especially important because appositive-like elaborations do not have to actually restrict the expression they elaborate on. *One*-appositives discussed in subsection 2.2.1 do seem to

⁹⁴This utterance is arguably degraded for reasons of articulatory integration.

always do that, which seems to be due to the fact that, as pointed out to me by Alan Munn (p.c.), the modifier within a *one*-NP typically has to be restricting, as illustrated in (6.14).⁹⁵

(6.14) A: The finals week is approaching. How are the students in your class doing?

B: (i) None of the stupid brats will pass the exam.⁹⁶

→ None of the students in my class will pass the exam.

→ All the students in my class are stupid.

(ii) None of the stupid ones will pass the exam.

↗ None of the students in my class will pass the exam.

↗ All the students in my class are stupid.

That is not, however, the case for other appositive-like elaborations. It is quite possible to have elaborations such that their truth-conditional contribution is equivalent to the expression they clarify and they are used to simply be more precise. Two examples of such “non-restricting” elaborations, a nominal and a verbal one, are given in (6.15). (6.15a) is possible in a context in which Stephanie only has one dog, so there is no way the elaboration is making a truth-conditionally non-vacuous contribution; however, it can be made as a second thought, for example, if Mia has a special affinity for poodles, the speaker might realize mid-utterance that a more precise description might actually be more explanatory. Similarly, (6.15b) can be uttered in a context in which Zoe only ever works out by lifting weights, and the speaker goes for a more precise expression mid-utterance, because that explains the consequent better, as not all forms of workout call for an increase in protein intake.

(6.15) a. If Stephanie brings her dog, her poodle, Mia will be happy.

b. If Zoe’s gonna work out today, lift weights, she’ll need a protein-rich lunch.

In fact, (6.15b) can be easily replicated with an appositive-like gesture:

⁹⁵I think a more accurate observation is that it always has to be in focus, which often forces it to be at-issue, and, as we have seen again and again, the only way for a modifier to be at-issue is to be restricting; see the discussion of focus at the end of subsection 2.1.2.3.

⁹⁶I believe the reading whereby *the stupid brats* refers only to a subset of the students in the speaker’s class is possible, but the point is that the non-restricting reading is available here, too, unlike in (6.14-ii).



- (6.16) If Zoe's gonna work out today, LIFT-WEIGHTS, she'll need a protein-rich lunch.

There is a question as to how many cases of what Schlenker calls “post-speech gestures” are in fact such appositive-like elaborations, and there is a further question as to how much we should assimilate such appositive-like elaborations to actual supplements. Are such elaborations as compositionally integrated into utterances as bona fide supplements, given that they don’t seem to be planned? In fact, are all supplements (e.g., appositive relative clauses) equally compositionally integrated into utterances (many of them don’t seem to be planned either)? What about *i.e.*, ... or *that is...* elaborations, as in (6.17), which seem to be more planned than corrections or elaborations discussed so far and which many instances of appositive-like gestures seem to resemble?

- (6.17) a. If Zoe's gonna work out today, *i.e.*, lift weights, she'll need a protein-rich lunch.
b. If Zoe's gonna work out today, *that is* lift weights, she'll need a protein-rich lunch.

I believe all these questions need to be asked in a rigorous and systematic way (once again, using proper linguistic tools at all levels of representation), and I hope they will be asked in future research.

Conclusion

7.1 Summary of the dissertation

7.1.1 Summarizing the main points of the dissertation

In this dissertation I have put forward a general approach to projection of content that has its own node in the morphosyntax from under semantic operators. I have argued that how such content projects is determined by how it composes in the syntax/semantics. I have furthermore applied this approach both to spoken and to gestural content.

I have focused on two composition strategies, which come with associated projection patterns. The two strategies are (subsective) modification and supplementation. I have classified several different types of projective or potentially projective content either as modifier or as a supplement. I summarize the modifier vs. supplement distinction and my claims about the different types of content in Table 7.1.

For spoken content, one can usually tell whether a given piece of content is a modifier or a supplement based on its morphosyntactic makeup (although in some cases the difference is only signalled by prosody). This dissertation demonstrates that the modifier vs. supplement distinction holds for gestural content (such as hand gestures and facial expressions) as well. Because said content is often morphosyntactically poor, one and the same string with one and the same gesture or facial expression can in principle be ambiguous between the modifier and the supplement construal; I have argued, in particular, that this is often the case for co-speech gestures.

Table 7.1: Modifiers vs. supplements: a final summary.

(Subselective) modifiers	Supplements
Compose with β , yielding α such that α entails β	Compose with β , return a proposition of a special kind about β
Can be restricting or non-restricting	Can never be restricting
Trigger projecting inferences when non-restricting; triggering is pragmatic	Almost always trigger projecting inferences; triggering is conventional
Examples (target content bolded):	Examples (target content bolded):
<ul style="list-style-type: none"> • adnominal adjectives <i>(a blond stuntwoman)</i> • restrictive relative clauses <i>(her dog that's large)</i> • degree modifier adverbs <i>(Mia got surprisingly drunk)</i> • modifiers of verbal projections <i>(Zoe shot at the target with a longbow)</i> • NP-level gestures <i>(her [[NP dog]LARGE])</i> • degree modifier facial expressions <i>(Mia got [[drunk]O-O])</i> • <i>phi</i>-features on pronouns <i>(her^{3,sg,fem} dog)</i> • height specifications on gestures <i>(punch_{PUNCH-HIGH})</i> 	<ul style="list-style-type: none"> • adnominal appositives <i>(her dog, (who is) a large animal)</i> • appositives with verbal anchors <i>(Zoe shot at the target, which [= this event of shooting at the target] she did with a longbow; Zoe ran a marathon, which [= running a marathon] is not an easy thing to do)</i> • sentence-level adverbs <i>(Surprisingly, Mia got drunk)</i> • DP-level gestures <i>([[DP her dog]LARGE; [[DP her dog], LARGE]])</i> • sentence-level facial expressions <i>([[Mia got drunk]O-O])</i>

The two construals do, however, come apart, for example, in cases when one construal cannot yield the right interpretation, as is the case with restricting gestures, for instance, which can only be modifiers, as supplements can't be restricting. They also come apart in the case of co-nominal gestures without an individual discourse referent to serve as an anchor for them, in which case the supplement construal is not possible either, as adnominal supplements require individual anchors. Next, the two construals come apart in many cases of prosodically independent (non-co-speech) gestures integrated into otherwise spoken utterances. Those tend to follow the rules of linearization, syntax/prosody mapping, and syntactic displacement of the spoken language they are embedded into. Combined with articulatory considerations of gesture integration into the spoken stream, these rules often block the modifier construal for prosodically independent gestures

in English, however, the syntactic and prosodic rules are language-specific, so it is easier to get modifier construals for prosodically independent gestures in some other languages (e.g., French and Russian). The clearest case of the modifier and supplement construals for gestural content coming apart I have identified so far is the distinction between degree modifier and sentence-level facial expressions encoding surprisal. The two interpretations associated with the two construals are very distinct, even if the exact alignment of the facial expression is not a very reliable predictor for which construal is intended.

What is clear, however, is that the compositional construal—and, consequently, the projection behavior—of a given piece of secondary modality content cannot be directly predicted based on whether it co-occurs with something in a more primary modality or whether it is syntactically optional (even though there might be some interface-mediated correlations with either factor). The “co-something” status might make a given piece of content preferably truth-conditionally vacuous (as seems to be the case for many co-speech hand gestures), but it doesn’t directly determine how this truth-conditional vacuity will be assured; furthermore, this preference can be overridden by other considerations. Similarly, how conventionalized a given piece of secondary modality content is does not directly affect its projection behavior. Conventionalized gestures can have conventionalized constraints on how they can compose, which will affect how they project, if at all, but once again, there is no direct link between level of conventionalization and projection. This dissertation, thus, corrects the tendency in the existing literature to directly link the projection behavior of secondary modality content to its “co-something” status, syntactic optionality, or level of conventionalization (e.g., Ebert & Ebert 2014; Ebert 2017; Tieu et al. 2017, 2018; Schlenker 2018a,b) and calls for more attention to linguistically meaningful interface interaction when studying such types of content.

7.1.2 Summarizing the contributions of the dissertation

Below I summarize what I believe to be the main theoretical and empirical contributions of this dissertation:

1. Theoretical contributions:

- (a) This dissertation offers a general approach to predicting how a given piece of content will project depending on how it composes in the syntax/semantics, which can be applied beyond the types of content covered in this dissertation. In particular, it makes predictions for whether local interpretations will be in principle available for a given piece of content (*yes* for modifiers, unless they always combine with expressions that cannot be restricted any further, as is the case with *phi*-features; *no* for supplements⁹⁷).
- (b) This dissertation offers a general, pragmatic theory of how inferences of non-restricting (i.e., truth-conditionally vacuous) modifiers are triggered and project. It builds on the definition of non-restricting modifiers in Leffel 2014 and the notion of cosuppositions in Schlenker 2018a,b,c in doing so. I refine Leffel's definition of non-restricting modifiers by divorcing their truth-conditional vacuity, which holds at the level of the entire utterance, from the inference that the expression being modified entails the result of modification (the non-restricting modifier inference), which is sensitive to local contexts. I also improve on Leffel's theory of non-restricting modifier inferences, which treats them as semantic presuppositions, by making the link between the pragmatic definition of non-restricting modifiers and the projecting inferences they give rise to more transparent. Also, my theory is less constrained by idiosyncratic syntactic considerations in the nominal domain and is, thus, more easily generalizable beyond intersective adnominal modification (in particular, I generalize it to non-intersective adnominal modifiers, degree modifiers, and modifiers in the verbal domain). In addition, it follows straight-forwardly from the proposed analysis that non-restricting modifier inferences cannot be interpreted locally under semantic operators. I also propose to equate cosuppositions with non-restricting modifier inferences, and by doing so, I both constrain cosuppositions (which are unconstrained in Schlenker's original formulation) so that they don't apply in non-modification cases (e.g., to DP-level

⁹⁷Modulo the exceptional behavior of some appositive relative clauses discussed in Schlenker 2013; Jasinskaja & Poschmann 2018, which is much more constrained than the availability of restricting interpretations for modifiers.

gestures, sentence-level facial expressions, or gestures adjoining to definite event descriptions), and make them more general so that they can apply to non-intersective modifiers.

- (c) In the process of fitting my theory of non-restricting modification to degree modifiers, I also make an adjustment to the analysis of degree modifiers in Kennedy & McNally 2005, which makes it compatible with multiple degree modifiers modifying the same expression and also renders modification of adjectival projections compositionally more similar to modification of nominal and verbal projections.
- (d) This dissertation offers a theory of how gestural content integrates into otherwise spoken utterances and how it projects. While building on previous empirical generalizations and theoretical conjectures (Ebert & Ebert 2014; Ebert 2017; Tieu et al. 2017, 2018; Schlenker 2018a,b, etc.), it corrects the ubiquitous assumption in earlier literature that how a given gesture projects is determined by its linearization. The proposed theory connects the projection behavior of gestures to their composition and asserts that gestures can compose in different ways, just like spoken content. While there exist various interface-based constraints on how different compositional configurations with gestures can be linearized, there is no direct link between projection and linearization. Unlike previous theories, this interface-based view of gestures effortlessly accounts for discrepancies in projection behavior of gestural expressions that are linearized in the same way (e.g., NP- vs. DP-level co-speech or pro-speech gestures, degree modifier vs. sentence-level co-speech or co-gesture facial expressions, co-speech gestures adjoining to event predicates vs. co-speech gestures adjoining to definite event descriptions). By rejecting a more general assumption that there is a single, uniform way in which gestures linearized in a certain way contribute to the meaning of spoken utterances, this view on gestures allows for both compositionally integrated and compositionally non-integrated co-speech gestures (cf. Hunter 2018, 2019).
- (e) This dissertation also offers a new theory of semantic composition and projection of *phi*-features on pronouns, which proposes that *phi*-features compose as modifiers within def-

inite descriptions, but because they always modify expressions whose extension is a singleton set, they are always non-restricting and, thus, always give rise to strongly projecting inferences. This theory builds on the ideas that *phi*-features on pronouns trigger presuppositions (Cooper 1983 et seq.) and that pronouns are definite descriptions (Elbourne 2005). However, it decomposes the presuppositional black box in the previous literature by stating that these presuppositions aren't lexical properties of *phi*-features, but arise via the general pragmatic mechanism for generating non-restricting modifier inferences. This theory is also more morphosyntactically plausible than the alternatives, as it treats *phi*-features on pronouns as modifiers with their own nodes in the morphosyntax, which is in line with the morphosyntactic evidence we have from pronouns in languages other than English.

- (f) This dissertation also outlines the direction for an analysis of gender and T–V features on pronouns as form indexicals, i.e., indexicals that are interpreted relative to indices within the context parameter that keep track of appropriate forms of expressions. While this sketch builds on some insights from the previous literature (e.g., Yanovich 2010; Schlenker 2007), it posits a more expressive indexical system, which allows capturing the more complex and variable empirical picture regarding the use of pronoun forms in various local contexts.
- (g) This dissertation also shows that the modifier analysis is extendable to height specifications on gestures, making the empirical parallels between those and *phi*-features on pronouns observed in Schlenker & Chemla 2018 non-accidental.

2. Empirical contributions:

- (a) Chapter 2 of this dissertation makes a novel observation that non-restricting modified inferences (i.e., inferences that the expression being modified entails the result of modification) project relative to a local context. This allows us to separate the definition of non-restricting modifiers as modifiers that don't affect the global truth conditions of the entire utterance from projection of non-restricting modifier inferences, which is sensitive to the internal composition of the utterance. It also observes that non-restricting modifier inferences can't

be interpreted locally under semantic operators, which is expected under the proposed analysis of how they are triggered.

- (b) Chapter 3 of this dissertation reports experimental data on acceptability of various interpretations of co-nominal gestures as compared to adjectives and appositives. Apart from these data serving as a testground for various theories of co-speech gestures in Chapter 4 of this dissertation, the experiment itself adds to the body of experimental literature on gestures in a non-trivial way, as it uses an acceptability-based task (cf. inference-based tasks in Tieu et al. 2017, 2018) and makes composition-based interpretative distinctions while drawing an explicit comparison with two types of spoken content (cf. Tieu et al. 2017, 2018; Zlogar & Davidson 2018; Esipova 2019a).
- (c) Chapter 4 of this dissertation adduces novel data on prosodically independent compositionally integrated gestures (pro-speech gestures) in English and other languages. It also identifies two different construals for facial expressions encoding surprisal associated with two different interpretations (sentence-level supplement vs. degree modifier), which mirrors a similar distinction for spoken adverbs like *surprisingly*. Finally, it discusses the projection behavior of some conventionalized gestures and how it, too, is determined by their composition.
- (d) Chapter 5 points out parallels in projection between non-restricting modifier inferences and inferences contributed by *phi*-features on pronouns as well as height specifications on gestures. In this chapter I also adduce novel data from joint fieldwork with Naomi Lee that show that pronoun-internal property-denoting modifiers in Khoekhoe (first described in Lee 2019) are obligatorily non-restricting and give rise to obligatorily projecting inferences. These data serve as a new piece of evidence that (i) pronouns need a property-denoting layer, and (ii) this property can be modified, but in an obligatorily non-restricting way. Finally, in this chapter I observe empirical parallels between gender and T–V features on pronouns with respect to their local context sensitivity, both in terms of the patterns

attested as well as the locus and amount of inter- and intra-speaker variation. I use this observation to argue that we need an indexical-based analysis of both types of content with more expressive power than previously assumed in the literature.

- (e) Chapter 6 makes a variety of novel observations about modifiers and supplements in the verbal domain. In particular, I show that restricting construals are possible for co-verbal gestures, but when a restricting construal is blocked, the inference contributed by the gesture has to project, which makes them similar to co-nominal gestures. I also adduce novel examples highlighting the role of prosodic grouping in interpretation of both spoken and gestural post-posed adjuncts combining with verbal projections.
- (f) Section 7.2 of the current chapter draws a parallel between the behavior of a given piece of content under ellipsis and its behavior under attitude verbs and makes some preliminary observations about constraints on multi-modal linearization, leaving the exploration of these data for future research.

7.2 Directions for future research

7.2.1 Metalinguistic focus

When studying projection, one must be wary of potential effects of different types of focus. I have discussed the interaction of QUD-addressing and non-QUD-addressing contrastive focus with projection at the end of subsection 2.1.2.3. There is, however, another type of focus that complicates the empirical picture when it comes to projection, namely, the so-called METALINGUISTIC FOCUS. Most types of content I have discussed in this dissertation (with the exception of, perhaps, appositives) can be targeted by metalinguistic focus, thus, preventing them from projecting. In fact, using metalinguistic focus can yield what on the surface looks like ‘Non-projecting non-restricting’ interpretations, which we have seen to be systematically unavailable, as exemplified in B’s responses in (7.1).

- (7.1) a. A: Stephanie brought her ginormous dog to the party.
 B: Stephanie didn't bring her **ginormous** dog—her dog is **average**-sized!
- b. A: Stephanie brought her dog_{LARGE} to the party.
 B: Stephanie didn't bring her dog_{LARGE}—her dog is of **this_{SMALL}** size!
- c. A: What about Skyler_i? Did she_i come to the party?
 B: **She_i** didn't come to the party, because **they_i** don't identify as a woman.
- d. A: Zoe punched the extra_{PUNCH-HIGH}.
 B: Zoe didn't punch him_{PUNCH-HIGH}—the guy is shorter than her!

The cases above don't have to be viewed as counterexamples to the observation that inferences contributed by the content above have to project, since they can be analyzed under a general assumption that the focus in them is targeting the form, not the content of the expressions (see, e.g., Li 2017 for a recent version of such an analysis).

While usually such metalinguistic focus is used under negation, as oftentimes its purpose is to object to the form of something in the antecedent utterance (which is why it is also sometimes called CORRECTIVE FOCUS), it is in principle conceivable that it can be used in other environments, too. In particular, as I said at the end of subsection 3.4.3, metalinguistic focus could have been used by some participants in the Chapter 3 experiment as a “rescuing” strategy for ‘Non-projecting non-restricting’ items. Of course, if this strategy is available for ‘Non-projecting non-restricting’ items, there is no a priori reason why it shouldn't be available for ‘Restricting’ items as well, and it very well might be. However, the crucial contrast between the two types of items in the ‘Gesture’ condition still remains and needs to be explained.

The question of how analyses of metalinguistic focus, which have been so far based mostly on spoken content, extend to gestural and other non-spoken content is worth investigating in the future. Furthermore, there is a methodological benefit in understanding what licenses metalinguistic focus, so as to be able to distinguish cases thereof from cases of bona fide non-projection.

7.2.2 Projective content under ellipsis and in attitude reports

At the end of section 5.6, I have alluded to the possibility of distinguishing between the two potential analyses of height specifications on gestures, one that assimilates them to *phi*-features on pronouns and one that assimilates them to ordinary modifiers, by comparing the behavior of all these types of content under ellipsis and *only*. However, in order to actually be able to draw any conclusions based on such comparisons, we need to understand ellipsis and alternative generation under *only* better.

Schlenker & Chemla (2018) claim that both height specifications on gestures and *phi*-features on pronouns can be ignored in these environments. For example, under the sloppy readings of the examples in (7.2), the inferences that arise for Zoe don't arise for Skyler.

(7.2) *Context: Zoe and Skyler are practicing face punches with sparring partners.*

- a. *Zoe_i punched_{PUNCH-HIGH} her_i sparring partner, but Skyler didn't.*
- b. *Only Zoe_i punched_{PUNCH-HIGH} her_i sparring partner.*

For (a,b): $\not\rightarrow$ Skyler's sparring partner's face is higher than Skyler's.

For (a,b): $\not\rightarrow$ Skyler is female.

It seems that non-restricting modifiers, both spoken and gestural, can at least sometimes be ignored under ellipsis and *only*, too, as shown in (7.3). However, these judgements are often more gradient and variable, and ultimately, we would want to obtain quantitative data on all the four types of content.

(7.3) a. *Context: Stephanie and Lucy went to the same party yesterday.*

- (i) *Stephanie brought her ginormous dog to the party, but Lucy didn't.*
- (ii) *Only Stephanie brought her ginormous dog to the party.*
- (iii) *Stephanie brought her dog_{LARGE} to the party, but Lucy didn't.*
- (iv) *Only Stephanie brought her dog_{LARGE} to the party.*

For (i–iv): \rightarrow Lucy has a dog.

For (i–iv): $? \% \not\rightarrow$ Lucy's dog is large.

- b. *Context: Zoe and Kim are participating in a multi-sport competition. For the shooting part of the competition, participants had to choose a longbow or a gun and then shoot at the target.*
- (i) Zoe has already shot at the target_{LONGBOW}, but Kim hasn't.
 - (ii) Only Zoe has shot at the target_{LONGBOW} yet.
- $? \% \not\rightarrow$ If Kim shoots at the target, she'll shoot a longbow.

Note that all the examples in (7.3a) still come with an inference that Lucy has a dog, which shows that not all projecting inferences are ignorable under ellipsis and *only*, and, in particular, existence inferences of definites aren't. Many lexical presuppositions⁹⁸ cannot be ignored in these environments either, which is also noted in Schlenker & Chemla (2018). For example, (7.4) shows that the lexical presuppositions of *stop*, *know*, and *regret* cannot be ignored under ellipsis or *only* (modulo local accommodation, in which case the presupposition still isn't ignored, but rather is interpreted locally):

- (7.4) a. Jackie stopped smoking, but Daisy didn't.
→ Daisy used to smoke.
- b. Out of these three women, only Jackie stopped smoking.
→ Each of these three women used to smoke.
- c. O-Ren knows that she is in danger, but Vernita doesn't.
→ Vernita is in danger. (under the sloppy reading)
- d. Out of these three women only O-Ren knows that she is in danger.
→ Each of these three women is in danger. (under the sloppy reading)
- e. Kim regrets cheating on the exam, but Abernathy doesn't.
→ Abernathy cheated on the exam.
- f. Out of these three women only Kim regrets cheating on the exam.

⁹⁸It is not obvious to me that existence inferences of definites are lexical.

→ Each of these three women cheated on the exam.

Appositives, expressives, and slurs all can be ignored under ellipsis:⁹⁹

(7.5) Stephanie brought her dog, a large beast, to the party, but Lucy didn't.

↗ Lucy's dog is a large beast.

(7.6) A: I saw your fucking dog in the park.

B: No, you didn't—you couldn't have. The poor thing passed away last week.

↗ B {dislikes their dog, is angry, etc.}.

(Potts et al. 2009, (32))

(7.7) A: Why are you marrying that mudblood?

B: Because I want to.

↗ B is prejudiced against muggleborns.

Note that Potts et al. (2009) attribute the behavior of *fucking* in (7.6) to it being an expressive and contrast its behavior to that of an ordinary adjective in (7.8).

(7.8) A: I saw a shaggy dog in the park.

B: I did too. # {The one I saw, It} had no hair.

(Potts et al. 2009, (33))

However, *shaggy* in (7.8) is likely a restricting adjective. If we reproduce Potts et al.'s example with a non-restricting adjective, it improves, as shown in (7.9), which might suggest that it is not the expressive nature of *fucking*—or not only its expressive nature—that makes it so easily ignorable under ellipsis, but its aggressively non-restricting nature.

(7.9) A: I saw Stephanie's ginormous dog in the park.

B: And I saw Lucy's. ?Hers is small, though.

⁹⁹It is hard, if not impossible, to test the behavior of the latter two types of content under *only* because of their performative nature.

The data above call for a more systematic investigation of the behavior of different types of content under ellipsis and *only*. Such investigations can help us better understand not only different types of projective content, but ellipsis and alternative generation under *only* as such.

There is a further question of how the behavior of a given type of content under ellipsis or *only* relates to its behavior in attitude reports. At first glance, there seem to be non-trivial parallels between the two. As a first approximation, if p is a projective piece of content embedded under an intensional attitude verb, whether or not there is an inference that the attitude holder believes that p correlates with whether or not p can be ignored under ellipsis and *only*. Some examples corroborating this generalization are given in (7.10) and (7.11).

(7.10) **Types of content that can be ignored under ellipsis (and possibly *only*)**

- a. Uma wants Stephanie to bring her ginormous dog.
 $\not\rightarrow$ Uma believes that Stephanie's dog is ginormous.
- b. Uma wants Stephanie to bring her dog_{LARGE}.
 $\not\rightarrow$ Uma believes that Stephanie's dog is large.
- c. Uma wants Skyler_i to bring her_i dog.
 $\not\rightarrow$ Uma believes that Skyler is female.
- d. Uma wants Zoe to punch the extra in the face_{PUNCH-HIGH}.
 $\not\rightarrow$ Uma believes that the extra's face is higher than Zoe's.
- e. Uma wants Zoe to shoot at the target_{LONGBOW}.
 $\not\rightarrow$ Uma believes that if Zoe shoots at the target, she'll shoot a longbow.
- f. Uma wants Stephanie to bring her dog, a large beast.
 $\not\rightarrow$ Uma believes that Stephanie's dog is a large beast.
- g. Uma wants Stephanie to bring her fucking dog.
 $\not\rightarrow$ Uma believes that I, the speaker, {dislike Stephanie's dog, am angry, etc.}.
 $\not\rightarrow$ Uma herself {dislikes Stephanie's dog, is angry, etc.}.
- h. Uma wants to marry that mudblood.

- ↗ Uma believes that I, the speaker, am biased against muggleborns.
- ↗ Uma herself is biased against muggleborns.

(7.11) **Types of content that cannot be ignored under ellipsis (and possibly *only*)**

- a. Uma wants Jackie to stop smoking.
→ Uma believes that Jackie smokes.
- b. Uma wants O-Ren to know that she is in danger.
→ Uma believes that O-Ren is in danger.
- c. Uma wants Kim to regret cheating on the exam.
→ Uma believes that Kim cheated on the exam.

However, *de dicto/de re* distinctions complicate things when it comes to attitude reports, as shown in (7.12).

(7.12) Uma wants Stephanie to bring her dog.

De dicto construal: → Uma believes that Stephanie has a dog. (The dog doesn't have to exist in reality.)

De re construal: ↗ Uma believes that Stephanie has a dog; e.g., Uma wants Stephanie to bring her dog, which she believes to be a wolf. (The dog has to exist in reality, and Uma still has to believe that Stephanie has some animal.)

In fact, we could go back to the data (7.10) and (7.11) and see how much mileage we can get out of the *de dicto/de re* distinction alone, checking if in all the cases in which we don't get the inference that the attitude holder believes that *p*, we have a *de re* construal (see also Sudo 2012 for a relevant discussion for *phi*-features on pronouns). Arguably, no *de dicto/de re* distinctions emerge under ellipsis or *only*, though. Thus, the nature of the link between how a given type of content behaves under ellipsis or *only* and how it behaves in attitude reports remains opaque and calls for further research.

7.2.3 Composition and projection of tense and aspect

Tense and aspect typically give rise to projecting inferences about the relation between the topic time (i.e., the time the sentence is “about”; see Reichenbach 1947) and the utterance time, and the relation between the event time and the topic time, respectively, as shown in (7.13).

(7.13) I doubt that Mia was sleeping.

- The topic time of the embedded clause precedes the utterance time.
- The topic time of the embedded clause is contained within the event time.

I did not, however, mention tense or aspect when discussing modifiers and supplements in the verbal domain, because under all standard compositional analyses of tense and aspect I am familiar with, they are not treated as modifiers (i.e., pieces of content that combine with an expression of type τ yielding an expression of type τ). Instead, aspect is typically taken to introduce an additional argument slot, for a time, which is then filled in by tense (see, e.g., Kratzer 1998).

Despite this, does the approach to projection I have developed in this dissertation have anything to say about tense and aspect? In line with Partee’s (1973) analysis of tenses as pronouns, Heim (1994) proposes to assimilate contributions of tenses to those of *phi*-features on pronouns and, consequently, to treat them as (lexical) presuppositions.

Naturally, if tenses are pronouns, we can decompose them in the same way as personal pronouns, thus, also decomposing the black box that the presuppositions posited by Heim are. Like *phi*-features on personal pronouns, contributions of features like [past] contained within tenses will then be obligatorily non-restricting modifiers and will thus give rise to non-restricting modifier inferences. The structure of $PAST_i$ under this view and the (eventual) projecting inference it gives rise to are given in (7.14) (c_t is the utterance time, supplied by the context parameter c_p ; $<$ indicates temporal precedence).

(7.14) **Past tense in the proposed system**

a.

$$\begin{array}{c}
 \text{PAST}_i \\
 \lambda w.\iota t.(t < c_t) \wedge t = g(i) \\
 \swarrow \qquad \searrow \\
 \lambda T\lambda w.\iota t.T(t, w) \quad \lambda t\lambda w.(t < c_t) \wedge t = g(i) \\
 \swarrow \qquad \searrow \\
 [\text{past}] \qquad \lambda t\lambda w.t = g(i) \\
 \lambda t\lambda w.t < c_t
 \end{array}$$

b. non-restricting modifier inference:

$$\forall t\forall w[\llbracket c'' \rrbracket(w) \rightarrow (t = g(i) \rightarrow (t < c_t))]$$

However, as things stand, the same story doesn't extend to aspect, as it is not typically analyzed as a pronoun. Nor can it be analyzed as a preferably or obligatorily non-restricting modifier itself, as it is not a modifier. Furthermore, projection behavior of aspect is, to my knowledge, understudied, even though (7.13) shows that at least in simple cases, contributions of aspect project in the same way as contributions of tense. I think further research clarifying the empirical picture regarding projection of tense and aspect and relating their projection to their composition is in order.

Another related question for future research is how tense and aspect behave under ellipsis and what it tells us, if anything, about how they compose (and, thus, project). It would seem that, unlike *phi*-features on pronouns, tense and aspect cannot be ignored under ellipsis, as shown in (7.15), but pending a better understanding of how different types of content behave under ellipsis and how ellipsis works in the first place, we can't really draw any conclusions from that.

(7.15) Zoe was running. {Kim, too., Who else?}

→ The topic time of the second sentence precedes the utterance time.

→ The event time of the second sentence is contained within the topic time.

7.2.4 Speech–gesture integration as code-switching/code-blending

Throughout Chapter 4, and especially in subsection 4.3.4, I was raising the question of how gestures integrate into otherwise spoken utterances at various levels of representation. In particular, I made some preliminary observations suggesting that at least in some cases pro-speech (i.e., prosodically independent, compositionally integrated) gestures follow language-specific prosodic grouping, linearization, and displacement rules.

However, this does not have to be the case a priori. The problem of speech–gesture integration can be viewed as a problem of code-switching (in the case of pro-speech gestures) or code-blending (in the case of co-speech gestures).¹⁰⁰ When switching between two spoken languages, it is not always the case that the embedded expression follows the rules of the matrix language. For example, below I adduce some preliminary data on adjective placement in English–Spanish and English–French code-switching.¹⁰¹

(7.16) Adjective placement in English, Spanish, and French

- a. My dad is making {green peppers, #peppers green}. (English; A-N)
- b. Mi papà hace {pimientos verdes, *verdes pimientos}.
my dad makes peppers green green peppers
'My dad is making green peppers.' (Spanish; N-A)
- c. Mon père prépare des {piments verts, *verts piments}.
my dad makes INDEF.PL peppers green green peppers
'My dad is making green peppers.' (French; N-A)

(7.17) Adjective placement in English–Spanish code-switching

- a. Mi papà hace {??pimientos green, green pimientos}. (Sp → Eng; A-N)
- b. My dad is making {*verdes peppers, peppers verdes}. (Eng → Sp; N-A)

¹⁰⁰The term “code-blending” was coined in Emmorey et al. 2005 to refer to simultaneous production of speech and sign in bimodal bilinguals.

¹⁰¹I thank Paloma Jeretič, a French–English–Spanish trilingual, for providing the judgements and discussing the data with me. The judgements for the English–Spanish examples have been confirmed to me by another English–Spanish bilingual.

(7.18) Adjective placement in English–French code-switching

- a. Mon père prépare des {?piments green, ??green piments}. (Fr → Eng; N-A)
- b. My dad is making {marron potatoes, potatoes marron}.¹⁰²(Eng → Fr; A-N or N-A)

English has predominantly pre-nominal adjective placement, as illustrated in (7.16a), while Spanish and French have pre-dominantly post-nominal adjective placement, as illustrated in (7.16b) and (7.16c), respectively. An English adjective embedded into a Spanish utterance has to be placed pre-nominally and a Spanish adjective embedded into an English utterance has to be placed post-nominally, as shown in (7.17). In other words, in the English–Spanish language pair, the embedded adjective followed the rules of the embedded, but not the matrix language. In the English–French language pair, the judgements were much harder and seemed to depend a lot on prosodic factors. However, as shown in (7.18), we observed an overall tendency going in the opposite direction: an English adjective embedded into a French utterance was preferably post-nominal (so, it followed the rules of the matrix language, not the embedded one, unlike in the English-Spanish pair), but a French adjective embedded into an English utterance could be either pre-nominal or post-nominal, with potential subtle meaning differences, which we haven't explored yet.

At this point it's unclear what drives the differences between Spanish and French above, although it does seem that articulatory and prosodic integration seems to play an important role, and French rules of prosodic grouping and prominence assignment differ from English ones more so than Spanish ones do.¹⁰³ It is also worth noting that the example that was the hardest to judge, (7.18a), was the only one that had an overt determiner, which likely complicated articulatory integration. What these data do illustrate, however, is that linearization preferences in code-switching depend on the language pair.

What does this observation mean for speech–gesture integration? In line with the observations I made in subsection 4.3.4, a naïve first attempt to coach the problem of speech–gesture

¹⁰²*Marron* is French for ‘brown’. This example is different from the rest of the paradigm, because my consultant found it hard to prosodically integrate a monosyllabic French adjective into an English utterance, which further highlights the role of prosody in code-switching.

¹⁰³For one thing, English and Spanish both have lexical stress while French doesn't.

integration in code-switching terms could be to say that gestures follow the linearization rules of the matrix language because they don't have rules of their own. This claim, however, would be based on assumptions that need not be right and require further research.

For one thing, while gesture in non-signers might not be a fully grammaticalized system, there might still be gesture-specific preferences for linearization, prosodic grouping, displacement and the like, which would emerge in gesture-only utterances. For example, there has been work showing that speakers of languages with various relative placement of Subject, Object, and Verb exhibit uniform preferences for SOV or SVO orders (depending on the nature of the information conveyed) in gesture-only utterances (Gibson et al. 2013; Futrell et al. 2015; Marno et al. 2015, a.o.). I am not aware of any similar work trying to identify default preferences, if any, for modifier placement in gesture-only utterances, but research in this direction would be of great interest.

Second, the data reported in subsection 4.3.4 are but preliminary data for only three languages (English, French, Russian) and for size gestures only. Different patterns could emerge once we start looking at gestures embedded into spoken utterances in other languages as well as gestures that encode other types of meaning.

Third, we have seen that articulatory and prosodic integration are of utmost importance for both speech–gesture integration and for code-switching between two spoken languages with sufficiently different prosodic systems. The role of articulatory and prosodic considerations in both processes needs to be clarified before any claims are made about integration at other levels.

As hinted at the beginning of this subsection, research on speech–gesture integration could benefit most obviously from being informed by studies on speech–sign integration in bimodal bilinguals, in particular, when it comes to investigating the issues of prosodic and articulatory integration, constraints on code-blending, and differences between code-blending and code-switching. It is worth noting in this respect that it has been argued in the literature that speech–sign code-blending in bimodal bilinguals arises via a single, unified speech–sign system rather than switching between two systems (e.g., Lillo-Martin et al. 2016; de Quadros et al. To appear). It would be interesting to see to what extent the same view can be extended to co-speech gestures (which

I assimilated to code-blending at the beginning of this subsection) and how things are different, if at all, for pro-speech gestures (which I assimilated to code-switching at the beginning of this subsection).

7.2.5 Linearization constraints in multi-modal composition

The issue of linearization preferences raised in subsections 4.3.4 and 7.2.4 can and should be extended beyond integration of speech and hand gestures. At the end of the day, given a multi-modal compositional structure, we want to be able to predict how it can be linearized.

The problem is far from trivial, and modality-specific effects certainly do arise in this domain. For example, when spoken degree modifiers modify gestures, the latter can be linearized as pro-speech (modulo potential difficulties of articulatory integration), but not as co-speech, as shown in (7.19).

- (7.19) a. Stephanie's dog is {surprisingly, impressively, very, extremely} SMALL.
b. *Stephanie brought her {surprisingly, impressively, very, extremely} dog_{SMALL}.

In contrast, facial expressions, don't seem to be constrained in the same way, as shown in (7.20).¹⁰⁴

- (7.20) Stephanie brought her dog_{SMALL}^{O_O}.
✓ Stephanie's dog is small to a {surprising, impressive, high} extent.

I believe that investigating such constraints in a systematic way is a promising and important direction for future research within the general program of studying multi-modal integration at various levels of representation.

¹⁰⁴ Alternative construals of (7.20) might be possible, whereby *O_O* actually adjoins to *her dog* and is interpreted as a supplement along the lines of 'who is surprising' (it's unlikely that it adjoins to *dog* and is interpreted akin to the adjective *surprising*, because then we would expect a restricting reading to be available here, but it doesn't seem to be). We then pragmatically establish that what is surprising about Stephanie's dog is its size, encoded by the gesture, but there is no compositional link between the gesture and the facial expression.

Notational conventions

Below I list the notational conventions used throughout this dissertation.

1. Gestures and facial expressions.

- (a) Approximate word equivalents of gestures are written in all caps.
- (b) Gestures with their own time slot are written as part of the linear string, without any additional markers.
- (c) Co-speech gestures are written as subscripts, with underlining indicating their approximate temporal alignment.
- (d) Only mirative facial expressions are discussed in this dissertation. They are systematically written as O_O, iconically representing eyes wide open; potential differences among various versions of O_O (e.g., in mouth or head movements) are not captured in the notation.
- (e) Co-speech or co-gesture facial expressions are written as superscripts, with overlining indicating their approximate temporal alignment.
- (f) In some cases pictures are added to illustrate gestural content (hand gestures or facial expressions); they are systematically placed before the approximate onset of the expression they are illustrating. In examples cited from other people's work, original pictures are used whenever available; no pictures are added if the original examples don't contain them.

2. Prosody.

- (a) Vocal prosodic prominence associated with focus or contrastive topic marking (primarily

(L+)H* pitch accent for English) is indicated by bolding the word whose stressed syllable is prominent. Note that when a co-speech gesture is semantically in focus, it is the co-occurring spoken material that gets marked prosodically. Any production peculiarities of semantically focused co-speech gestures (e.g., kinetic emphasis) are not captured in the notation.

- (b) Gestural prosodic prominence associated with focus marking is only indicated for gestures with their own time slot. It usually involves more prominent, accelerated movement. A vocalization (iconic or not) might be added to facilitate prominence marking and/or articulatory integration of prominent gestural material into a spoken utterance.
- (c) Prosodic grouping is only indicated when it is particularly relevant. (PrP ...) is used to mark prosodic phrases, without further distinction among different types of prosodic phrases.

Glossary of terms

In this Appendix I define some of the recurrent terms that I use throughout this dissertation, in particular, those that (i) are new, (ii) are ambiguous, (iii) are not very well-known, (iv) are not used in a homogeneous way in the literature, or (v) are used non-standardly by me. I don't include commonplace terms that are used in the traditional way here.

- **ANCHOR.** An expression that a supplement composes with. For example, adnominal appositives compose with individuals as their anchors, sentence-level adverbs compose with propositions as their anchors, etc.
- **AT-ISSUE.** Following some of the previous literature, I sometimes use this term to describe a piece of content that doesn't project and is, thus, truth-conditionally non-vacuous.
- **CO-NOMINAL GESTURE.** A content-bearing gesture that co-occurs with and associates with a nominal spoken expression.
- **CO-SPEECH GESTURE.** A content-bearing gesture that co-occurs with and associates with a spoken expression.
- **CO-VERBAL GESTURE.** A content-bearing gesture that co-occurs with and associates with a verbal spoken expression.
- **COSUPPOSITION.** An assertion-dependent inference that projects like a PRESUPPOSITION. I propose to both expand and constrain cosuppositions by equating them with NON-RESTRICTING MODIFIER INFERENCES.
- **EXPRESSIVE.** A piece of content that (i) conveys information about the speaker's emotional state such that (ii) that emotional state doesn't have to be brought about by the denotation of what this piece of content adjoins to syntactically. For example *fucking, damn, or bloody* (used in their

non-literal sense) are expressives, but *lovely* isn't. I thus use the term more narrowly than Potts (2005).

- FORM INDEXICAL. An indexical interpreted with respect to an index within the context parameter that keeps track of the preferred/appropriate/etc. form of a given expression for a given individual/pair of individuals/etc.
- HEIGHT SPECIFICATION. The endpoint of a directional gesture whose relative height iconically represents some property of the event depicted and/or one of the event participants.
- LEXICAL PRESUPPOSITION. A SEMANTIC PRESUPPOSITION that is triggered as part of the lexical entry of a specific lexical item.
- LOCAL ACCOMMODATION. A last-resort mechanism posited for presuppositions that allows for them to be interpreted locally under semantic operators when they can't project for some reason.
- METALINGUISTIC FOCUS (a.k.a. CORRECTIVE FOCUS). Focus that targets the form rather than the content of an expression.
- MODIFIER. A piece of content that composes with an expression of type τ with the result of the composition being of type τ , too. However, throughout this dissertation I often use the term MODIFIER to mean SUBSECTIVE MODIFIER, as I don't discuss non-subsective modifiers. Furthermore, additional syntactic information might be needed to distinguish between modifiers and things they modify (especially, if we maintain that some modifiers compose with the things they modify intersectively).
- NON-RESTRICTING MODIFIER. A non-restricting modifier is a SUBSECTIVE MODIFIER that is intended by the speaker as not affecting the truth conditions of the utterance in which it appears (and is furthermore licensed by some discourse relation).
- NON-RESTRICTING MODIFIER INFERENCE. An inference contributed by a NON-RESTRICTING MODIFIER that the expression being modified entails the result of modification. I propose that such inferences project as PRESUPPOSITIONS, i.e., relative to local contexts, but are triggered pragmatically rather than semantically. I also equate COSUPPOSITIONS with non-restricting modifier inferences.

- NON-RESTRICTIVE. A non-restrictive piece of content is one that doesn't have a compositional potential to be restricting.
- NON-SUBLEXICAL. A non-sublexical piece of content is one that has its own node in the morphosyntax, as opposed to being part of a larger lexical entry.
- PRAGMATIC PRESUPPOSITION. A presupposition that is triggered via pragmatic reasoning on the speaker's beliefs or intentions. Cf. SEMANTIC PRESUPPOSITION.
- PRESUPPOSITION. Any inference that by default needs to be entailed by its local context, regardless of how this requirement is triggered.
- PRO-SPEECH GESTURE. A content-bearing gesture that has its own time slot and is compositionally integrated into an otherwise spoken utterance.
- PROJECTION. A piece of content that projects from under an operator (e.g., negation, question operator, conditional operator, modal, etc.) if it is interpreted outside that operator's semantic scope despite appearing to be in its syntactic scope.
- PROSODICALLY INDEPENDENT GESTURE. Gesture that has its own time slot and doesn't co-occur with any spoken content (but might co-occur with vocalizations, for iconicity-related and/or articulatory reasons). Prosodically independent gestures have to integrate into the speech stream prosodically on their own, in which they differ from CO-SPEECH GESTURES, which are prosodically parasitic on the spoken string they co-occur with.
- RESTRICTING. A restricting modifier is a SUBSECTIVE MODIFIER that (i) is not intended by the speaker as truth-conditionally vacuous, (ii) is used by the speaker to pick out a potentially non-empty part of the denotation of the expression it modifies.
- RESTRICTIVE. A restrictive piece of content is one that has a compositional potential to be restricting.
- SEMANTIC PRESUPPOSITION. A presupposition that is triggered semantically, e.g., as part of the lexical entry of a specific lexical item or a semantic composition rule. Cf. PRAGMATIC PRESUPPOSITION.
- STRONG. A strongly projecting inference is one that PROJECTS by default and cannot be inter-

preted locally even under pressure.

- **SUBSECTIVE.** A subsective modifier is a MODIFIER that composes with the expression it modifies β such that the output contains a conjunct that applies β to all the upcoming arguments. See (2.6) for a formal definition. An obligatory property of subsective modifiers that follows from this definition is SUBSECTIVE ENTAILMENT.
- **SUBSECTIVE ENTAILMENT.** An entailment relation that holds when the result of modification entails the expression being modified.
- **SUPPLEMENT.** A piece of content that composes with an ANCHOR and yields a proposition of a special kind about it. Nominal appositives, appositive relative clauses, sentence-level adverbs are supplements. Full-clause parentheticals aren't. I thus use this term more narrowly than Potts (2005).
- **WEAK.** A weakly projecting inference is one that PROJECTS by default, but can be interpreted locally under pressure relatively easily.

Instructions for Experiment in Chapter 3

Below I provide the instructions that the MTurk participants in the experiment from Chapter 3 saw before completing the study (some of the aspects of the formatting, such as, for example, the fonts, were not preserved). Please find all the videos used in the experiment, including the example videos from the instructions, here: <https://osf.io/fr5xt/>

Instructions

To complete this study you will need a computer with Internet connection and ability to play embedded videos, as well as headphones or speakers. Please make sure your screen is big enough to watch videos.

This study has 11 trials and should take about 8–10 minutes to complete. In each trial you will read a short context paragraph and then watch a video of a person uttering a sentence in English in this context. (Depending on your browser, you might have to click twice on the play button for the video to play.) Your task will be to assess how natural this sentence is in the given context by dragging a slider on a continuous scale from 'Totally unnatural' to 'Totally natural'. (If you want to leave the slider where it is, i.e., in the middle of the scale, you will need to at least click on it, otherwise you won't be allowed to proceed.)

Some of the sentences in the videos will contain gestures. Some of the contexts will be similar

or identical to each other. **Please pay close attention to the details of the contexts and to the gestures (when present) when assessing the naturalness of the sentences.**

Please note that there are no right or wrong answers here. When assessing the naturalness of the sentences you are expected to rely on your own linguistic intuitions rather than some prescriptive rules of “good English”.

Below are some examples of how to use the scale (your judgements might be different from those provided in the examples). Please note that the contexts and videos in the actual trials will not be analogous to the examples below.

Example A

Context: Anna and Maria are watching a movie. Anna just saw the main character Zoe punch a guy much taller than her and standing right in front of her in the face. Anna turns to Maria and says:



[Utterance in the video: Wow! Did you see her punch him the face_{PUNCH-HIGH}?]

For most speakers of English this is a coherent, natural sentence in the given context. If you are among those speakers, you might rate this sentence very high on the naturalness scale, for example:

Totally unnatural

Drag the slider

Totally natural

Example B

Context: Anna and Maria are watching a movie. Anna just saw the main character Zoe punch a guy much taller than her and standing right in front of her in the face. Anna turns to Maria and says:



[Utterance in the video: Wow! Did you see her punch him the face_{PUNCH-LOW}?]

The context in Example B is the same as in Example A, but Anna's gesture in the video is different; unlike the one in Example A, this gesture doesn't reflect the guy's height relative to Zoe's. Some speakers of English might find this sentence a bit odd in the given context. If you are among those speakers, you might rate this sentence somewhat lower on the naturalness scale, for example:

Totally unnatural

Drag the slider

Totally natural

Example C

Context: Anna and Maria are watching a movie. Anna just saw the main character Zoe punch a guy much shorter than her and standing right in front of her in the face. Anna turns to Maria and

says:



[Utterance in the video: Wow! Did you see her punch him the face_{PUNCH-HIGH}?]

The sentence in the video is the same as in Example A, but the context is different. Anna's gesture contradicts the information about the guy's height relative to Zoe's. Most speakers of English would probably find this sentence quite odd in the given context. If you are among those speakers, you might want to rate this sentence very low on the naturalness scale, for example:

Totally unnatural

Totally natural

Drag the slider

In order to get paid, please make sure that you complete all 11 trials.

Items for Experiment in Chapter 3

Below I list all the test items used in the experiment from Chapter 3. No illustrations are provided for the gestures. Please find all the videos used in the experiment here: <https://osf.io/fr5xt/>

Scenario type: ‘Brother’

Projecting non-restricting

(D1) *Context: We are planning a ‘Harry Potter’-themed event at which all the invitees are supposed to wear floor-length wizard robes. Anna and Maria are responsible for ordering the robes for everyone. Maria just told Anna that Simon, who has two siblings, a somewhat short sister and a very tall brother, has invited one of them to the event. Anna, who has seen both Simon’s siblings before, says:*

Do you know which one of Simon’s siblings will be joining us? ’Cause if he invited...

- a. his short **sister**
- b. his **sister**, a **short** person
- c. his **sister**_{SHORT}

..., we’re good, but if he invited...

- a. his tall **brother**
- b. his **brother**, a **tall** person
- c. his **brother**_{TALL}

..., we should get a longer robe for him.

Restricting

(D2) *Context: We are planning a ‘Harry Potter’-themed event at which all the invitees are supposed to wear floor-length wizard robes. Anna and Maria are responsible for ordering the robes for everyone. Maria just told Anna that Simon, who has two brothers, a somewhat short guy and a very tall one, has invited one of them to the event. Anna, who has seen both Simon’s brothers before, says:*

Do you know which one of Simon’s brothers will be joining us? ’Cause if he invited...

- a. his **short** brother
- b. his brother, a **short** guy
- c. his brother_{SHORT}

..., we’re good, but if he invited...

- a. his **tall** brother
- b. his brother, a **tall** guy
- c. his brother_{TALL}

..., we should get a longer robe for him.

Non-projecting non-restricting

(D3) *Context: We are planning a ‘Harry Potter’-themed event at which all the invitees are supposed to wear floor-length wizard robes. Anna and Maria are responsible for ordering the robes for everyone. Maria just told Anna that Simon has invited his brother to the event. Anna knows that Simon only has one brother, but has never seen him. She says:*

Do you know how tall Simon’s brother is? ’Cause if he invited...

- a. his **short** brother
- b. his brother, a **short** guy
- c. his brother_{SHORT}

..., we're good, but if he invited...

- a. his **tall** brother
- b. his brother, a **tall** guy
- c. his brother_{TALL}

..., we should get a longer robe for him.

Scenario type: ‘Car’

Projecting non-restricting

(D4) *Context: Maria just told Anna that their friend Kate is coming to visit New York for a few days. Kate will be driving from Massachusetts and is looking for a place where to leave her car for while she's in New York. **Kate has two cars, a small smart car and a large SUV. Anna, who has seen both Kate's cars before**, says:*

Do you know which one of her cars Kate will be using to get here? 'Cause if she's coming in...

- a. her small **smart** car
- b. her **smart** car, a **small** vehicle
- c. her smart car_{SIMILAR}

..., it'll fit in my garage, but if she's coming in...

- a. her large **SUV**
- b. her **SUV**, a **large** vehicle
- c. her SUV_{LARGE}

..., she'll have to look elsewhere.

Restricting

(D5) *Context: Maria just told Anna that their friend Kate is coming to visit New York for a few days. Kate will be driving from Massachusetts and is looking for a place where to leave her car for while she's in New York. **Kate has two cars, a small smart car and a large SUV.** Anna, who has seen both Kate's cars before, says:*

Do you know which one of her cars Kate will be using to get here? 'Cause if she's coming in...

- a. her **small** car
- b. her car, a **small** vehicle
- c. her car_{SMALL}

..., it'll fit in my garage, but if she's coming in...

- a. her **large** car
- b. her car, a **large** vehicle
- c. her car_{LARGE}

..., she'll have to look elsewhere.

Non-projecting non-restricting

(D6) *Context: Maria just told Anna that their friend Kate is coming to visit New York for a few days. Kate will be driving from Massachusetts and is looking for a place where to leave her car for while she's in New York. **Anna knows that Kate only has one car, but has never seen it.** She says:*

Do you know how big Kate's car is? 'Cause if she's coming in...

- a. her **small** car
- b. her car, a **small** vehicle
- c. her car_{SMALL}

..., it'll fit in my garage, but if she's coming in...

- a. her **large** car
- b. her car, a **large** vehicle
- c. her car_{LARGE}

..., she'll have to look elsewhere.

Scenario type: 'Dog'

Projecting non-restricting

- (D7) *Context: We are going on a group tour. Anna and Maria are responsible for renting a van. Maria just told Anna that Stephanie, who has two pets, a small cat and a large dog, is planning to bring along one of her pets. Anna, who has seen both Stephanie's pets before, says:*

Do you know which one of Stephanie's pets is coming with us? 'Cause if she's bringing...

- a. her small **cat**
- b. her **cat**, a **small** animal
- c. her cat_{SMALL}

..., we'll be fine, but if she's bringing...

- a. her large **dog**
- b. her **dog**, a **large** animal
- c. her dog_{LARGE}

..., we should get a bigger van.

Restricting

- (D8) *Context: We are going on a group tour. Anna and Maria are responsible for renting a van. Maria just told Anna that Stephanie, who has two dogs, a small Pug and a large Great Dane, is planning to bring along one of her dogs. Anna, who has seen both Stephanie's*

dogs before, says:

Do you know which one of Stephanie's dogs is coming with us? 'Cause if she's bringing...

- a. her **small** dog
- b. her dog, a **small** animal
- c. her **dog**_SMALL

..., we'll be fine, but if she's bringing...

- a. her **large** dog
- b. her dog, a **large** animal
- c. her **dog**_LAUGE

..., we should get a bigger van.

Non-projecting non-restricting

(D9) *Context: We are going on a group tour. Anna and Maria are responsible for renting a van. Maria just told Anna that Stephanie is planning to bring along her dog. Anna knows that Stephanie only has one dog, but has never seen it. She says:*

Do you know which one of Stephanie's dogs is coming with us? 'Cause if she's bringing...

- a. her **small** dog
- b. her dog, a **small** animal
- c. her **dog**_SMALL

..., we'll be fine, but if she's bringing...

- a. her **large** dog
- b. her dog, a **large** animal
- c. her **dog**_LAUGE

..., we should get a bigger van.

Scenario type: ‘Dog’

Projecting non-restricting

- (D10) *Context: Maria just told Anna that their friend Paloma, who is an amateur musician, is going to France for a semester. Paloma has a small electric piano and a large tuba. She wants to bring one of her instruments with her to France so that she can practice occasionally, but she doesn’t know if she can fly with it. Anna, who has seen both Paloma’s instruments before, says:*

Do you know which one of her instruments Paloma wants to take with her? ’Cause if she’s flying with...

- a. her small **piano**
- b. her **piano**, a **small** instrument
- c. her piano_{SMALL}

..., she should be able to take it as carry-on, but if she’s flying with...

- a. her large **tuba**
- b. her **tuba**, a **large** instrument
- c. her tuba_{LARGE}

..., she’ll have to check it in.

Restricting

- (D11) *Context: Maria just told Anna that their friend Paloma, who is an amateur pianist, is going to France for a semester. Paloma has two electric pianos, a small portable one and a large full-size one. She wants to bring one of her pianos with her to France so that she can practice occasionally, but she doesn’t know if she can fly with it. Anna, who has seen both Paloma’s pianos before, says:*

Do you know which one of her pianos Paloma wants to take with her? ’Cause if she’s

flying with...

- a. her **small** piano
- b. her piano, a **small** instrument
- c. her piano_{SMALL}

..., she should be able to take it as carry-on, but if she's flying with...

- a. her **large** piano
- b. her piano, a **large** instrument
- c. her piano_{LARGE}

..., she'll have to check it in.

Non-projecting non-restricting

(D12) *Context: Maria just told Anna that their friend Paloma, who is an amateur pianist, is going to France for a semester. Paloma wants to bring her electric piano to France so that she can practice occasionally, but she doesn't know if she can fly with it. Anna knows that Paloma only has one piano, but has never seen it. She says:*

Do you know how big Paloma's piano is? 'Cause if she's flying with...

- a. her **small** piano
- b. her piano, a **small** instrument
- c. her piano_{SMALL}

..., she should be able to take it as carry-on, but if she's flying with...

- a. her **large** piano
- b. her piano, a **large** instrument
- c. her piano_{LARGE}

..., she'll have to check it in.

Bibliography

- Abrusán, Márta. 2011. Predicting the presuppositions of soft triggers. *Linguistics and Philosophy* 34(6). 491–535. doi: 10.1007/s10988-012-9108-y.
- AnderBois, Scott, Adrian Brasoveanu & Robert Henderson. 2013. At-issue proposals and appositive impositions in discourse. *Journal of Semantics* 32(1). 93–138. doi: 10.1093/jos/fft014.
- Anvari, Amir. 2018. Logical Integrity: from *Maximize Presupposition!* to *Mismatching Implications*. Ms., Institut Jean-Nicod. URL <https://ling.auf.net/lingbuzz/003866>.
- Aristodemo, Vita Maria Valentina. 2017. *Gradable constructions in Italian Sign Language*: Institut Jean-Nicod dissertation. URL http://www.theses.fr/2017EHES0147/abes/Aristodemo-Vita_Maria_Valentina_these_2017.pdf.
- Beaver, David I. 2001. *Presupposition and assertion in dynamic semantics*. Stanford, CA: CSLI publications.
- Beckman, Mary E. & Gayle Ayers. 1997. Guidelines for ToBI labelling. Version 3.0. The Ohio State University Research Foundation.
- Carlson, Gregory N. 1984. Thematic roles and their role in semantic interpretation. *Linguistics* 22(3). 259–280. doi: 10.1515/ling.1984.22.3.259.
- Cepollaro, Bianca. 2017. *The semantics and pragmatics of slurs and thick terms*: PSL Research University dissertation. URL <https://tel.archives-ouvertes.fr/tel-01508856/document>.

Chemla, Emmanuel & Lewis Bott. 2013. Processing presuppositions: Dynamic semantics vs pragmatic enrichment. *Language and Cognitive Processes* 28(3). 241–260. doi: 10.1080/01690965.2011.615221.

Cooper, Robin. 1983. *Quantification and syntactic theory*. Dordrecht: Reidel. doi: 10.1007/978-94-015-6932-3.

Davidson, Kathryn. 2015. Quotation, demonstration, and iconicity. *Linguistics and philosophy* 38(6). 477–520. doi: 10.1007/s10988-015-9180-1.

Delais-Roussarie, Elisabeth, Brechtje Post, Mathieu Avanzi, Caroline Buthke, Albert Di Cristo, Ingo Feldhausen, Sun-Ah Jun, Philippe Martin, Trudel Meisenburg, Annie Rialland, Rafèu Sichel-Bazin & Hi-Yon Yoo. 2015. Intonational phonology of French: Developing a ToBI system for French. In Sónia Frota & Pilar Prieto (eds.), *Intonation in Romance*, Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199685332.003.0003.

Duncan, Daniel. 2019. The role of gesture in the English *ish*-construction. *Proceedings of the Linguistic Society of America* 4(1). 1–12. doi: 10.3765/plsa.v4i1.4462.

Ebert, Cornelia. 2017. Co-speech vs. post-speech gestures. Talk given at *Language and cognition workshop in memory of Peter Bosch*, Osnabrück, February.

Ebert, Cornelia & Christian Ebert. 2014. Gestures, demonstratives, and the attributive/referential distinction. Talk given at *Semantics and Philosophy in Europe (SPE 7)*, Berlin, June.

Elbourne, Paul D. 2005. *Situations and individuals*. Cambridge, MA: MIT Press.

Emmorey, Karen, Helsa B Borinstein & Robin Thompson. 2005. Bimodal bilingualism: Code-blending between spoken English and American Sign Language. In James Cohen, Kara T McAlister, Kellie Rolstad & Jeff MacSwan (eds.), *Proceedings of the 4th International Symposium on Bilingualism*, 663–673. Somerville, MA: Cascadilla Press.

Esipova, Maria. 2017. Co- and post-speech gestures: A prosody/syntax approach. Talk given at the workshop *Week of Signs and Gestures*, University of Stuttgart. URL <https://osf.io/e8syf/>.

Esipova, Maria. 2018a. Focus on what's not at issue: Gestures, presuppositions, supplements under contrastive focus. In Uli Sauerland & Stephanie Solt (eds.), *Proceedings of Sinn und Bedeutung 22*, 385–402. Berlin: Leibniz-Centre General Linguistics.

Esipova, Maria. 2018b. Gestural adjuncts at the interfaces. Poster presented at the workshop *Linguistic Explorations Beyond Language*, University of Crete. URL <https://osf.io/tr9kg/>.

Esipova, Maria. 2018c. Gestures at the interfaces. Talk given at the workshop *Foundational Topics in Semantics*, New York University. URL <https://osf.io/9dp36/>.

Esipova, Maria. 2019a. Acceptability of at-issue co-speech gestures under contrastive focus. *Glossa: A Journal of General Linguistics* 4(1). 11. doi: 10.5334/gjgl.635.

Esipova, Maria. 2019b. Composition and projection of co-speech gestures. Poster presented at *Semantics and Linguistic Theory (SALT) 29*, UCLA. URL <https://osf.io/8x75s/>.

Esipova, Maria. 2019c. Composition determines projection across modalities. Talk given at the workshop *Linguistic Investigations Beyond Language: Gestures, Body Movement and Primate Linguistics*, ZAS Berlin. URL <https://osf.io/pm287/>.

Esipova, Maria. 2019d. Polar responses in Russian across modalities and across interfaces. Talk given at *Formal Approaches to Slavic Linguistics (FASL) 28*, Stony Brook University. URL <https://osf.io/gr9pe/>.

Futrell, Richard, Tina Hickey, Aldrin Lee, Eunice Lim, Elena Luchkina & Edward Gibson. 2015. Cross-linguistic gestures reflect typological universals: A subject-initial, verb-final bias in speakers of diverse languages. *Cognition* 136. 215–221. doi: 10.1016/j.cognition.2014.11.022.

Geurts, Bart. 1996. Local satisfaction guaranteed: A presupposition theory and its problems. *Linguistics and Philosophy* 19(3). 259–294. doi: 10.1007/BF00628201.

- Gibson, Edward, Steven T Piantadosi, Kimberly Brink, Leon Bergen, Eunice Lim & Rebecca Saxe. 2013. A noisy-channel account of crosslinguistic word-order variation. *Psychological Science* 24(7). 1079–1088. doi: 10.1177/0956797612463705.
- Grice, H. Paul. 1975. Logic and conversation. In Peter Cole & Jerry Morgan (eds.), *Syntax and semantics*, vol. 3: Speech acts, 41–58. New York: Academic Press.
- Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Kenneth Hale & S. Jay Keyser (eds.), *The view from Building 20*, 111–176. Cambridge, MA: MIT Press.
- Heim, Irene. 1983. On the projection problem for presuppositions. In Michael Barlow, Daniel Flickinger & Michael Wescoat (eds.), *Proceedings of the Second West Coast Conference on Formal Linguistics*, 114–125.
- Heim, Irene. 1991. Artikel und Definitheit. In Arnim von Stechow & Dieter Wunderlich (eds.), *Semantics: An international handbook of contemporary research*, 487–535. Berlin: Walter de Gruyter.
- Heim, Irene. 1994. Comments on Abusch’s theory of tense. In Hans Kamp (ed.), *Ellipsis, tense and questions*, 143–170. Amsterdam: University of Amsterdam.
- Heim, Irene & Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell.
- Hunter, Julie. 2018. Relating gesture to speech: reflections on the role of conditional presuppositions. *Linguistics and Philosophy* 1–16. doi: 10.1007/s10988-018-9244-0.
- Hunter, Julie. 2019. Relating gesture to speech. Talk given at the workshop *Linguistic Investigations Beyond Language: Gestures, Body Movement and Primate Linguistics*, ZAS Berlin.
- Jasinskaja, Katja & Claudia Poschmann. 2018. Attachment in syntax and discourse: towards an explanation for the variable scope of non-restrictive relative clauses. In Sireemas Maspong,

Brynhildur Stefánsdóttir, Katherine Blake & Forrest Davis (eds.), *Proceedings of Semantics and Linguistic Theory (SALT) 28*, 433–453. doi: 10.3765/salt.v28i0.4438.

Karttunen, Lauri. 1974. Presuppositions and linguistic context. *Theoretical Linguistics* 1. 181–194. doi: 10.1515/thli.1974.1.1-3.181.

Kennedy, Christopher & Louise McNally. 2005. Scale structure, degree modification, and the semantics of gradable predicates. *Language* 81(2). 345–381. doi: 10.1353/lan.2005.0071.

Koev, Todor. 2013. *Apposition and the structure of discourse*: Rutgers University dissertation. URL <https://ling.rutgers.edu/images/dissertations/Koev-2013-Dissertation-Final-Version.pdf>.

Kratzer, Angelika. 1998. More structural analogies between pronouns and tenses. In Devon Strolovitch & Aaron Lawson (eds.), *Proceedings of Semantics and Linguistic Theory (SALT) 8*, 92–110. Ithaca, NY: Cornell University. doi: 10.3765/salt.v8i0.2808.

Larson, Richard. 1991. The projection of DP (and DegP). Ms., Stony Brook University.

Larson, Richard & Franc Marušič. 2004. On indefinite pronoun structures with APs: Reply to Kishimoto. *Linguistic Inquiry* 35(2). 268–287. doi: 10.2307/4179274.

Lee, Naomi. 2019. Khoekhoe pronominal morphosyntax: Gender on Root-attached little *n*. Talk given at *The 93rd Annual Meeting of the LSA*, New York City.

Leffel, Timothy. 2014. *The semantics of modification: Adjectives, nouns, and order*: New York University dissertation. URL <https://ling.auf.net/lingbuzz/002212>.

Lewis, David. 1979. Scorekeeping in a language game. *Journal of Philosophical Logic* 8. 339–359. doi: 10.1007/BF00258436.

Li, Haoze. 2017. Semantics of metalinguistic focus. In Alexandre Cremers, Thom van Gessel & Floris Roelofsen (eds.), *21st Amsterdam Colloquium*, 345–354.

Lillo-Martin, Diane, Ronice Müller de Quadros & Deborah Chen Pichler. 2016. The development of bimodal bilingualism. *Linguistic Approaches to Bilingualism* 6(6). 719–755. doi: 10.1075/lab.6.6.01lil.

Loehr, Daniel P. 2004. *Gesture and intonation*: Georgetown University dissertation.

Marno, Hanna, Alan Langus, Mahmoud Omidbeigi, Sina Asaadi, Shima Seyed-Allaei & Marina Nespor. 2015. A new perspective on word order preferences: the availability of a lexicon triggers the use of SVO word order. *Frontiers in Psychology* 6. 1183. doi: 10.3389/fpsyg.2015.01183.

Matushansky, Ora. 2008. On the linguistic complexity of proper names. *Linguistics and Philosophy* 31(5). 573–627. doi: 10.1007/s10988-008-9050-1.

McCawley, James D. 1998. *The syntactic phenomena of English*, vol. 2. Chicago IL: The University of Chicago Press.

Morzycki, Marcin. 2008. Nonrestrictive modifiers in nonparenthetical positions. In Christopher Kennedy & Louise McNally (eds.), *Adjectives and adverbs: Syntax, semantics, and discourse* Studies in Theoretical Linguistics, 101–122. Oxford: Oxford University Press.

Morzycki, Marcin. 2015. *Modification*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511842184.

Nouwen, Rick. 2014. A note on the projection of appositives. In Eric McCready, Katsuhiko Yabushita & Kei Yoshimoto (eds.), *Formal Approaches to Semantics and Pragmatics*, 205–222. Dordrecht: Springer. doi: 10.1007/978-94-017-8813-7_10.

Parsons, Terence. 1990. *Events in the semantics of English*. Cambridge, MA: MIT Press.

Partee, Barbara & Mats Rooth. 1983. Generalized conjunction and type ambiguity. In Rainer Bäuerle, Christoph Schwarze & Arnim von Stechow (eds.), *Meaning, use and interpretation of language*, 361–383. Berlin: de Gruyter. doi: 10.1515/9783110852820.361.

Partee, Barbara Hall. 1973. Some structural analogies between tenses and pronouns in english. *The Journal of Philosophy* 70(18). 601–609. doi: 10.2307/2025024.

Postal, Paul. 1966. On so-called pronouns in English. In Francis P Dinneen (ed.), *Report on the 17th Annual Round Table Meeting of Language and Linguistics Studies*, 177–206. Washington, D.C.: Georgetown University Press.

Potts, Christopher. 2005. *The logic of conventional implicatures*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199273829.001.0001.

Potts, Christopher. 2007. The expressive dimension. *Theoretical linguistics* 33(2). 165–198. doi: 10.1515/TL.2007.011.

Potts, Christopher, Ash Asudeh, Seth Cable, Yurie Hara, Eric McCready, Luis Alonso-Ovalle, Rajesh Bhatt, Christopher Davis, Angelika Kratzer, Tom Roeper et al. 2009. Expressives and identity conditions. *Linguistic Inquiry* 40(2). 356–366. doi: 10.1162/ling.2009.40.2.356.

de Quadros, Ronice, Kathryn Davidson, Diane Lillo-Martin & Kathryn Emmorey. To appear. Code-blending with depicting signs. *Linguistic Approaches to Bilingualism*.

R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing Vienna, Austria. URL <https://www.R-project.org>.

Reichenbach, Hans. 1947. *Elements of symbolic logic*. Macmillan Co.

Roberts, Craige. 2012. Information structure: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics* 5(6). 1–69. doi: 10.3765/sp.5.6.

Schlenker, Philippe. 2005. *Minimize restrictors!* (Notes on definite descriptions, condition C and epithets). In Emar Maier, Corien Bary & Janneke Huitink (eds.), *Proceedings of Sinn und Bedeutung* 9, 385–416.

Schlenker, Philippe. 2007. Expressive presuppositions. *Theoretical Linguistics* 33(2). 237–245. doi: 10.1515/TL.2007.017.

Schlenker, Philippe. 2009. Local contexts. *Semantics and Pragmatics* 2(3). 1–78. doi: 10.3765/sp.2.3.

Schlenker, Philippe. 2010. Presuppositions and local contexts. *Mind* 119(474). 377–391. doi: doi:10.1093/mind/fzq032.

Schlenker, Philippe. 2013. Supplements without bidimensionalism. Ms., Institut Jean-Nicod and New York University. URL <http://www.semanticsarchive.net/Archive/jgwMjNmM/>.

Schlenker, Philippe. 2018a. Gesture projection and cosuppositions. *Linguistics and Philosophy* 41(3). 295–365. doi: 10.1007/s10988-017-9225-8.

Schlenker, Philippe. 2018b. Iconic pragmatics. *Natural Language & Linguistic Theory* 36(3). 877–936. doi: 10.1007/s11049-017-9392-x.

Schlenker, Philippe. 2018c. Iconic presuppositions. Ms., Institut Jean-Nicod and New York University. URL <https://ling.auf.net/lingbuzz/004116>.

Schlenker, Philippe. 2018d. Indexicals. In Sven Ove Hansson & Vincent F. Hendricks (eds.), *Introduction to formal philosophy*, 297–321. Springer. doi: 10.1007/978-3-319-77434-3_14.

Schlenker, Philippe. 2018e. Visible meaning: Sign language and the foundations of semantics. *Theoretical Linguistics* 44(3–4). 123–208. doi: 10.1515/tl-2018-0012.

Schlenker, Philippe. To appear. The semantics and pragmatics of appositives. In Lisa Matthewson, Cécile Meier, Hotze Rullmann & Thomas Ede Zimmermann (eds.), *Blackwell companion to semantics*, Wiley. URL <http://ling.auf.net/lingbuzz/002538>.

Schlenker, Philippe & Emmanuel Chemla. 2018. Gestural agreement. *Natural Language & Linguistic Theory* 36(2). 587–625. doi: 10.1007/s1104.

Selkirk, Elisabeth. 2005. Comments on intonational phrasing in English. In Sónia Frota, Marina Vigário & Maria João Freitas (eds.), *Prosodies: with special reference to Iberian languages*, 11–58. Mouton de Gruyter.

- Simons, Mandy, Judith Tonhauser, David Beaver & Craige Roberts. 2010. What projects and why. In Nan Li & David Lutz (eds.), *Proceedings of SALT 20*, 309–327. doi: 10.3765/salt.v20i0.2584.
- Stalnaker, Robert. 1974. Pragmatic presuppositions. In Milton Kunitz & Peter Unger (eds.), *Semantics and philosophy*, 197–213. New York University Press.
- Stalnaker, Robert. 1975. Indicative conditionals. *Philosophia* 5(3). 269–286. doi: 10.1007/BF02379021.
- Starke, Michal. 2010. Nanosyntax: A short primer to a new approach to language. *Nordlyd* 36(1). 1–6. doi: 10.7557/12.213.
- Sudo, Yasutada. 2012. *On the semantics of phi features on pronouns*: Massachusetts Institute of Technology dissertation. URL <https://dspace.mit.edu/bitstream/handle/1721.1/77805/828407318-MIT.pdf>.
- Szabolcsi, Anna. 1980. Az aktuális mondattagolás szemantikájához. *Nyelvtudományi Közlemények* 82. 59–83.
- Tieu, Lyn, Robert Pasternak, Philippe Schlenker & Emmanuel Chemla. 2017. Co-speech gesture projection: Evidence from truth-value judgment and picture selection tasks. *Glossa: A Journal of General Linguistics* 2(1). 102. doi: 10.5334/gjgl.334.
- Tieu, Lyn, Robert Pasternak, Philippe Schlenker & Emmanuel Chemla. 2018. Co-speech gesture projection: Evidence from inferential judgments. *Glossa: A Journal of General Linguistics* 3(1). 109. doi: 10.5334/gjgl.580.
- Truckenbrodt, Hubert. 1999. On the relation between syntactic phrases and phonological phrases. *Linguistic inquiry* 30(2). 219–255. doi: 10.1162/002438999554048.
- Umbach, Carla. 2006. Non-restrictive modification and backgrounding. In Beáta Gyuris, Kálmán László, Cristoph Piñón & Károly Varasdi (eds.), *Proceedings of the Ninth Symposium on Logic and Language*, 152–159. Budapest: Hungarian Academy of Sciences.

Wang, Linton, Brian Reese & Eric McCready. 2005. The projection problem of nominal appositives. *Snippets* 10(1). 13–14.

Yanovich, Igor. 2010. On the nature and formal analysis of indexical presuppositions. In Kumiyo Nakakoji, Yohei Murakami & Eric McCready (eds.), *New Frontiers in Artificial Intelligence. JSAI-isAI 2009. Lecture Notes in Computer Science*, 272–291. Springer, Berlin, Heidelberg. doi: 10.1007/978-3-642-14888-0_22.

Zlogar, Christina & Kathryn Davidson. 2018. Effects of linguistic context on the acceptability of co-speech gestures. *Glossa: A Journal of General Linguistics* 3(1). 73. doi: 10.5334/gjgl.438/.