COMPOSING QUESTIONS

COMPOSING QUESTIONS

Hadas Kotek

The MIT Press
Cambridge, Massachusetts
London, England

Contents

List of Figures

List of Tables

Notation

Semantic types

e	Individual
t	Truth-value
s	World
σ, τ	Arbitrary semantic types
$\langle \sigma, \tau \rangle$	A function from $D_\sigma \rightarrow D_\tau$

Functions

$[\![\cdot]\!]$, $[\![\cdot]\!]^o$	The ordinary semantic value function
$[\![\cdot]\!]^f$	The focus-semantic value function

Judgments

✓	Acceptable (presumed grammatical)
$*$	Unacceptable (presumed ungrammatical)
$*^{\text{PL}}$	Does not have a pair-list reading

Examples and trees

italic	*Wh*-word
bold	Intervener
X_1	Element base-generated in a higher position
X_2	Element base-generated in a lower position
X *t*	Overt movement
X *t*	Covert movement
X *t*	Area of in-situ composition

Glosses

ABS	absolutive
ACC	accusative
ASP	aspect
AUX	auxiliary
C	complementizer
CL	classifier
CONJ	conjunction
DAT	dative
DEF	definite
EMPH	emphatic
ERG	ergative
EXPL	expletive
F	feminine
FUT	future
IMPF	imperfective
M	masculine
NEG	negation
NOM	nominative
PL	plural
PREP	preposition
PRES	present
PRT	particle
PST	past
Q	question particle
REL.PRO	relative pronoun
SG	singular
VM	verbal marker

Acknowledgements

This book has benefited from the continued support and input of many friends and colleagues over several years. Some made passing observations; others provided detailed commentary. Many had to tolerate repeated queries for complex data and judgments. During its five year development, from early drafts of my dissertation to its current book form, this work has gone through several versions and underwent quite a few revisions. If it is any closer to correct now than it was at the beginning, no doubt the following people deserve the credit: My MIT dissertation committee – David Pesetsky, Martin Hackl, Danny Fox, and Irene Heim; senior colleagues at Yale and NYU – especially Bob Frank, Raffaella Zanuttini, Lucas Champollion, Chris Collins, Stephanie Harves, and Anna Szabolsci; and my many MIT friends and colleagues who supported me throughout graduate school. Special thanks go to Chris Tancredi for detailed comments on an earlier draft of the book, and to Michael Yoshitaka Erlewine for all kinds of support and advice, as well as for detailed comments.

Parts of the dissertation and book manuscript were taught in seminars at McGill University (Fall 2014), Yale University (Fall 2016), and NYU (Fall 2017). Participants in all of these classes made invaluable contributions to my thinking and to the development of this work from dissertation-form to book-form. Parts of this work were also presented at various conferences and colloquia, whose audiences I would like to thank: LSA 87 in Boston (2012), NELS 44 at the University of Connecticut (2013), LSA 88 in Minneapolis (2013), LSA 89 in Portland (2014), DGfS 37 workshop in Leipzig (2015), SuB 20 in Tübingen (2015), IATL 31 at Bar-Ilan University (2015); The Tokyo Semantics Research Group (2012), Tel Aviv University (2014), McGill University (2014), University of Calgary (2014), University of Ottawa (2015), University of Arizona at Tucson (2015), Queen Mary University of London (2015), CUNY graduate center (2015), UC Irvine (2017), and the MIT panel on questions at Pesetsky@60 (2017).

This book is a modified and extended version of my dissertation, Kotek (2014a). Chapters 6 and 7 are revised versions of work that was reported in Kotek (2017a) and Kotek (2016), respectively. Chapter 8 is an extended version of joint work in Kotek and Hackl (2013). The book's title, as the dissertation title, is a play on Hagstrom's (1998) dissertation title 'Decomposing Questions,' which served as inspiration for this work. Parts of this work were supported by NSF Dissertation Improvement Grant #1251717 in 2013–2014 and by a Mellon Postdoctoral Fellowship in the years 2014–2016.

8 Processing evidence for covert scrambling

In this chapter I provide novel experimental evidence in support of the view of covert *wh*-movement as covert scrambling.[1] I begin by contrasting the predictions made by a theory in which *wh*-phrases undergo covert *wh*-movement at LF with one where they remain in-situ. As ample off-line evidence has shown, we expect to find evidence for covert movement of the in-situ *wh*-phrase in superiority-obeying English questions, but not in superiority-violating questions.

Once I establish the methodology and basic processing signature of movement, I then use the intervention effect diagnostic to test the nature of this movement more closely. In particular, as section 7.2 spelled out, if the covert *wh*-scrambling proposal is on the right track, we expect the size of movement to correlate with the presence and height of interveners: a higher intervener should require a longer movement step than a lower intervener in the structure. I show that this prediction is indeed borne out.

I conclude this chapter by briefly entertaining the source of short *wh*-scrambling: by combining the Hamblin-set approach to *wh*-expressions with Karttunen-based approaches which view *wh*-expressions as quantifiers, we may be able to combine the results developed throughout this book into a single coherent proposal. The denotations of *wh*-expressions that I eventually propose, then, consist of *a set of alternative quantificational interpretations* (i.e. of alternative $\langle et, t\rangle$ denotations).

8.1 Three approaches to covert movement

As we have seen throughout in this book, the literature on the syntax-semantics of *wh*-questions provides two conceptual approaches to the interpretation of in-situ *wh*-phrases: covert movement and in-situ composition. Under the covert movement approach, all *wh*-phrases must be structurally adjacent to the head that interprets them in the CP periphery. The in-situ approach invokes a mechanism that interprets *wh*-phrases without any movement. I have argued in this book for a combined approach, where covert *wh*-movement is modeled as short scrambling, with in-situ composition affecting the interpretation of everything between the landing site of scrambling and C. Here I will briefly sketch LFs based on these three approaches to movement, and spell out predictions for sentence processing, which will be the main focus of attention in this chapter.

The predicted LF under the traditional covert movement approach, where all *wh*-phrases must occur in the CP periphery by LF, is given in (265).[2]

(265) *A covert movement approach to question formation:* (= 237)

a. *Which* student read *which* book?

b. LF: [$_{\text{CP}}$ *Which* student *which* book C [$_{\text{TP}}$ ___ read ___]]

This view of covert movement predicts pervasive covert movement in multiple *wh*-questions, parallel to the overtly observed behavior of multiple fronting languages such as Bulgarian (Richards 1997:a.o.) (see section 7.2.1). Moreover, movement is always triggered for one and the same reason — the semantic needs of the *wh*-phrases themselves — and it always targets the same syntactic position at LF: C. This approach makes the prediction in (266).

(266) *A prediction of the 'traditional' covert movement approach:*
All *wh*-phrases in a question must (overtly or covertly) move to C for interpretation.

Against this backdrop stands the traditional in-situ approach to *wh*-in-situ, according to which the meaning of a question can be calculated through a mechanism that passes the meanings of *wh*-words up the structure until they reach C, where they can be interpreted (Hamblin 1973; Rooth 1985, 1992). No movement needs to be assumed for the interpretation of *wh*-phrases in the structure.

(267) *An in-situ approach to question formation:*
[$_{\text{CP}}$ *Which* student [C [$_{\text{TP}}$ ___ read *which* book]]]

This approach to questions makes no particular predictions about the position of *wh*-phrases at LF. Following standard assumptions in the theoretical literature that the simplest syntactic structure for a sentence is always preferred to a less simple one (cf. Chomsky 1995, 2000; Collins 2001; Epstein 1992; Fox 2000; Kitahara 1997; Reinhart 2006), it is predicted that *wh*-phrases occupy the position at which they were merged into the syntactic structure, unless another syntactic process triggers their movement.

(268) *A prediction of the in-situ approach:*
Wh-phrases can be interpreted in-situ and do not require any movement.

Evidence from previous chapters in this book have led me to argue for a combined account of *wh*-in-situ in English questions. We have seen evidence that *wh*-phrases covertly move in superiority-obeying questions, but that movement need not target interrogative C. I have argued that movement should be modeled as covert scrambling: *wh*-in-situ undergoes a short covert movement step when it is merged into the structure, and additional movement is motivated if it is necessary for the interpretation of the *wh*-phrase — for example, in order to avoid an intervention effect.

(269) *A covert scrambling approach to question formation:*
[$_{CP}$ *which* student [C [$_{TP}$ ... *which* book [$_{vP}$ ___ read ___]]]]

This is a logical consequence of a liberal interpretation of Hamblin (1973): *wh*-phrases may be able to undergo a *partial* movement step to positions other than interrogative C, and then be interpreted through the projection of Rooth-Hamblin alternatives between this landing site and C. Such movement will not be triggered for the reasons proposed in Karttunen (1977), as this would necessitate movement to C, but other factors may drive covert movement at LF.

(270) *A prediction of the scrambling approach:*
Wh-phrases may undergo a short movement step, and will only move further if there are external reasons requiring this movement.

8.2 Experimental background

This section provides an introduction to the experimental methodology that will be central to the arguments made in the remainder of this chapter. In section 8.3 I present a series of experiments supporting the view that English *wh*-in-situ undergo a short covert scrambling step and can be interpreted without any further movement between their landing site and C. Before turning to the evidence, I first present background on Antecedent Contained Deletion (ACD), which will be central to the experiments, the Hackl et al. (2012) paradigm, and a general discussion of how syntax and online processing interact.

8.2.1 Antecedent Contained Deletion

ACD is a phenomenon found in certain VP ellipsis contexts. For VP-ellipsis to be licensed, a pronounced antecedent VP must exist that is identical to the missing VP.[3] This is straightforward in examples like (271), where the only pronounced VP in the sentence—*read a book*—can serve as an antecedent for the missing VP. In ACD cases like (272), however, the missing VP appears to be properly contained inside the only possible antecedent VP in the sentence. Matching the missing VP with the antecedent VP should be impossible since the missing VP is properly contained inside its antecedent and so cannot be identical to it.

(271) *VP-ellipsis and its resolution:*

a. Sue read a book and Mary did ⟨missing VP⟩ , too.

b. Sue read a book and Mary did ~~⟨read a book⟩~~, too.

(272) *ACD and the containment problem:*

a. Sue read every book that Mary did ⟨missing VP⟩ .

b. Sue read every book that Mary did ⟨??⟩.

To solve the containment problem in (273a), the standard analysis of ACD assumes that the object is covertly moved out of TP_1 to a syntactic position in the higher TP_2, yielding the LF in (273b). The resulting VP, containing only the Verb and the trace of covert movement, can then be used as an antecedent for the missing VP (cf. Fox 2003; Larson and May 1990; May 1985; Sag 1976; Williams 1974, 1977).

(273) *Resolution of ACD using covert movement:*

a. [$_{\mathrm{TP1}}$ Sue read [every book that Mary did ⟨missing VP⟩].

b. [$_{\mathrm{TP2}}$ [every book that Mary did ~~read t~~] [$_{\mathrm{TP1}}$ Sue read t]].

For additional details on the interaction of ACD with multiple *wh*-questions, see the appendix to chapter 2.

8.2.2 Self-paced reading and the Hackl et al. (2012) paradigm

In this section I present the methodology that is used in the experiments in section 8.3. These experiments use the paradigm developed in Hackl et al. (2012), which studies the real-time processing of sentences where covert movement has been argued to occur, concentrating on the universal quantifier *every*.

The Hackl et al. (2012) paradigm takes advantage of the inherently linear organization of real-time sentence processing: the human parser integrates material that occurs earlier in a sentence before it encounters material occurring later. The paradigm relies on two underlying assumptions about the economy of structure building: (a) the linguistic parser always builds the simplest syntactic structure consistent with the linguistic input, and (b) structures without covert movement are simpler than structures with covert movement.

Given these assumptions, we expect the parser not to postulate covert movement in the parse of a sentence until the point at which the parser determines that it is necessary. From that point on, one might expect to find online consequences (detectable as a delay in Reading Times, RTs) of the *reanalysis* of the structure to a less preferred parse.

Hackl et al. (2012) compare the processing of sentences that contain non-quantificational objects (*the*, 274a-c) to sentences with quantificational objects (*every*, 274d-f). This factor is crossed with three different gap sizes inside an attached relative clause: (a) *no ellipsis* using a lexical <u>*verb*</u>, (b) *small ellipsis* marked by <u>*did*</u>, where the antecedent of the ACD is the embedded VP_2 headed by *treat*, and (c) *large ellipsis* marked by <u>*was*</u>, where the antecedent of the ACD is the matrix VP_1 headed by *reluctant*.

(274) *The Hackl et al. (2012) paradigm:*

The doctor was [$_{\text{VP1}}$ reluctant to [$_{\text{VP2}}$ treat ...

a. **the** patient that the recently hired nurse <u>*admitted*</u> *(no ellipsis)*
b. **the** patient that the recently hired nurse <u>*did*</u> *(small ellipsis)*
c. **the** patient that the recently hired nurse <u>*was*</u> *(large ellipsis)*
d. **every** patient that the recently hired nurse <u>*admitted*</u> *(no ellipsis)*
e. **every** patient that the recently hired nurse <u>*did*</u> *(small ellipsis)*
f. **every** patient that the recently hired nurse <u>*was*</u> *(large ellipsis)*

... after looking over the test results.

Under a standard analysis, ACD resolution involves at least three steps: (i) creating a structure in which antecedent containment is undone; (ii) identifying an appropriate antecedent for the ellipsis; and (iii) filling the antecedent into the gap and computing the resulting meaning. Steps (ii) and (iii) are required in all cases of VP ellipsis. Step (i)—reanalysis of the structure to undo antecedent containment—is only required in the case of ACD. Hackl et al. (2012) generate specific predictions for language processing in real time based on these properties and the assumptions that (a) step (i) of ACD resolution requires covert movement, and (b) quantificational objects (but not non-quantificational objects) require covert movement (in particular, Quantifier Raising) for their interpretation, and that this movement targets the lowest position in the structure where the object can be interpreted (cf. Fox 1995, 2000; Heim and Kratzer 1998; May 1985).

In the *definite* conditions in (274a–c), no covert movement is predicted to take place when the definite article is processed. The parser will only assume covert movement if and when it determines that the sentence contains an instance of ACD. This will happen after encountering the auxiliaries *did* and *was* in (274b–c), which, together with the immediately following word, signal the presence of an ACD site and thus trigger reanalysis in order to resolve ACD. This reanalysis should incur a processing cost detectable as a slowdown in RTs following the ellipsis site, compared to the baseline with *no ellipsis* (274a). Furthermore, the difference in the locality of covert movement and, concomitantly, the size of the antecedent VP should also be reflected in RTs: the covert movement in (274c) must target a non-local position, above the matrix VP_1 headed by *reluctant*, in order to make the matrix VP available for ACD resolution while the covert movement in (274b) targets a closer position above the embedded VP_2 headed by *treat*. On the assumption that non-local movement and the retrieval of a larger antecedent VP are more costly than local movement and the retrieval of a smaller antecedent VP, and that the non-local VP instantiates a more complex meaning than the local VP, (274c) is expected to produce longer RTs than (274b). These predictions are summarized in (275). Here and in (276) below, triggers for movement are boxed:

(275) *Predictions for the* definite *conditions:*

a. The Dr. was reluctant to treat **the** patient that the nurse *admitted.*
⇒ Lexical verb condition: No movement assumed at all.

b. The doctor was reluctant to treat **the** patient that the nurse $\boxed{\textit{did.}}$
⇒ Movement assumed after **did** is encountered. *late trigger*
Movement targets a position above *treat*. *short movement*

c. The doctor was reluctant to treat **the** patient that the nurse $\boxed{\textit{was.}}$
⇒ Movement assumed after **was** is encountered. *late trigger*
Movement targets a position above *reluctant*. *long movement*

For the *quantificational* conditions, different predictions are made. In all of (274d–f), the parser must assume covert movement as soon as it encounters the quantificational object headed by *every*. Furthermore, the movement is expected to be *local* and to target a position just above the embedded VP_2

headed by *treat*: this is the first position where quantifiers in object position can be integrated into the structure without causing a type-mismatch. Importantly, this position is high enough to preemptively undo antecedent containment in the case of *small ellipsis* (274e) but not in the case of the *large ellipsis* (274f): the movement triggered by *every* targets the same position that would be independently targeted for the resolution of the *small ellipsis*. Because this movement has already happened earlier in the parse, step (i) of ACD resolution can be avoided, and only steps (ii)–(iii) must apply when the ellipsis marker *did* is encountered. This means that at the point of identifying the ACD site, no reanalysis is predicted for this sentence. Hence, ACD resolution is expected to be easier compared to (274b) since part of the work necessary to resolve ACD has already happened prior to encountering the ACD marker *did* with *every*, but all three steps of ACD resolution must happen when *did* is encountered with *the*. In other words, we expect a facilitation effect of ACD resolution, detectable as relatively fast RTs, in the *every*-*did* condition.

In the case of the *large ellipsis* in (274f), by contrast, the covert movement step that is assumed following the processing of *every* is not sufficient to furnish a suitable antecedent VP: following this local movement, the missing VP is still contained inside its antecedent. Hence, when the auxiliary *was* is reached, the parser must again reanalyze the structure, covertly moving the object a second time, from its position above the embedded VP to a position higher than the matrix VP. This means that no facilitation of ACD resolution is expected in (274f) even though the host DP is quantificational in nature. These predictions are summarized in (276):

(276) *Predictions for the* every *conditions:*

a. The Dr was reluctant to treat **every** patient that the nurse *admitted.*
⇒ Movement assumed after **every** is encountered. *early trigger*
Movement targets a position above *treat*. *short movement*

b. The doctor was reluctant to treat **every** patient that the nurse *did.*
⇒ Movement assumed after **every** is encountered. *early trigger*
Movement targets a position above *treat*. *short movement*

c. The Dr. was reluctant to treat **every** patient that the nurse *was.*
⇒ Movement assumed after **every** is encountered. *early trigger*
Movement targets a position above *treat*. *short movement*
⇒ **Additional** movement assumed after **was** is encountered. *late trigger*
Movement targets a position above *reluctant*. *long movement*

Hackl et al. (2012) used sentences in a paradigm as in (274) in a self-paced reading study (Just, Carpenter, and Wooley 1982): participants read sentences that appeared on the screen one word at a time using a moving window display. Residual Reading Times (RRTs) were analyzed for each word in the sentence. Note that RRTs measure whether a given word in a given trial in the experiment was read at the average speed for all words of the same length throughout the experiment for each participant (represented by RRT=0 in the graph below), faster than expected (negative RRT values), or slower than expected (positive RRT values).[4]

Figure 8.1 shows RRTs two words after the ellipsis site. These results indicate that the predictions described above are borne out: When the object is *definite*, ACD resolution is associated with longer RRTs compared to the baseline condition, *no ellipsis* (*Verb*). The increase in RRTs is linear across the three gap size levels (*the*-*verb* vs. *the*-*did* vs. *the*-*was*). When the object is *quantificational*, however, an interaction pattern emerges. No increase is observed between the *every*-*verb* condition and the *every*-*did* condition (*small ellipsis*), while a large increase is observed between the *every*-*did* condition and the *every*-*was* condition (*large ellipsis*).

This is unexpected if the two factors (quantificational vs. non-quantificational object and ACD size) are not linked in some form, but it is the expected result under the assumptions that (a) covert movement is required to accommodate both a quantificational object and an ACD site, as the first step in resolving antecedent-contained ellipsis, and that (b) quantifiers undergo a short QR step sufficient to facilitate ACD resolution with *did*, but insufficient in the case of *was*.[5]

8.2.3 How syntax and online processing interact

At this point, let us step back and more carefully spell out the assumptions that self-paced reading paradigms rely on. This discussion will serve as the basis for the experiments presented in section 8.3 below.

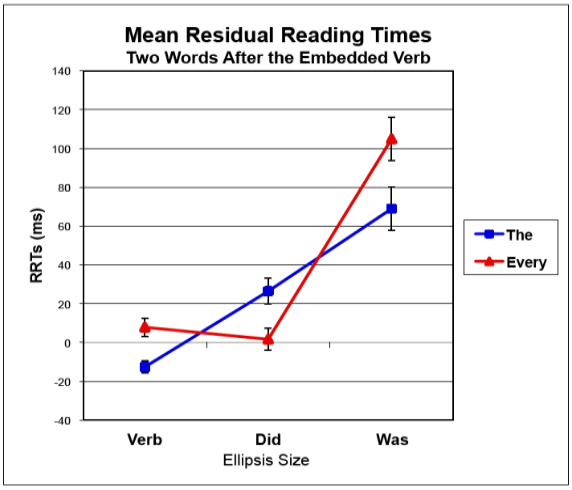


Figure 8.1
Results of Hackl et al. (2012)

Structure building and processing

First let us consider how online sentence processing corresponds to structure building in syntax. The predictions laid out above assume that linguistic structure can be built in real time from left to right. This is enforced rather straightforwardly by the Self-Paced Reading paradigm: participants encounter material on the left before they reach material on the right, and hence processing of linguistic material proceeds similarly from left to right (Phillips 1996).

Next, consider when the parser begins to build a parse for the material it encounters. A consensus in the literature is that the parser does not wait until a full sentence has been uttered in order to begin imposing a structure on it. Instead, structure is built 'online,' in tandem with reading or hearing the sentence. The literature offers two approaches to this process: *serial processing*, where the parser pursues just one parse of a sentence at a time (e.g. Ford, Bresnan, and Kaplan 1982; Frazier and Rayner 1982; Pritchett 1992) and *parallel processing*, where the parser pursues multiple analyses at once (e.g. MacDonald, Perlmutter, and Seidenberg 1994; Tanenhaus and Trueswell 1995). An intermediate third way of viewing processing is to assume a *serial* parser, in the sense that it assumes one particular parse of the sentence at a time, but at each step multiple potential parses are evaluated before that step is made.

Here I will be using rhetoric that assumes a serial parse, but the conclusions I draw are compatible with parallel models as well. The important assumption—a standard and well-motivated one—is that the parse is built simultaneously with reading or hearing the sentence, not after it has been completed.

Although processing happens from left-to-right and is very fast, it may not be instantaneous or strictly incremental in the sense that each new word is fully processed and integrated before the next word is processed. In fact, effects in the self-paced reading paradigm often show up not on the word expected to trigger an effect, but downstream from that word. The more costly the expected computation, the longer it takes for the effect to show up. For this reason, it is advisable to employ precaution measures that make it possible to isolate the hypothesized effects of specific triggers in a sentence. First, possible triggers for slowdown—for example, suspected triggers of covert movement—are separated from one another by additional lexical material. In the Hackl et al. (2012) paradigm in (274), two factors are varied: the quantifier heading the relative clause (*the* vs. *every*), and the ellipsis marker at the right edge of that clause (*verb* vs. *did* vs. *was*). These two factors are separated by 6 words, to ensure that effects from the earlier trigger do not interfere with the later trigger.[6] Second, a "continuation" is used following the second trigger, to allow for a spillover region where effects of this trigger might show up, if they are not instantaneous. These can sometimes make for a long and perhaps cumbersome sentence, but this complication is an inherent component of all SPR experiments, and experimenters work to ensure that while long, the sentences are as easy to process as possible.

Covert movement and Self-Paced Reading

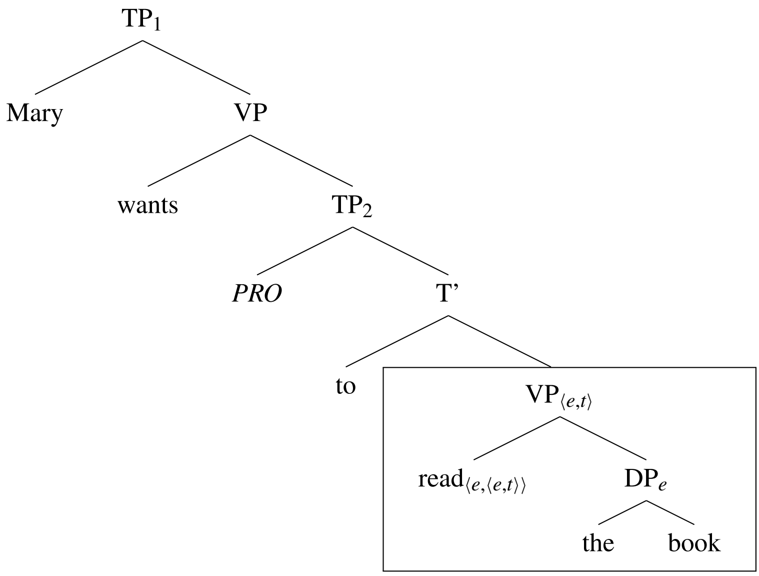
With this in mind, consider how movement is integrated into an online parse of a sentence. Overt movements (to the left) constitute *filler-gap dependencies*, which have a well-known processing signature: the parser encounters a linguistic element in a position other than its theta-position; this element must be stored in memory until its gap position is found, so that it can be integrated into the structure there. There is a rich literature surrounding the processes which are implicated in processing such filler-gap dependencies. For example, it is known that the difficulty of processing these dependencies increases with distance (Gibson 1998), interference effects (van Dyke and McElree 2006; Gordon, Hendrick, and Johnson 2004), as well as processing load imposed by other referential entities along the filler-gap path (Warren and Gibson 2002).

Much less is known about the effects of covert movement operations on real time sentence processing. To get off the ground, we need to make explicit two assumptions that will play a role in the reasoning below: First, I assume, following Frazier and Fodor (1978) and much subsequent work, that processing is not just incremental from left-to-right but also biased towards simplicity along the lines of the principle of Minimal Attachment, (277) (see also Bever 1970; Frazier and Rayner 1982; Phillips 2003:a.o.). Second, I assume that, all else being equal, a parse without covert movement is simpler than a parse with covert movement, and that a parse with local (short) covert movement is simpler than a parse with non-local (long) covert movement (cf. Anderson 2004; Fox 1995, 2000; Frazier 1999; Tunstall 1998).

(277) *The minimal attachment principle:*
Attach each new item into the current phrase marker postulating only as many syntactic phrase nodes as is required by the grammar.

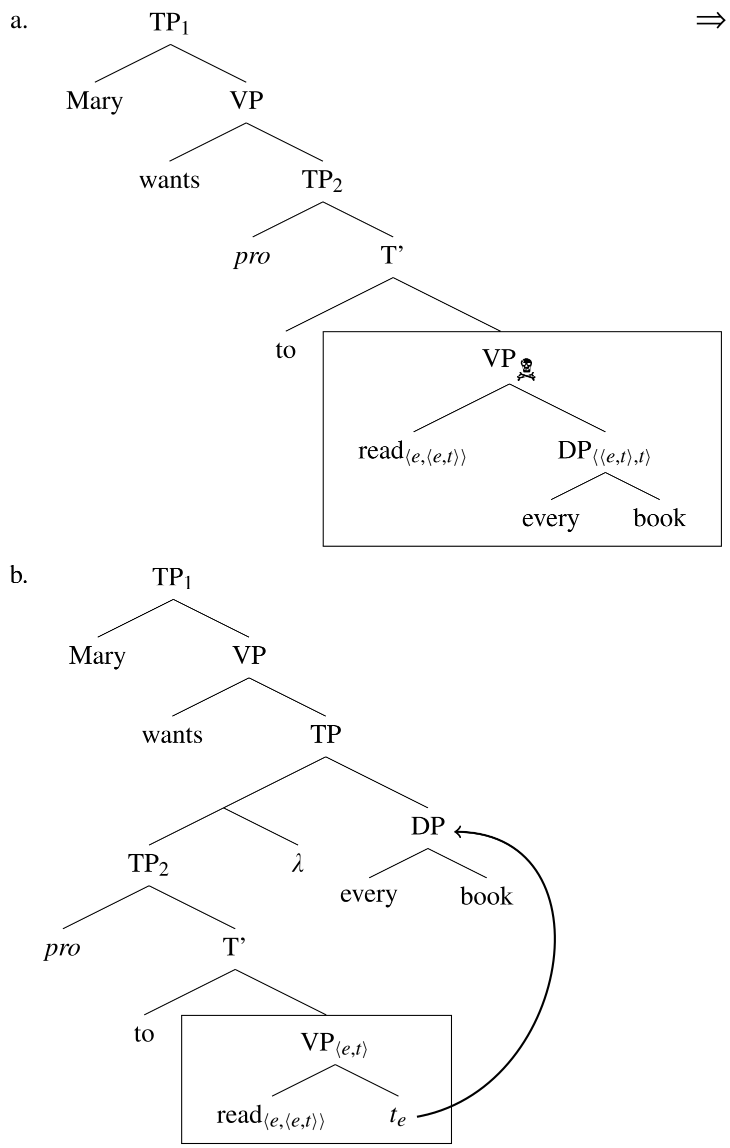
Given these assumptions, we expect the parser not to postulate covert movement in the parse of a sentence until the point at which the parser determines that it is necessary. Therefore, we expect the parser to assume no covert movement upon encountering a definite object DP, since definite descriptions are able to semantically merge with a verb in-situ. This step in the parse is illustrated (278).[7]

(278) *Integrating a definite object into the parse in online processing:*



Conversely, we expect movement to be assumed as soon as a quantificational object is encountered. Following the principles laid out above, we expect this movement to be as short as possible, to avoid assuming more structure than necessary. This step in the parse is illustrated in (279). The quantificational object cannot be integrated into the structure as the sister of the verb, because of a type mismatch. The object is instead integrated into the first position at which it is interpretable—here, immediately above TP_2.[8] As in chapter 7, I illustrate QR as movement to the right in these trees, although nothing hinges on this assumption.

(279) *Integrating a quantificational object into parse in online processing:*



Since going from the original parse in (279a) to the alternative one in (279b) amounts to a *reanalysis* triggered by the quantificational DP 'every book,' one might expect to find online consequences of this move to a less preferred parse.

Alternatively, a slowdown might be attributed simply to the presence of the quantifier in the structure: since quantifiers are less frequent in text and speech than definite articles, a slowdown in RTs might reflect low-level surprisal due to lexical frequency. Moreover, it is worth pointing out that observing no slowdown after a quantificational object does not mean that reanalysis as described in (279) did not take place. It could simply be that this sort of reanalysis is preformed by the parser with sufficient automaticity that SPR is not sensitive enough to detect it. As a result, an effect or the lack of an effect of *every* vs. *the* cannot be conclusively interpreted to be associated with a covert movement step in the derivation.

Finally, unlike in cases of overt movement, we do not expect the traditional costs associated with filler-gap dependencies: once the quantificational object is integrated into the structure, the parser already knows where the gap position of movement is, and therefore it need not hold onto the filler until a gap is located. That is, once the movement is performed, the meaning of the structure can be computed immediately and no further operation is needed.

Ellipsis resolution and Self-Paced Reading

The Hackl et al. (2012) paradigm crucially relies on a process of ellipsis resolution that is triggered once the auxiliaries *did* or *was* are encountered in the sentence. As discussed in the previous section, ellipsis resolution involves at least two steps: (i) identifying an appropriate antecedent for the ellipsis; and (ii) filling the antecedent into the gap and computing the resulting meaning of the sentence. In the case of Antecedent Contained Deletion, an additional step is necessary: an appropriate antecedent must first be constructed, before steps (i) and (ii) above can proceed. In the Hackl et al. (2012) paradigm, this involves covert movement of the object to a position above the verb heading the antecedent VP.

The movement step involved in this process may proceed in more than one way: the parser may always assume movement above the most local VP, as a first step. It then attempts to use this VP as an antecedent. If the resulting meaning is convergent, the parse may continue. Otherwise, the parser must assume additional movement, and will attempt to use the next higher VP as the antecedent for ellipsis. Thus, in the case of *small ellipsis*, the first assumption about movement is sufficient, but in the case of *large ellipsis*, a second movement step must be assumed in order to form an appropriate antecedent for the ellipsis. Alternatively, we may assume that the parser is able to immediately target the correct antecedent in both the *small* and *large* ellipsis conditions. In this case too, the large ellipsis has a richer content than the smaller ellipsis, and therefore integrating it into the structure and computing the resulting meaning of the sentence may be more costly.[9]

Under either formulation, we may expect to find slower RTs with larger ellipsis than with smaller ellipsis. As we will see below, such an effect is indeed present in all of the experiments presented here. Thus, a main effect of ellipsis size alone will not be taken as an indication of movement. Instead, the crucial question will be whether an earlier trigger for movement facilitates ACD resolution downstream—the RT-signature of facilitation being an interaction. This is expected under a movement analysis of ACD, but is unexpected if ACD resolution does not involve movement.

The timing of effects in Self-Paced Reading paradigms

Finally, we can now ask when we might expect to detect an effect of covert movement in a Self-Paced Reading paradigm. One obvious prediction is that we should not detect an effect *before* a trigger is encountered. However, it is common to find an effect not *on* the trigger word itself, but in the spillover region *following* that word (e.g. Clifton, Frazier, and Deevy 1999; Kazanina and Phillips 2000; Polinsky, Gallo, Graff, and Kravtchenko 2012; Polinsky and Potsdam 2014; Wagers and Phillips 2009; Xiang, Harizanov, Polinsky, and Kravtchenko 2011:among others). The more costly the computation associated with the trigger, the further downstream an effect may occur.

In the case of the Hackl et al. (2012) paradigm, the crucial region of interest starts with the later trigger for movement—the auxiliary verbs *did* and *was*, which signal the presence of Antecedent Containment in the structure. More accurately, the earliest point at which the parser has conclusive evidence for the presence of an ACD site is arguably the word following the auxiliary verb and not the auxiliary itself: at this point in the sentence, readers encounter a conjunction or disjunction that can only be attached at the clausal level, indicating that the previous clause contained ellipsis. For ease of exposition, however, I will assume in what follows that the trigger is the auxiliary verb itself and start counting words in the spillover region from there on.

The question, then, is when we should expect to detect effects of covert movement, as revealed by facilitation of processing an ACD site. As shown above, the Hackl et al. (2012) effects appear not on the auxiliary verb itself, but two words after the verb/AUX. Below we will see that all five experiments presented in this chapter also exhibit an effect in this same region: effects consistently show up two and three words after the auxiliary verb.

In this context, it is again important to recall this inherent limitation of the SPR paradigm: the SPR paradigm does not allow for an a-priori prediction as to the precise location of an effect, as long as the effect happens inside the spillover region. To interpret the effects of experiments based on the Hackl et al. (2012) paradigm as stemming from covert movement triggered by the need to resolve ACD, the effect must happen only after the parser encounters the triggers *did*/*was*, and they must happen within a region that is uniform across all conditions so that nothing else in that region could have generated the effects. The materials below were constructed to ensure this. Moreover, an important strategy for confirming the validity of results of SPR experiments is replication and consistency: crucial effects reported in the experiments below survive replication and show up in the same region across all experiments.

8.3 Experimental evidence for covert scrambling

With this background established, in this section I present experimental evidence in support of the covert scrambling view of covert *wh*-movement. All of the experiments rely on the Hackl et al. (2012) paradigm, building on multiple questions that host an ACD site, such as in (280):

(280) *ACD hosted by in-situ* wh-*phrase:*
Mary asked *which soloist* [$_{\text{VP1}}$ wanted to [$_{\text{VP2}}$ perform *which concerto* that the brilliant protégé did Δ]]. (ambiguous)
Δ = VP_2 "perform *t*" *(small ellipsis)*
Δ = VP_1 "wanted to perform *t*" *(large ellipsis)*

Recall our predictions from section 8.1, summarized here for convenience:

(281) *Predictions for online sentence processing:*

a. If covert movement targets interrogative C, then all *wh*-phrases move *non-locally* to C. As a result, we predict that both *small ellipsis* (targeting VP_2) and *large ellipsis* (targeting VP_1) should be relatively easy to process, because antecedent containment will have been undone by covert *wh*-movement before the parser reaches the gap site. When the parser encounters the gap site, all it needs to do is find an antecedent for the missing VP and nothing more.

b. If *wh*-phrases are interpreted in-situ, the parser will only reanalyze the structure and covertly move the *wh*-object to undo antecedent containment once an ellipsis site is reached, in order to construct an appropriate antecedent for the ellipsis. Hence, a high processing cost should be associated with the resolution of the ACD site in both the *small* and *large* ellipsis cases (with a potentially larger cost for the large ellipsis).

c. If *wh*-phrases covertly scramble a short distance, using a movement step akin to QR, we expect the movement to target a position above VP_2 but below VP_1. As a result, we predict that *small ellipsis* (targeting VP_2) should be relatively easy to process, but *large ellipsis* (targeting VP_1) should incur a high processing cost.

The three experiments in this section are designed to test these predictions.

8.3.1 Experiment 1: *every* vs. *which*

Experiment 1 compares ACD facilitation effects in multiple *wh*-questions to ACD facilitation with quantifiers. Since Hackl et al. (2012) looks at QR with *every*, this experiment will provide a baseline: we compare the behavior of *every* to that of *wh*-in-situ and ask whether the effects are stronger (favoring the 'traditional' covert movement view), weaker (favoring an in-situ view), or parallel (favoring a covert scrambling view). I elaborate on these predictions below.

Design and materials

Experiment 1 presented participants with (embedded) *wh*-questions headed by a subject *wh*-phrase. Two factors were crossed: (a) *determiner*: whether the embedded question contained the quantificational determiner *every*, yielding a simplex *wh*-question, or a second *wh*-phrase, yielding a multiple *wh*-question; and (b) *ellipsis size*: whether the sentence contained a *small ellipsis* marked by *did*, where the antecedent of the ACD site is the embedded VP_2, or *large ellipsis* marked by *was*, where the antecedent of the ACD site is the matrix VP_1. A sample item is given in (282) below:[10,11]

(282) *Sample target item in Experiment 1:*

The conductor asked *which soloist* was [$_{\text{VP1}}$ willing to [$_{\text{VP2}}$ perform...

a. **every** concerto that the brilliant protégé *did* *(small ellipsis)*

b. **which** concerto that the brilliant protégé *did* *(small ellipsis)*

c. **every** concerto that the brilliant protégé *was* *(large ellipsis)*

d. **which** concerto that the brilliant protégé *was* *(large ellipsis)*

...and restructured the rehearsal accordingly.

There were 28 sentence templates modeled similarly to the sample paradigm in (282). Each sentence in a template employed either *every* or *which* as the determiner of the object DP. This DP hosted a relative clause with an ACD site marked with an auxiliary verb. The auxiliary *did* marked a *small ellipsis* corresponding to the embedded predicate, VP_2, and the auxiliary *was* marked a *large ellipsis* corresponding to the matrix predicate, VP_1. After the ellipsis site, the sentences had continuations beginning with a clausal conjunction or disjunction, which varied in length but were at least 5 words long, providing a spillover region for detecting possible processing difficulties associated with ACD resolution. See the appendix for a full list of the materials for all the experiments presented in this chapter.

Because the experiments were conducted online, it was not possible to control the participants' screen size. Consequently, in order to ensure that the region of interest was read without interruptions that may artificially affect the data, all the sentences were presented on two lines, with the line break in target sentences always placed immediately following the verbal complex (that is, the first line of the sentence was the first line in (282), and the second line contained the text in lines a–d and the continuation following these lines).[12]

Target items were counterbalanced across four lists using a Latin Square design and combined with 48 filler sentences of various types resulting in a total of 76 sentences. Non-target items included sentences that were similar to the target items in structure, in length and in containing quantifiers. 18 filler sentences resembled the target sentences in all aspects but contained a lexical verb instead of ellipsis (*did*, *was*). These filler sentences additionally contained line breaks in different positions, making it impossible for participants to anticipate where a line break might occur, or whether or not there will be ellipsis in the sentence.[13] The remaining filler sentences were taken from an unrelated study.

Each experimental item was followed by a yes/no comprehension question. The questions asked about different aspects of the sentences, including about material inside the relative clause and about the predicates used in the sentences, to ensure that participants were processing all parts of the sentence at a deep level. The correct answers to half of the questions was *yes* and to the other half *no*.

Methods

Experiment 1 used the moving window self-paced reading methodology and was hosted on Ibex Farm.[14] Participants were presented with sentences that appeared on the screen one word at a time in a moving window display. After the final word of each sentence, the comprehension question appeared. Participants responded by pressing "1" for "Yes" or "2" for "No." No feedback was given about whether the answer to the question was correct or incorrect.

Before beginning the experiment, participants were given detailed instructions about the experiment and then read and accepted a consent statement. Participants were instructed to read the sentences at a natural rate to ensure understanding. They were also instructed to answer the comprehension questions as accurately as possible. There were three practice items before the experiment began. The experiment took approximately 30 minutes to complete.

Participants were recruited through Amazon Mechanical Turk and were paid $1.5 for their participation. Participants were asked about their native language but were told that payment was not contingent on their response. To further ensure that only native speakers of English participated in the experiments, IP addresses of participants were restricted to the US using Amazon Mechanical Turk's user interface. Only Turk Workers with an overall approval rate of over 95% of all their submissions were allowed to participate.

Predictions

The *every* conditions are expected to replicate the results of Hackl et al. (2012). That is, we expect to find a main effect of *ellipsis size*, such that *small ellipsis* is easier to process than *large ellipsis*: since *every* triggers covert movement to a position above the embedded VP_2 (headed by *perform* in (282)) as soon as the quantifier is encountered, we expect antecedent containment to be preemptively undone in the case of the *small ellipsis* (282a), leading to facilitation of ACD resolution. However, since this movement does not target a position high enough to undo antecedent containment in the case of the *large ellipsis* (282c), we expect ACD resolution to be relatively more difficult in this case: once the auxiliary *was* is reached, the parser must perform a second reanalysis, covertly moving *every* from its QRed position above the embedded VP_2 to a position above the higher VP_1 (headed by *willing* in (282)), in order to allow for ACD resolution. These two conditions thus provide us with a baseline contrast against which to compare the *which* conditions in (282b,d).

For the *which* conditions, the three approaches to in-situ *wh*-phrases make different predictions. Under the 'traditional' covert movement approach, all *wh*-phrases must move *non-locally* to C for interpretation. Both *small* and *large* ellipsis are predicted to be relatively easy to process because antecedent containment is undone by covert *wh*-movement before the parser reaches the ellipsis site. We thus expect an interaction, such that the *which-did* and *which-was* conditions pattern with *every-did* and exhibit facilitation effects, whereas we predict participants to exhibit increased difficulty with ACD resolution in the remaining *every-was* condition.

Under the in-situ approach, the in-situ *wh*-phrase can be interpreted without any movement. Hence, only upon reaching the ellipsis site will the need for reanalysis be apparent. Thus processing costs for ACD resolution are predicted to reflect both covert movement of the *wh*-object to a position above the missing VP as well as the retrieval of the appropriate antecedent for the elided VP and so should be relatively higher for both the *small* and *large* ellipsis conditions. We thus expect to find a main effect of the object type, such that the *which* conditions are more difficult to process than the *every* conditions. This main effect may be accompanied by an interaction, such that the two *which* conditions pattern with *every-was* and are more difficult to process than *every-did*, or they may be even more difficult than the *every-was* condition.

Finally, under the covert scrambling approach, *wh*-in-situ undergoes a short scrambling step akin to QR. Hence, we predict that the *which* conditions in the experiments should show similar results to the *every* conditions. That is, we expect to find a main effect of *ellipsis size* such that the *did* conditions are easier to process that the *was* conditions, but no effect of determiner.

Results

61 native speakers of English participated in this study. The following exclusion criteria were used to filter the results of this experiment and all subsequent

ones: participants who held the spacebar continuously pressed instead of reading the sentences one word at a time as instructed, participants who participated in the study more than once, participants who submitted the entire survey in less than 10 minutes, participants with an average reaction time of over 700ms,[15] and participants with low accuracy rates in response to comprehension questions (<75% on filler trials and <75% on target trials) were excluded from the study. Twenty participants in Experiment 1 were excluded from the analysis for these reasons.[16] In addition, two target sentences were excluded from the analysis because of low accuracy (<60% across participants).

Questions across the full experiment (targets and fillers) were answered correctly 87.5% of the time across participants; questions for target items were answered correctly on 83.3% of trials. The data was trimmed as follows: RTs from the first and last words of all items, RTs faster than 90ms or slower than 2000ms, and any RTs that were more than 2 standard deviations faster or slower than the average RTs for each subject (calculated per condition) were excluded from the analysis. Overall, less than 1% of the data were lost due to these criteria.[17] Figure 8.2 shows the mean residual reading times (RRTs)[18] for the two regions of interest for the four target conditions.

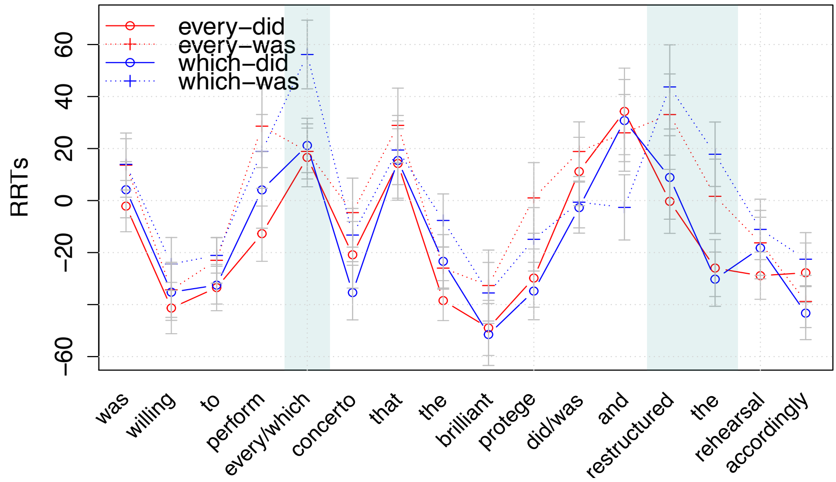


Figure 8.2
Residual reading times for target items in Experiment 1

A linear mixed effects model was fit to the data using *R* and the *R* package *lme*4 (Bates and Sarkar 2007). The model predicted RRTs from the two factors

Predictor	Coefficient	Standard Error	t value
Intercept	-31.591	18.851	-1.676
Determiner	3.586	16.018	0.224
Ellipsis size	39.8416	15.651	2.546
Determiner×*Ellipsis size*	2.665	20.628	0.129

Table 8.1
Results of Experiment 1

of interest: *determiner* (*every* vs. *which*) and *ellipsis size* (*small ellipsis* marked by *did*, vs. *large ellipsis* marked by *was*).[19] The model contained random intercepts and slopes for both predictors for subjects and items (Baayen 2004; Barr, Levy, Scheepers, and Tily 2013).

The results show a main effect of *determiner* at the slot at which the determiner appeared in the sentence (log likelihood tests comparing a model with and without the effect of *determiner*, p<0.05). This result is driven by the fact that RTs in the *which* condition were slower than RTs in the *every* condition, across both ellipsis conditions.[20]

The results additionally show a main effect of *ellipsis size* two words and three words after the auxiliary verb (log likelihood tests, p's<0.05). This result is driven by the fact that the resolution of *small ellipsis* is faster than the resolution of *large ellipsis* for both *every* and *which*. The results of the model for the third word after the auxiliary verb are summarized in Table 8.1.[21]

Discussion

There are two effects of interest in Experiment 1. First, the main effect of *determiner* in the first region of interest may be attributed to the relatively higher complexity of a multiple question compared to that of a simplex question, or to the relatively higher frequency of simplex as opposed to multiple questions. Regardless of the cause of this effect, it shows that the participants were processing the sentences at least at a depth sufficient for detecting the difference in determiner.

Second, we observe a main effect of *ellipsis size* in the second region of interest, following the auxiliary verb, such that sentences with a *small ellipsis* are read faster than sentences with a *large ellipsis*. This is the case for both *every* and *which*. To see this more clearly, observe Figure 8.3, which compares

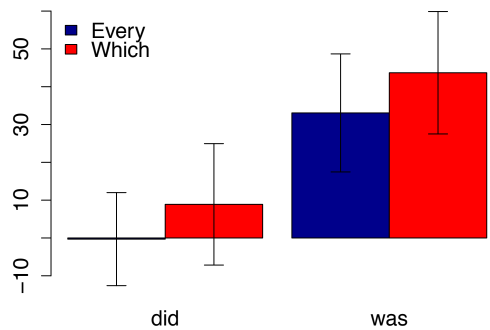


Figure 8.3
Residual reading times two words after AUX in Experiment 1

reading times for *every* and *which* two words after the auxiliary verb. As we can see, the *ellipsis size* manipulation affects the two determiners equally.

This is the result predicted by the covert scrambling view of covert movement — covert *wh*-movement is a short movement that facilitates *small* ellipsis, but is insufficient to preemptively undo antecedent containment in the case of *large* ellipsis.

8.3.2 Experiment 2: *the* vs. *every* and *the* vs. *which*

Experiments 2a–b provide an important baseline to strengthen the results of Experiment 1. In particular, this experiment compares the behavior of the non-quantificational determiner *the* with that of *every* and of *which*, to eliminate two possible concerns regarding Experiment 1: (a) that the paradigm it uses is not sensitive enough to detect ACD resolution facilitation effects, because of the added complexity of the embedded questions in the paradigm as compared to the original Hackl et al. (2012) paradigm; and (b) that participants are not processing the sentences at a deep level.

Design

To allow for a direct comparison with Experiment 1, Experiment 2 uses the same materials and methods as in Experiment 1, with minor changes to accommodate the experimental manipulation of this experiment. *Determiner* was treated as a between-subject factor: Experiment 2a compares *every* and *the*, and Experiment 2b compares *which* and *the*. The missing comparison, *which* and *every*, was the focus of Experiment 1.

(283) *Sample target item in Experiment 2a:*

The conductor asked *which soloist* was [$_{VP1}$ willing to [$_{VP2}$ perform...

a. **every/the** concerto that the brilliant protégé <u>*did*</u> *(small ellipsis)*

b. **every/the** concerto that the brilliant protégé <u>*was*</u> *(large ellipsis)*

and restructured the rehearsal accordingly.

(284) *Sample target item in Experiment 2b:*

The conductor asked *which soloist* was [$_{VP1}$ willing to [$_{VP2}$ perform...

a. **which/the** concerto that the brilliant protégé <u>*did*</u> *(small ellipsis)*

b. **which/the** concerto that the brilliant protégé <u>*was*</u> *(large ellipsis)*

...and restructured the rehearsal accordingly.

The same 48 filler items from Experiment 1 were used. The comprehension questions to some items were minimally changed to accommodate the change of determiners in the sentences.

Predictions

We expect Experiment 2a to replicate the results of Hackl et al. (2012). In particular, we expect to see an effect of *ellipsis size*, such that *small* ellipsis is easier to process than *large* ellipsis for both *the* and *every*. In addition, we expect to find a difference in the processing of ACD in sentences with a relative clause headed by *every* and sentences with a relative clause headed by *the*, such that the processing of ACD in sentences with *every* is facilitated compared to sentences with *the*. This is because we assume QR of *every* but no movement of *the* in our sentences.

We furthermore expect to find the same behavior pattern with *which* in Experiment 2b: if *which*, like *every*, undergoes a short movement step targeting a position above the lower VP_2, we expect to find that *which*, like *every*, facilitates the resolution of *small* ellipsis with <u>*did*</u>, but not of *large* ellipsis with <u>*was*</u>, for the same reasons as described above for *every*. We expect *the* not to facilitate ACD resolution of any size, and hence we predict that sentences with *the* to be relatively more difficult to process than sentences with *which*.

Results

165 native speakers of English participated in this study: 84 subjects participated in Experiment 2a and 81 participated in Experiment 2b. 21 subjects were excluded from the analysis of Experiment 2a and 24 subjects were excluded from the analysis of Experiment 2b using the same exclusion criteria specified in Experiment 1. Three target sentences and one filler sentences were excluded from the analysis of Experiment 2a and three target sentences and two filler sentence were excluded from the analysis of Experiment 2b because of low accuracy (<60% across all participants).

Questions across the full experiment (targets and fillers) were answered correctly 86.7% of the time in Experiment 2a and 85.8% of the time in Experiment 2b; questions for target items were answered correctly on 84.7% of trials in Experiment 2a and 83.8% of the time in Experiment 2b. The data was trimmed using the same criteria described for Experiment 1. Overall, less than 1% of the data was excluded from the analysis. Figure 8.4 shows the mean RRTs for the region of interest for the four target conditions in Experiment 2a, comparing the processing of sentences with *the* and *every*.

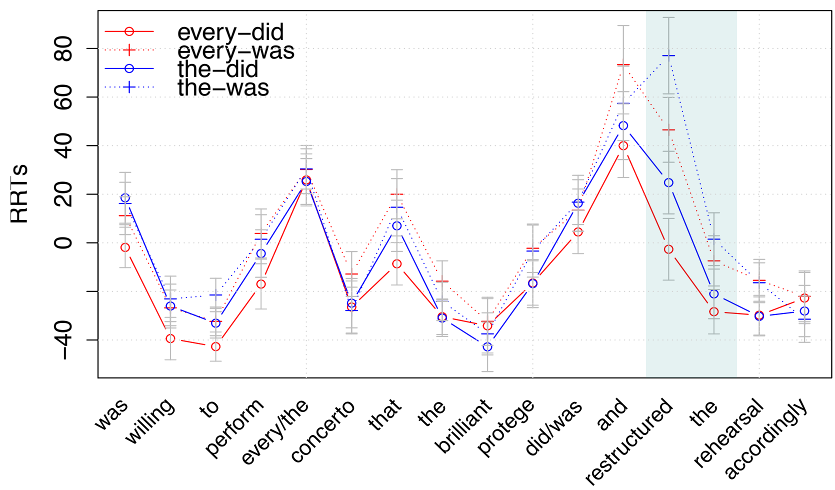


Figure 8.4

Residual reading times for target items in Experiment 2a

A linear mixed effects model with random intercepts and slopes for *ellipsis size* for subjects and items was fit to the data.[22] The model predicted RRTs from the two factors of interest: *determiner* (*every* vs. *the*) and *ellipsis size*

Predictor	Coefficient	Standard Error	t value
Intercept	3.304	20.959	0.158
Determiner	21.780	16.517	1.319
Ellipsis size	39.008	19.510	1.999
Determiner × *Ellipsis size*	12.207	23.441	0.521

Table 8.2
Results of Experiment 2a

(*small ellipsis* marked by *did*, vs. *large ellipsis* marked by *was*). The results show a main effect of *ellipsis size* two and three words after the auxiliary verb and a main effect of *determiner* two words after the auxiliary verb (log likelihood tests, p's < 0.05). These results are driven by the fact that the resolution of *small ellipsis* is faster than the resolution of *large ellipsis* for both determiners, and furthermore that the resolution of ACD in the *every* conditions is faster than in the *the* conditions. The results of the model for the second word after the auxiliary verb are summarized in Table 8.2.

Next, we examine the results of Experiment 2b, comparing the determiners *the* and *which*. Figure 8.5 shows the mean RRTs for the regions of interest for the four target conditions.

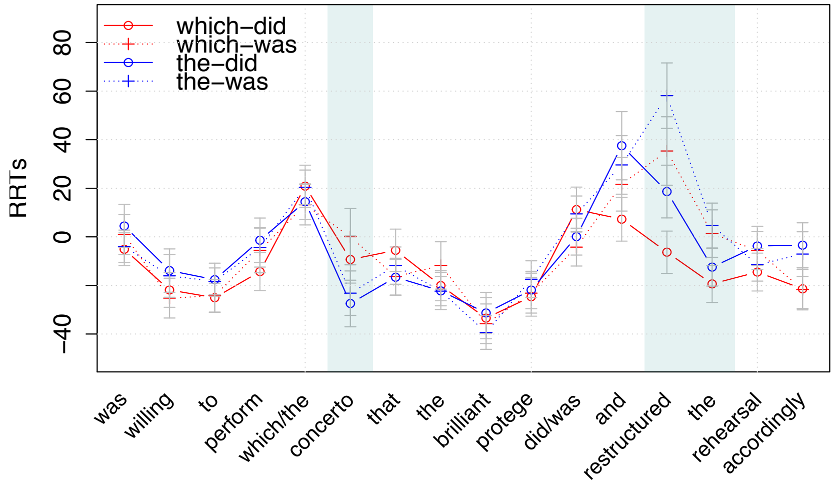


Figure 8.5
Residual reading times for target items in Experiment 2b

Predictor	Coefficient	Standard Error	*t* value
Intercept	16.572	15.716	1.054
Determiner	22.748	14.213	1.600
Ellipsis size	42.714	15.949	2.678
Determiner × *Ellipsis size*	5.117	20.483	0.250

Table 8.3
Results of Experiment 2b

A linear mixed effects model with random intercepts and slopes for *ellipsis size* for subjects and items was fit to the data. The model predicted RRTs from the two factors of interest: *determiner* (*which* vs. *the*) and *ellipsis size* (*small ellipsis*, vs. *large ellipsis*). The results show a main effect of *ellipsis size* two and three words after the auxiliary verb and a main effect of *determiner* two words after the auxiliary verb (log likelihood tests, p's <0.05). These results are driven by the fact that the resolution of *small ellipsis* is faster than the resolution of *large ellipsis* for both determiners, and furthermore that the resolution of ACD in the *which* conditions is faster than in the *the* conditions. The results of the model for the second word after the auxiliary verb are summarized in Table 8.3.

Discussion

The results of Experiment 2a–b confirm that the Hackl et al. (2012) paradigm extends to the context of an embedded question. In particular, we find—in addition to the main effect of *ellipsis size*—a main effect of *determiner*, such that sentences with *every* are processed faster than sentences with *the*, and sentences with *which* are similarly processed faster than sentences with *the*. That is, *every* and *which* pattern together and facilitate ACD resolution more than parallel sentences with *the*. To see this more clearly, consider Figure 8.6, which compares the reading times of *the*, *every* and *which* two words after the auxiliary verb site.

For the results of Experiment 2a, I adopt Hackl et al.'s (2012) explanation and assume that these results are the way they are because the parser must assume covert movement with *every* as soon as the quantifier is encountered, but no such covert movement is assumed with *the*.[23] As a result, covert movement is always assumed to take place at the point of ACD resolution in the

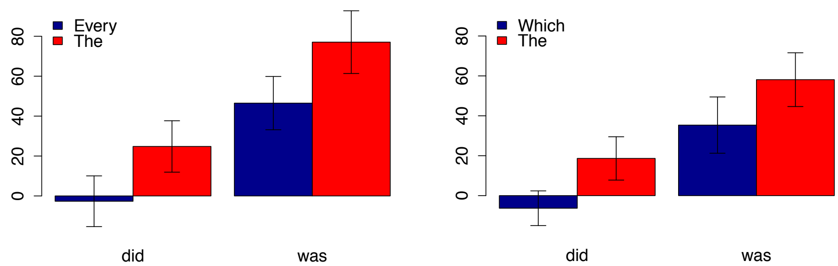


Figure 8.6
Residual reading times two words after AUX in Experiments 2a (left) and 2b (right)

case of *the* but not in the case of *every*: when a *small ellipsis* is encountered, the parser need not assume any additional covert movement, since sufficient movement has taken place earlier in the parse. To explain the fact that non-local ACD resolution is less difficult with *every* compared to *the*, I hypothesize that the second QR step to a position above the matrix VP_1 is easier in the case of *every*, perhaps because the object must move a shorter distance than in the case of *the*.

The results of Experiment 2b again show that *which* patterns with *every*. These results can be understood if we assume that the in-situ *wh*-phrases in the target items of Experiment 2b undergo a covert movement step similar in extent to that of *every*—that is, movement appears to target the lower VP_2, and not the higher VP_1 or interrogative C.

8.3.3 Experiment 3: *every* vs. *which* with interveners

This final experiment tests a further prediction of the covert *wh*-scrambling proposal. To get us started, recall that we predicted a short covert scrambling step of a (phonologically) in-situ *wh*-phrase similar to the overt scrambling step observed in German, as in the example below, repeated from (262a–b).

(285) *Obligatory overt short* wh-*scrambling in German (Hallman 1997):*

a. *Wer* hat denn (das Buch) gestern (das Buch) gelesen?
who has *denn* (the book) yesterday (the book) read
'Who read the book yesterday?'

b. *Wer* hat denn (*was*) gestern (**was*) gelesen?
who has *denn* (what) yesterday (what) read

'Who read what yesterday?'

We additionally predict that the presence of an intervener in the question can drive additional movement beyond this initial target position of movement. Again, following the overt behavior exhibited by German questions as in e.g. (259), repeated below for convenience, we expect covert scrambling to target the first position above the intervener where the *wh* can be integrated into the structure. I assume here that this is the first node with (extensional) propositional type in the structure.

(286) *German: intervention above* wh-*in-situ, avoided by scrambling*

a. * *Wer* hat **niemanden** $\boxed{\textit{wo}}$ angetroffen?
who has nobody where met

b. ✓ *Wer* hat $\boxed{\textit{wo}}$ **niemanden** ___ angetroffen?
who has where no one met
'Who didn't meet anybody where'? (Beck 1996a)

This prediction will be the focus of Experiments 3a–b, and we will see that it is indeed borne out.

Design

Experiment 3 is an extension of Experiment 1, adding interveners to the questions. Specifically, I use the focus-sensitive adverb *also* as an intervener.[24,25] Three factors are crossed: (a) *determiner* (*every* vs. *which* as the head of the object relative clause hosting the ACD site); (b) *ellipsis size* (*small ellipsis* marked by <u>*did*</u> vs. *large ellipsis* marked by <u>*was*</u>); and (c) *position of also* (*low also* occurring above the embedded VP_2 vs. *high also* occurring above the matrix VP_1). To simplify the design and the analysis of the results, Experiment 3 is divided into two sub-experiments, each using just one determiner, *every* or *which*. Sample items for Experiments 3a-b are given in (287)–(288).

(287) *Sample target item in Experiment 3a (*every*)*[26]

The conductor asked *which soloist* was ...

a. [$_{\text{VP1}}$ **also** willing to [$_{\text{VP2}}$ perform every concerto that the brilliant protégé <u>*did*</u> *(small ellipsis)*

b. [$_{\text{VP1}}$ willing to [$_{\text{VP2}}$ **also** perform every concerto that the brilliant protégé <u>*did*</u> *(small ellipsis)*

c. [$_{\text{VP1}}$ **also** willing to [$_{\text{VP2}}$ perform every concerto that the brilliant protégé <u>*was*</u> *(large ellipsis)*

d. [$_{\text{VP1}}$ willing to [$_{\text{VP2}}$ **also** perform every concerto that the brilliant protégé <u>*was*</u> *(large ellipsis)*

and restructured the rehearsal accordingly.

(288) *Sample target item in Experiment 3b (*which*)*

The conductor asked *which soloist* was ...

a. [$_{\text{VP1}}$ **also** willing to [$_{\text{VP2}}$ perform which concerto that the brilliant protégé <u>*did*</u> *(small ellipsis)*

b. [$_{\text{VP1}}$ willing to [$_{\text{VP2}}$ **also** perform which concerto that the brilliant protégé <u>*did*</u> *(small ellipsis)*

c. [$_{\text{VP1}}$ **also** willing to [$_{\text{VP2}}$ perform which concerto that the brilliant protégé <u>*was*</u> *(large ellipsis)*

d. [$_{\text{VP1}}$ willing to [$_{\text{VP2}}$ **also** perform which concerto that the brilliant protégé <u>*was*</u> *(large ellipsis)*

...and restructured the rehearsal accordingly.

Experiment 3 contained the same 48 filler items from Experiment 1. The target items from Experiment 1 were minimally changed to add the intervener *also* in the appropriate places. No changes were made to the comprehension questions.

Predictions

We expect interveners to interact with *wh*-phrases and force covert scrambling of the in-situ *wh*-phrase to a position above the intervener. Following our assumptions about the parser, we expect it to posit the shortest movement step possible. Hence, in the *low also* condition, we expect *wh*-scrambling to target a position above *also* but below VP_1, while in the *high also* condition we expect long-distance movement above *also* and thus above VP_1.

Crucially for this experiment, we do not expect interveners to interact with *every*, since the literature provides no evidence that conventional quantifiers interact with focus interveners. Hence, we expect *every* to undergo a small QR step and not move any further unless and until the need for long-distance movement becomes apparent, once the parser reaches an ACD site marked by <u>*was*</u>.

As a consequence, we expect Experiment 3a to replicate the results of Experiment 1, since interveners should not affect the parsing of sentences with *every*. That is, we expect to find a main effect of *ellipsis size* and, crucially, no effect of *also*. Different predictions are made in the case of Experiment 3b. Here, we expect the position of *also* to affect the extent of covert *wh*-movement in the question, irrespective of the size of ACD.

In the *low also* conditions (288b,d), we expect the results to resemble those of Experiment 1 and of Experiment 3a. The position of *also* above the embedded VP_2 will force scrambling to a position above *also* in the embedded VP_2 but below the matrix VP_1. This position is high enough to preemptively undo antecedent containment in the case of *small* ellipsis (288b) but not in the case of the *large* ellipsis (288d). Consequently, we expect ACD resolution to be facilitated in the case of *small ellipsis* but not in the case of the *large ellipsis*.

In the *high also* conditions (288a,c), the position of *also* is above the matrix VP_1. As always, the parser must assume covert movement as soon as it encounters *which*. Unlike in the case of *low also*, here the movement must be non-local and target a position above the matrix VP_1. This position is high enough to preemptively undo antecedent containment both in the case of *small* ellipsis (288a) and *large* ellipsis (288c). As a result, when the parser reaches the gap site, no reanalysis is necessary in order to construct an appropriate antecedent for the ellipsis. We thus expect ACD resolution to be facilitated in both the *small* and *large* ellipsis conditions.

Results

243 native speakers of English participated in this study: 123 subjects participated in Experiment 3a and 120 subjects participated in Experiment 3b. These details represent the aggregate results from two separate replications of each experiment, performed to ensure that the results reported here are robust. 27 subjects were excluded from the analysis of Experiment 3a and 25 subjects were excluded from the analysis of Experiment 3b using the same exclusion criteria specified for Experiment 1. Two target sentences and three filler sentences were excluded from the analysis of both experiments because of low accuracy ($<$60% across all participants).

Questions across the full experiment (targets and fillers) were answered correctly 84.9% of the time across participants in Experiment 3a and 86% of the time across participants in Experiment 3b; questions for target items were answered correctly on 82.42% of trials in Experiment 3a and on 82.47% of

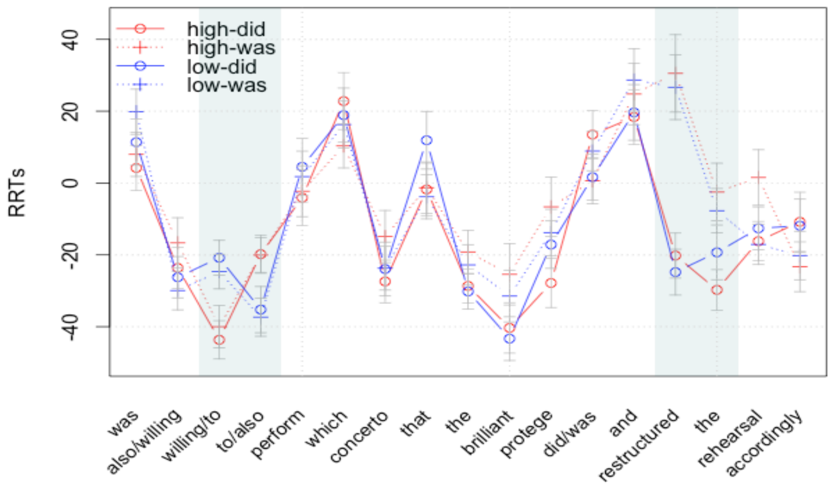


Figure 8.7

Residual reading times for target items in Experiment 3a

the trials in Experiment 3b. The data was trimmed as described for Experiment 1 above. Overall, less than 1% of the data was excluded. Below I present first the results of Experiment 3a and then those of Experiment 3b. Figure 8.7 shows the mean RRTs for the region of interest for the four target conditions in Experiment 3a.

A linear mixed effects model with random intercepts and slopes for the effect of *ellipsis size* and the effect of *also* (for subjects only) was fit to the data[27]. The model predicted RRTs from the two factors of interest: *position of also* (*low*, above the embedded verb vs. *high*, above the matrix verb) and *ellipsis size* (*small* vs. *large*). The results show a main effect of *also* at the third and fourth words in the region of interest (where *also* occurs in the sentence). The results additionally show a main effect of *ellipsis size* two and three words after the auxiliary verb (log likelihood tests, p's<0.05). These results are driven by the fact that the resolution of *small ellipsis* is faster than that of *large ellipsis* for both *also* conditions. The results of the model for the third word after the auxiliary verb are summarized in Table 8.4.

Predictor	Coefficient	Standard Error	*t* value
Intercept	-28.342	11.909	-2.380
Position of also	7.449	8.282	0.900
Ellipsis size	25.073	10.653	2.354
Position of also×*Ellipsis size*	-14.070	13.034	-1.079

Table 8.4
Results of Experiment 3a

Next, we examine the results of Experiment 3b. Recall that in this experiment, the determiner heading the object relative clause in all four target conditions was *which*. Figure 8.8 shows the mean RRTs for the region of interest for the four target conditions.

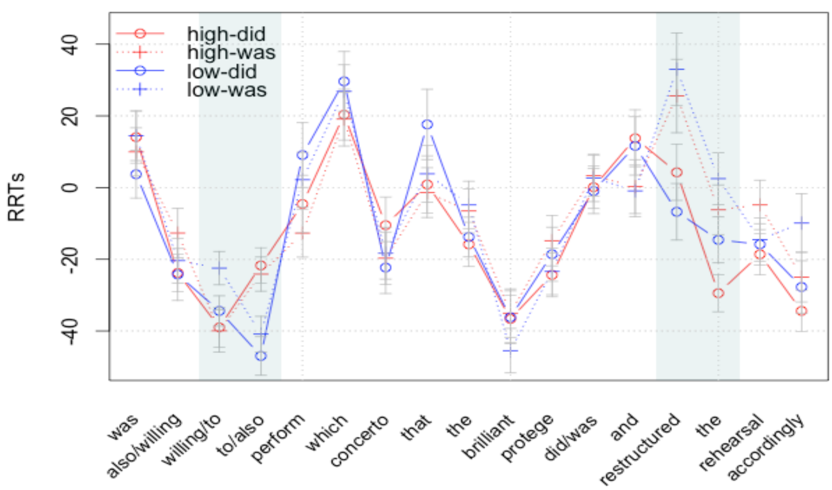


Figure 8.8
Residual reading times for target items in Experiment 3b

A linear mixed effects model with random slopes and intercepts for subjects and items was fit to the data. The model predicted RRTs from the two factors of interest: *position of also* (*low* vs. *high*) and *ellipsis size* (*small* vs. *large*). The results show a main effect of *also* at the third and fourth word in the region of interest (where *also* occurs in the sentence). The results additionally show a main effect of *ellipsis size* two and three words after the auxiliary verb

Predictor	Coefficient	Standard Error	*t* value
Intercept	-29.634	9.056	-3.272
Position of also	17.017	7.584	2.244
Ellipsis size	21.984	9.215	2.386
Position of also×*Ellipsis size*	-6.791	11.810	-0.575

Table 8.5
Results of Experiment 3b

and a main effect of *also* three words after the auxiliary (log likelihood tests, p's<0.05).[28] The main effect of *ellipsis size* reflects the fact that the resolution of *small ellipsis* is faster than the resolution of *large ellipsis* for both *also* conditions, while the main effect of the *also* is caused because the processing of the *high also* conditions is overall faster than the processing of the *low also* conditions. The results of the model for the third word after the auxiliary verb are summarized in Table 8.5.

Finally, an analysis was conducted pooling data from the *high also* condition across experiments 3a–b. A linear mixed effects model predicting RRTs from the two factors of interest: *determiner* (*every* vs. *which*) and ellipsis size (*small ellipsis* vs. *large ellipsis*) was fit to the data. The model included random slopes and intercepts for the factors of interest. The results show a main effect of *ellipsis size* two and three words after the auxiliary site, such that *small ellipsis* was processed more quickly than *large ellipsis* (log likelihood tests, p's<0.05). In addition, the results show a *determiner*×*ellipsis size* interaction two words after the auxiliary verb, driven by the fact that the difference between the *every* conditions is greater than the *which* conditions across the two *ellipsis size* conditions. The results of the model for the second word after the auxiliary verb are summarized in Table 8.6.[29]

Discussion

Several effects can be observed in both Experiments 3a–b. First, participants react similarly to reading the word '*also*' itself. We can take this to indicate that participants were paying attention to this experimental manipulation. Additionally, there is a main effect of *ellipsis size* in both experiments, occurring on the second and third words after the ellipsis site. This effect is more pronounced

Predictor	Coefficient	Standard Error	*t* value
Intercept	-25.08	14.32	-1.751
Determiner	34.87	11.12	1.810
Ellipsis size	52.06	19.27	4.680
Determiner×*Ellipsis size*	-30.06	15.20	-1.978

Table 8.6

Results of comparison of the *high also* conditions in Experiments 3a–b

on the second word following the auxiliary verb (see Figure 8.9). This is again consistent with the effects observed in Experiments 1 and 2a–b.

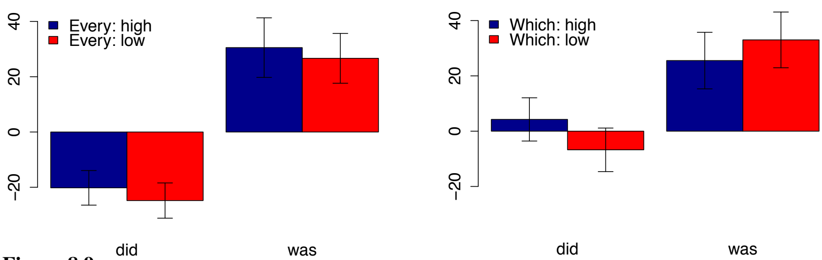


Figure 8.9

Residual reading times two words after AUX in Experiments 3a (*every*, left) and 3b (*which*, right)

In both Experiments 3a–b, we observe the main effect of *ellipsis size* not only on the second word following the auxiliary verb but also on the third word. More importantly, we find that on this word, the presence and position of *also* affects the resolution of ACD differently in the two experiments. Specifically, *also* does not have an effect in Experiment 3a (*every*), but in the case of Experiment 3b (*which*), we find a main effect of *also* in addition to the main effect of *ellipsis size*. To see this more clearly, consider Figure 8.10.

The main effect of *also* in Experiment 3b reflects a facilitation effect of non-local ACD resolution with *high also*, such that it is processed at a similar speed to that of the *local* ACD with *low also* (i.e. of the did-*low also* condition). This is consistent with our covert scrambling proposal for English *wh*-in-situ: if interpretability requirements force long-distance covert scrambling, we expect ACD to be facilitated in the entire domain of movement. In Experiment 3b, the *high also* forces scrambling above VP_1 and as a result, antecedent containment

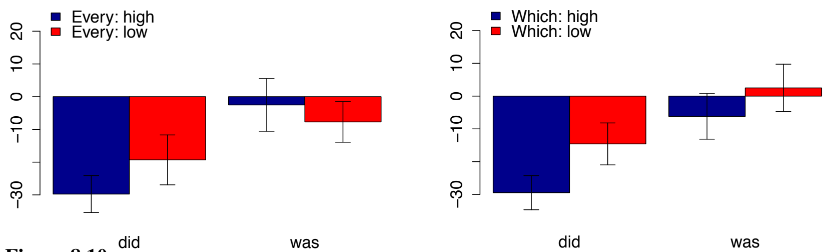


Figure 8.10
Residual reading times three words after AUX in Experiments 3a (*every*, left) and 3b (*which*, right)

is preemptively undone not only in the case of small ellipsis but also in the case of large ellipsis.

As discussed above, a comparison of the *high also* conditions across the two experiments yields an interaction, attributable to the fact that the difference between the *small* and *large* ellipsis conditions with *every* is greater than the difference between the two conditions with *which*. This, again, is predicted if *high also* facilitates the resolution of ellipsis in both the *small* and *large* conditions with *which*, but with *every* only *small* ellipsis resolution is facilitated.

8.3.4 An alternative account

Before concluding this chapter, I would like to address one alternative interpretation of the experimental results presented above. This alternative interpretation accepts the finding that *wh*-in-situ is in fact not interpreted in-situ at LF, but it attempts to reconcile the data with the traditional covert movement view, under which covert *wh*-movement necessarily targets interrogative C.[30]

Let us return to our sample item from Experiment 1, repeated here for convenience. Recall that this experiment compared the behavior of *every* with that of *which* with regard to facilitation of ACD resolution, and found that they facilitate ACD resolution to the same extent.

(289) *Sample target item in Experiment 1*

The conductor asked *which soloist* was [$_{\text{VP1}}$ willing to [$_{\text{VP2}}$ perform...

a. **every** concerto that the brilliant protégé *did* ... *(small ellipsis)*
b. **which** concerto that the brilliant protégé *did* ... *(small ellipsis)*
c. **every** concerto that the brilliant protégé *was* ... *(large ellipsis)*
d. **which** concerto that the brilliant protégé *was* ... *(large ellipsis)*

and restructured the rehearsal accordingly.

Our starting point here was the assumption that *every* undergoes a short QR movement step to a position above the lower VP_2—here, *perform*—in order to resolve a type-mismatch as soon as it is encountered by the parser. This assumption was based on the results of Hackl et al. (2012), surveyed above in section 8.2.2. Specifically, Hackl et al. (2012) showed that such movement indeed takes place in their items. However, the items used in the present study differ from Hackl et al.'s (2012) in that they are embedded inside a question.

Questions with quantifiers are ambiguous between two readings:

(290) *The two readings of a question with a quantifier:*
Which book did *every boy* read?

a. Which book is such that all the boys read it? *which* > ∀
b. For each boy, which book did he read? ∀ > *which*

The reading in (290a) involves narrow scope for *every* with respect to *which*, and is consistent with our assumptions that *every* only undergoes a short movement step. On the other hand, the reading in (290b) involves wide scope for *every* with respect to *which*, and would require *every* to take scope above the question. There are several theories of how this scope is obtained, but it is at least possible that it is obtained through long-distance covert movement of *every* (cf Chierchia 1993; Krifka 2006; Nicolae 2013).

The Hackl et al. (2012) methodology does not allow us to test which interpretation was accessed by the experimental participants. As a result, we must entertain the idea that at least some speakers accessed this reading for some of the items. This is made more plausible by the fact that multiple questions themselves admit a similar ambiguity, between a single-pair and a pair-list reading. At least the pair-list reading, and perhaps also the single-pair, have been argued to require covert movement for their interpretation (see chapter 3).

(291) *The two readings of a multiple question:*
Which book did *which* student read?

a. Which book and which student are such that the student read the book? *single-pair*
b. For each book, which student read it? *pair-list*

If the items in the present experiments were read with a pair-list reading, we might expect *every* to undergo long-distance movement to a position above the question—and hence above both the embedded VP_2 and the higher VP_1.

Consequently, ACD should be preemptively undone in both the *small* and *large* ellipsis conditions, and ACD resolution should be facilitated when the parser encounters both *did* and *was*. The fact that *every* patterns with *which*, then, teaches us not that *every* and *which* both underwent a short QR movement step just above VP_2, but instead that they both underwent long-distance movement, targeting a position higher than VP_1. The difference between the *small* and *large* ellipsis could be attributed to the fact that the larger ellipsis is richer and therefore may be more costly to integrate into the structure than a simpler antecedent, and not to any additional covert movement in the structure.

The results of Experiment 2 could be consistent with this view of interrogative syntax: we observe that *every* and *which* are both processed faster than corresponding sentences with *the*, and this could be explained both under a theory where both have undergone a short movement step (as I proposed above) or with the theory entertained here, that they both underwent a long-distance covert movement step.

However, this alternative proposal is inconsistent with the results of Experiment 3. If in-situ *wh*-phrases always covertly move to C, *also* should not be able to affect *which*, since sufficient movement to avoid an intervention effect would be performed regardless of the presence and position of *also*.

To maintain the view that *wh*-movement always targets C in view of the results of Experiment 3, we may propose that *also* interacts with *every* so as to slow down the processing of the items in Experiment 3a. One way this could come about is if *also* marks the possible scope of *every*: *low also* would force a short QR step, and *high also* would be compatible with a longer movement step. Crucially, we would have to assume that such a constraint does not affect the movement of *which*. However, if this were the case, we would expect an interaction in Experiment 3a, such that the *low also-was* condition suffers, as it would be the only condition in that experiment where movement of *every* is unable to preemptively undo antecedent containment. This is not the result observed in Experiment 3a: we instead observe a main effect of ellipsis size but no interaction. This result was replicated twice, as described in section 8.3.3, to ensure its validity.[31]

To summarize, the idea that questions with *every* are assigned a pair-list reading, consistent with a long-distance step of the quantifier, may be true in some cases; but a proposal that attempts to explain the experimental data in this chapter based solely on pair-list LFs in which *every* and *which* take non-local scope is not consistent with the results of Experiment 3. As a consequence, I propose to adopt the *covert scrambling* approach to *wh*-in-situ, allowing *wh*-in-situ to be interpreted not only at C but at other positions in the structure as well.

8.4 A Hamblin semantics with covert scrambling

The experiments presented here lend further support to the idea that (phonologically) in-situ *wh*-phrases in superiority-obeying English questions undergo a short covert scrambling step immediately upon being integrated into the structure, parallel to the overt scrambling step observed in German. Further movement can then be forced by interpretability considerations, such as the need to avoid intervention effects or to construct an appropriate antecedent for ellipsis.

Following Beck's (2006) influential proposal, *wh*-phrases are only sensitive to intervention effects if they are interpreted via focus alternative computation but not if they undergo (covert) movement to a position above the intervener, (292). The need to avoid intervention effects is thus the driving force behind the extended scrambling we observed in Experiment 3: *wh* is unable to remain below the intervener and be interpreted using Rooth-Hamblin alternatives, and instead may only project alternatives from a position above the intervener.

(292) *Only* wh*s interpreted via focus alternatives are subject to intervention:*

a. ✓ LF: [$_{\text{CP}}$ C ... *wh* **intervener** ... *t*]

b. * LF: [$_{\text{CP}}$ C ... **intervener** ... *wh*]

Finally, I introduce one additional extension to the proposal: the notion that *wh*-phrases may not be simple Hamblin sets of individuals, as suggested above, but instead *interrogative quantifiers* of some kind. Such proposals have been previously made for a variety of languages (cf Karttunen 1977, Baker 1970; Dornisch 2000 for Polish, Huang 1995; Kim 1991 for Korean, Rullmann and Beck 1998). This proposal is suggestive, given the parallels between *every* and *which* observed in Experiments 1–2. However, such a proposal is not able to explain the interaction with intervention effects observed for *wh*-phrases in Experiment 3b, nor why *every* is not similarly sensitive to intervention effects, as observed in Experiment 3a and independently known in the literature.[32]

A way to integrate both sets of results from this book, then, would be to suggest that *wh*-expressions have as their interpretation *a set of alternative quantificational interpretations* (i.e., a set of alternative $\langle et, t\rangle$ denotations). This could be implemented by replacing, for example, the set of individuals {Abby, Betty, ...} as the denotation of *who* with the type-lifted denotation in (294) (Partee 1986; Partee and Rooth 1983; Rooth and Partee 1982):

(293) Wh-*expressions as sets of type-lifted individual alternatives:*
$$\left\{\lambda P.\ P(\text{Abby}),\ \lambda P.\ P(\text{Betty}),\ \ldots\right\}$$

An even more ambitious proposal would be to allow *wh*-expressions to denote sets of existential quantifiers, modeled after Karttunen-style denotations for *wh*-words.[33]

(294) Wh-*expressions as sets of existential alternatives:*
$$\left\{\lambda P.\ \exists x\ (P(x)\ \&\ x = \text{Abby}),\ \lambda P\ .\ \exists x\ (P(x)\ \&\ x = \text{Betty}),\ \ldots\right\}$$

This provides an automatic explanation for the short movement of object *wh*-expressions observed in chapters 7 and 8, although at this point this solution is merely a technical one, and one that will perhaps only be allowed if the simpler set-of-individuals interpretation is unavailable. I leave for future work a more detailed argument in favor of this more complex quantificational set denotation over the more traditional one, as a set of individuals, or alternatively a proposal which motivates short scrambling in English and German on different grounds, independent of quantification.[34]

8.5 Appendix: Materials for Experiments 1–3

Below is the full list of items used in Experiments 1–3. The same templates were used in all three experiments, with different choices of determiner and placement/position of *also* as described in the body of the chapter. Recall that in all experiments, the sentences were presented on two lines, with the line break always placed immediately following the verbal complex.

1. The orderly learned which doctor was (also) planning to (also) monitor *the*/*every*/*which* patient that the duty nurse *did*/*was* and immediately updated the charts.
2. The principal determined which instructor was (also) able to (also) teach *the*/*every*/*which* class that the substitute teacher *did*/*was* and accordingly finalized the schedule.
3. The conductor asked which soloist was (also) willing to (also) play *the*/*every*/*which* piece that the brilliant protégé *did*/*was* and restructured the rehearsal accordingly.

4. The coordinator learned which tutor was (also) scheduled to (also) teach *the/every/which* topic that the Physics professor *did/was* and assigned them to classrooms.
5. The prosecutor asked which witness was (also) told to (also) discredit *the/every/which* defendant that the corrupt detective *did/was* but only one witness revealed anything.
6. The teacher found_out which student was (also) eager to (also) attend *the/every/which* trip that the class president *did/was* and organized the field trips accordingly.
7. The detective found_out which guard was (also) willing to (also) hassle *the/every/which* prisoner that the sadistic warden *did/was* and included the names in his report.
8. The analyst predicted which investor was (also) prepared to (also) buy *the/every/which* stock that the hedge fund *did/was* and then sent a_memo to the bank management.
9. The realtor asked which trainee was able to (also) show *the/every/which* property that the experienced secretary *did/was* but nobody was (also) available that weekend.
10. The carpenter asked which apprentice was (also) qualified to (also) use *the/every/which* technique that the licensed electrician *did/was* and then assigned personnel to projects.
11. The choreographer determined which dancer was (also) ready to (also) perform *the/every/which* dance_routine that the Russian ballerina *did/was* and then started the dance recital.
12. The organizers found_out which announcer was (also) willing to (also) cover *the/every/which* game that the notorious commentator *did/was* and then finalized the broadcasting schedule.
13. The librarian learned which teacher was (also) planning to (also) borrow *the/every/which* book that the visiting scholar *did/was* and accordingly shortened the loan periods.
14. The attorney clarified which witness was (also) supposed to (also) support *the/every/which* alibi that the undercover informant *did/was* and then gave his closing argument.

15. The dispatcher clarified which apprentice was (also) scheduled to (also) accompany *the/every/which* crew that the experienced engineer *did/was* and sent the crews on their way.
16. The programmer realized which update was (also) certain to (also) solve *the/every/which* problem that the old software *did/was* but surprisingly decided not to tell anyone.
17. The focus-group explained which discount was (also) likely to (also) attract *the/every/which* demographic that the Spring sale *did/was* and then several TV ads were launched.
18. The secretary found_out which professor was (also) going to (also) question *the/every/which* student that the disciplinary committee *did/was* and then scheduled the hearings.
19. The general forgot which unit was (also) scheduled to (also) attack *the/every/which* target that the nuclear submarine *did/was* and sent a messenger to headquarters.
20. The biologist discovered which reptile was (also) likely to (also) have *the/every/which* gene that the Tyrannosaurus Rex *did/was* and proposed additional tests.
21. The admiral specified which ship was (also) ordered to (also) attack *the/every/which* position that the navy jet *did/was* and then the joint army-navy exercise began.
22. The engineer explained which apprentice was (also) asked to (also) service *the/every/which* engine that the sick crew_member *did/was* and then called the train company.
23. The colonel explained which officer was (also) ordered to (also) interrogate *the/every/which* prisoner that the CIA agent *did/was* and then described what methods not to use.
24. The log showed which detective was (also) sent to (also) arrest *the/every/which* suspect that the FBI agent *did/was* and additionally where the arrest took place.
25. The detective discovered which mobster was (also) about to (also) blackmail *the/every/which* business that the street gang *did/was* and immediately informed his superiors.

26. The sheriff knew which marshal was (also) excited to (also) chase *the*/*every*/*which* fugitive that the state police *did*/*was* but doubted that the fugitives would be caught.
27. The scientist discovered which antibody was (also) likely to (also) attack *the*/*every*/*which* virus that the standard medication *did*/*was* but needed funding to complete her study.
28. The warden guessed which inmate was (also) trying to (also) smuggle *the*/*every*/*which* contraband that the corrupt guard *did*/*was* and therefore intensified the security screens.

See further discussion of long-distance scrambling/QR in Wurmbrand (To appear) and arguments against long-distance scrambling in Miyagawa (2011).

28. I thank Phil Brannigan (p.c.) for bringing this data to my attention.

29. Note that the *wh* here and in the d variant can be interpreted as *wh*-indefinites (Martin Hackl and Hagen Blix, p.c.).

30. See additional discussion and predictions for the behavior of short vs. long scrambling as A-movement vs. $\bar{\text{A}}$-movement, respectively, in Mahajan 1990.

31. Overt scrambling in German is also possible for other, information-structural, reasons. Such parallel movement may occur covertly in English; detecting whether this is in fact the case is left for future work.

Chapter 8

1. This chapter reports finding based on work conducted under the supervision of Martin Hackl; preliminary results were reported in Kotek and Hackl (2013). The research reported in this chapter was supported by National Science Foundation dissertation improvement grant No. 1251717.

2. As in other parts of the book, these LFs are simplified in some ways that are not relevant to the point at hand. I do not represent *v*P unless it is relevant to the LF, nor to I show subject raising from *v*P to TP, successive-cyclic *wh*-movement, or T-to-C movement.

3. How to define identity in the domain of ellipsis is a much debated question. For example, experiments have shown that voice mismatches are possible in some contexts but not others (e.g. Arregui, Clifton, Frazier, and Moulton 2006; Kehler 2001; Kertz 2010; San Pietro, Xiang, and Merchant 2012:a.o.). It has also been argued that such mismatches are only possible in VP ellipsis but not in sluicing (Merchant 2013). I will not attempt to contribute to the definition of identity here. I simply assume that traces count as identical for the purpose of ellipsis parallelism if they are bound from parallel positions (e.g. Fox 2002).

4. More specifically, for all participants included in the analysis, a regression equation predicting reading times from word length using all items is fit to adjust for differences in word length and differences in participantsâĂŹ natural reading rates. At each word position, the reading time predicted by the participantâĂŹs regression equation is subtracted from the actual measured reading time to obtain a residual reading time. Outlier results for each participant — in this case, beyond two standard deviations from the mean for a given condition and position — are excluded from the analysis (here, accounting for less than 5% of the data).

5. There is a debate in the sentence processing literature about whether or not the size of the antecedent site affects the processing of the ellipsis site (cf. Frazier and Clifton 2000, 2001; Martin and McElree 2008, 2009, 2011; Murphy 1985). Note that the results discussed here do not depend on the answer to this question. Although the current thinking in the literature suggests that a larger antecedent does not necessarily incur a larger cost in the processing of ellipsis, it is possible to attribute the increased processing cost that Hackl et al. (2012) observe for the non-local ACD cases either to the longer covert movement step associated with the *large ellipsis* condition, or to the fact that the *large ellipsis* involves a more complex VP and hence more complex meaning than the *small ellipsis*.

6. See graphs and discussion in Hackl et al. (2012) that show that this is indeed the case in their experiments. As we will see below, this is also the case in the experiments presented here.

7. Extensional types are shown just for the relevant part of the structure. For simplicity, I abstract away from the VP-internal subject hypothesis.

8. Note that I am illustrating movement here as targeting TP, since I am abstracting away from *v*P and *v*P-internal subjects. Nothing hinges on this assumption.

9. In fact, this assumption may be made even if we entertain a theory of ACD resolution that does not involve any movement.

10. To simplify the discussion, in what follows I ignore the highest predicate embedding the question. I thus refer to the lower VP inside the embedded question as the "embedded VP" and to the larger VP inside the embedded question as the "matrix VP."

11. Previous processing work, as well as rating and corpus studies, show that questions with D-linked *wh*-phrases (Pesetsky 1987) are easier to process than questions with bare *wh*-pronouns (Arnon, Snider, Hofmeister, Jaeger, and Sag 2006; Clifton, Fanselow, and Frazier 2006; Fanselow, Lenertova, and Weskott 2008; Featherston 2005a,b; Frazier and Clifton 2002; Hofmeister, Jaeger, Sag, Arnon, and Snider 2007). In order to make the experimental items as easy to process as possible, the items in Experiments 1–3 all use D-linked *wh*-phrases.

12. Testing on a number of standard monitors showed this method to consistently avoid line breaks inside the region of interest, and was furthermore found to be the most natural among several other line break options in a pilot study.

13. These sentences are not included in the analysis because the difference in line breaks made them no longer parallel to the target sentences.

14. Ibex: Internet Based Experiments, accessible at `http://spellout.net/ibexfarm/`.

15. The average completion times for all of the experiments we present here were above 25 minutes. Particularly slow reading times in our experiments tended to reflect long breaks and distracted behavior of the participants, introducing extraneous noise into the results. Particularly fast reading times were similarly contributed by distracted participants.

16. Although this is exclusion rate is quite high, I take it to be necessary in order to filter the relatively higher noise level in Amazon Mechanical Turk participants as compared to lab participants. This more parsimonious criterion helps ensure the validity of the results I report here.

17. Here and in the other experiments, the results remain statistically unaltered if this step is not performed.

18. RRTs were calculated based on a regression equation predicting reading time from word length using all words from all experimental items, except for the first word and the last word of the sentence.

19. Similar results are obtained for log-transformed RTs in the data. Such a transformation is performed to normalize the data and reduce the effect any outliers have on the results. I report the results for RRTs for convenience.

20. Note, in particular, that the *which-was* condition appears to be slower than the other three conditions. However, the *determiner*×*ellipsis size* interaction is not significant (p=0.344).

21. As we will see below, this effect consistently shows up on the second and third words after the verb. This was furthermore confirmed for this experiment in a pilot experiment using the same items presented here. See also discussion in section 8.2.3.

22. A more specified model that includes the effect of *determiner* yielded a false convergence.

23. However, an important difference between these results and Hackl et al.'s (2012) results is that while Experiment 2 yields two main effects, Hackl et al. observed an interaction, such that *every-was* and *the-was* were not significantly different from each other. This played a role in Hackl et al.'s argument that only local QR facilitates ACD resolution. Below, the results of Experiment 3 will provide us with the missing evidence to make this claim in this chapter here as well. I return to this point after introducing that experiment.

24. Although it has not been a major focus of attention in the literature, I argue that *also* acts as an intervener in English, exhibiting the same grammaticality paradigm as better known interveners such as *only*: A superiority-obeying question with *also* c-commanding the in-situ *wh*-phrase is grammatical, (iia), but a superiority-violating question with *also* c-commanding the in-situ *wh*-phrase is ungrammatical (iib).

    (i) *Baseline: sup.-obeying and sup-violating questions both grammatical*

    I know that the teacher punished some students last week. Do you know...

    a. *which* punishment the teacher gave ____ to *which* student on Friday?

    b. *which* student the teacher gave *which* punishment to ____ on Friday?

    (ii) Also *is an intervener in English:*

    I know that the teacher punished some students on Thursday. Do you know...

    a. *which* punishment the teacher **also** gave ____ to *which* student on Friday?

    b. * *which* student the teacher **also** gave *which* punishment to ____ on Friday?

    However, see Haida and Repp (2013) for an argument that *also* is not an intervener in German.

25. Several other options were entertained as possible interveners in Experiment 3. The use of sentential negation was rejected to avoid confounds due to *neg*-raising. The use of adnominal focus elements such as *only*-DPs was rejected because of the added complexity they would introduce, with the need to accommodate yet another participant in an already rich context. Universal quantifiers such as *every* could not be used because they were a part of the *determiner* manipulation. Other adverbs such as *never* or *always* presented challenges in keeping the items uniform across experiments. On the other hand, *also* allowed for minimal alterations to the items of Experiment 1, while still affording a clear truth-conditional difference between sentences with *also* in different positions. These design considerations eventually led to the choice of *also* as the intervener in Experiment 3.

26. There are several options for what element in (287)–(288) might be chosen by a reader to be the associate of *also*. One natural choice is *protégé*, which contrasts with the *soloist* occurring higher in the sentence. Other options native speakers have reported to me include the noun *concerto* and the verb *perform*. The actual choice made by readers is immaterial to the predictions for this experiment: all that matters is that *also* acts as a focus-sensitive operator, forcing movement above it in the case of *which* but not in the case of *every*.

27. A more specified model including the effect of *also* for items did not converge.

28. In one of the two individual iterations of this experiment, the result at this word additionally showed an *ellipsis size*×*also* interaction, driven by slower reading times of large ellipsis (<u>*was*</u>) with *low also* compared to the other three conditions. In the other iteration of the experiment and in the aggregate this interaction did not rise to the level of significance, and therefore I refrain from over-interpreting it here.

29. I thank Alexandre Cremers (p.c.) for suggesting this analysis to me.

30. I thank Alexandre Cremers for suggesting this idea and discussing its implications with me.

31. In fact, the interaction is never even close to significance.

32. Even if a theory of intervention compatible with this view of *wh*-phrases is put forth in the future, the main conclusion about the nature of covert *wh*-movement nonetheless stands: movement need not target interrogative C, but is normally shorter and more limited in scope.

33. This, then, leads to important and interesting questions concerning what other $\langle et, t\rangle$-elements might be included in this set. If, for example, the quantifier "no one" might be included, then the answer "no one" to a question such as "*who came?*" is a straightforward full answer to the question. Under current thinking within the Hamblin semantics tradition, this is not the case, but this remains an open and contested question.

34. I thank Chris Tancredi for proposing this point to me.

Chapter 9

1. In particular, this work suggests that all interveners are focus-sensitive; this may be appropriate for some interveners, such as *only*, *also*, and perhaps focus-sensitive negation, but it is less obviously appropriate for other interveners, such as *every*.

2. Likewise, if a theory of intervention effects based on choice functions as the affected mode of in-situ composition in a question is developed, or a theory couched within Inquisitive Semantics, this will again impact how the data may be interpreted.

3. Another point of related determinism discussed in Kotek and Erlewine (2016) concerns covert pied-piping in multiple questions: the constituent that is covertly pied-piped along with the *wh* is the largest among the options that would be available in overt pied-piping.

Bibliography

Abels, Klaus. 2012. *Phases: An essay on cyclicity in syntax (linguistische arbeiten 543)*. de Gruyter.

Akahane, Hitoshi. 2007. A phase-theoretic analysis of object omission. Ms.

Anderson, Catherine. 2004. The structure and real-time comprehension of quantifier scope ambiguity. Doctoral Dissertation, Northwestern University.

Aoun, Joseph, Norbert Hornstein, and Dominique Sportiche. 1981. Some aspects of wide scope quantification. *Journal of Linguistic Research* 1:67–95.

Aoun, Joseph, and Yen-hui Audrey Li. 1993. *Wh*-elements in situ: Syntax or LF? *Linguistic Inquiry* 24:199–238.

Aoun, Joseph, and Yen-hui Audrey Li. 2003. *Essays on the representational and derivational nature of grammar: The diversity of wh-constructions*. Linguistic Inquiry Monographs. MIT Press.

Arnon, Inbal, Neal Snider, Philip Hofmeister, Florian T. Jaeger, and Ivan A. Sag. 2006. Cross-linguistic variation in a processing account: The case of multiple *wh*-questions. In *Proceedings of BLS 32*.

Arregi, Karlos. 2002. Focus on Basque movements. Doctoral Dissertation, Massachusetts Institute of Technology.

Arregui, Ana, Charles Jr. Clifton, Lyn Frazier, and Keir Moulton. 2006. Processing elided verb phrases with flawed antecedents: The recycling hypothesis. *Journal of Memory and Language* 55:232–246.

Baayen, R. H. 2004. Statistics in psycholinguistics: A critique of some current gold standards. In *Mental Lexicon Working Papers*, 1–45. Edmonton.

Bachrach, Asaf, and Roni Katzir. 2009. Right-node raising and delayed spellout. In *Interphases: Phase-theoretic investigations of linguistic interfaces*, ed. Kleanthes K. Grohmann. Oxford, UK: Oxford University Press.

Baker, Carol L. 1970. Notes on the description of English questions: the role of an abstract question morpheme. *Foundations of Language* 6:197–217.

Barr, Dale J., Roger Levy, Christoph Scheepers, and Harry Tily. 2013. Random effects structure for confirm-atory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68:225–278.

Barss, Andrew. 2000. Minimalism and asymmetric *wh*-interpretation. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 31–52. MIT Press.

Bates, Douglas M., and Deepayan Sarkar. 2007. lme4: Linear mixed-effects models using S4 classes. R package version 0.9975-12, URL http://CRAN.R-project.org/.

Beaver, David, and Brady Clark. 2008. *Sense and sensitivity: How focus determines meaning*. Oxford, UK: Wiley-Blackwell.

Beck, Sigrid. 1996a. Quantified structures as barriers for LF movement. *Natural Language Semantics* 4:1–56.

Beck, Sigrid. 1996b. *Wh*-construction and transparent Logical Form. Doctoral Dissertation, Universität Tübingen.

Beck, Sigrid. 2006. Intervention effects follow from focus interpretation. *Natural Language Semantics* 14:1–56.

Beck, Sigrid, and Shin-Sook Kim. 1997. On *wh*-and operator scope in Korean. *Journal of East Asian Linguistics* 6:339–384.

Beck, Sigrid, and Shin-Sook Kim. 2006. Intervention effects in alternative questions. *Journal of Comparative German Linguistics* 9:165–208.

Bennett, Ryan. 2009. Two subject asymmetries in defaka focus extraction. Qualifying paper, Rutgers University.

Bennett, Ryan, Akinbiyi Akinlabi, and Bruce Connell. 2012. Two subject asymmetries in defaka focus constructions. In *Proceedings of WCCFL 29*.

Bever, Tom. 1970. The cognitive basis for linguistic structure. In *Cognition and the development of language*, ed. J. R. Hayes, 279–362. New York: Wiley.

Bobaljik, Jonathan. 1995. Morphosyntax: The syntax of verbal inflection. Doctoral Dissertation, Massachusetts Institute of Technology.

Bolinger, Dwight. 1978. Asking more than one thing at a time. In *Questions*, 107–150. Dordrecht: Reidel.

Bošković, Željko. 2000. Sometimes in [Spec,CP], sometimes in situ. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 53–87. Cambridge, MA: MIT Press.

Bošković, Željko. 2002. On multiple *wh*-fronting. *Linguistic Inquiry* 33:351–383.

Bošković, Željko. 2003. On the interpretation of multiple questions. *Linguistic Variation Yearbook* 1:1–15.

Bošković, Željko. 2007. A note on *wh* typology. In *Proceedings of FDSL 6*, 159–170.

Bouton, Lawrence F. 1970. Antecedent-contained pro-forms. In *Papers from the Sixth regional meeting, Chicago Linguistic Society*. Chicago, IL.

Branan, Kenyon. 2017a. In-situ *wh*-phrases in superiority violating contexts don't have to be in-situ. In *A pesky set: Papers for David Pesetsky*, ed. Claire Halpert, Hadas Kotek, and Coppe van Urk, volume 80, 353–359. Cambridge, MA: MITWPL.

Branan, Kenyon. 2017b. Prosodic intervention. Manuscript, MIT.

Breakstone, Micha, Alexandre Cremers, Danny Fox, and Martin Hackl. 2011. On the analysis of scope ambiguities in comparative constructions: Converging evidence from real-time sentence processing and offline data. In *Proceedings of SALT 21*, 712–231.

Brody, Michael. 1990. Some remarks on the focus field in Hungarian. In *Ucl working papers in linguistics*, volume 2.

Bruening, Benjamin. 2001. Syntax at the edge: Cross-clausal phenomena and the syntax of passamaquoddy. Doctoral Dissertation, Massachusetts Institute of Technology.

Buccola, Brian, and Benjamin Spector. 2016. Modified numerals and maximality. *Linguistics and Philosophy* 39:151–199.

Büring, Daniel. 1996. Drinking, accents, and negation. In *Proceedings of Workshop on Focus*, ed. Elena Benedicto, Maribel Romero, and Satoshi Tomioka, number 21 in University of Massachusetts Occasional Papers in Linguistics, 37–50. UMass, Amherst: GLSA.

Büring, Daniel. 2003. On D-trees, beans, and B-accents. *Linguistics and Philosophy* 26:511–545.

Butler, Alastair. 2001. Intervention effects in English questions. In *Proceedings of Szklarska Poreba Workshop 2*.

Cable, Seth. 2007. The grammar of Q. Doctoral Dissertation, Massachusetts Institute of Technology.

Cable, Seth. 2010. *The grammar of Q: Q-particles, wh-movement, and pied-piping*. Oxford, UK: Oxford University Press.

Chang, Lisa. 1997. Wh-in-situ phenomena in French. Master's thesis, University of British Columbia.

Charlow, Simon. 2017. The scope of alternatives: Indefiniteness and islands. Manuscript, Rutgers.

Cheng, Lisa Lai-Shen. 1991. On the typology of *wh*-questions. Doctoral Dissertation, Massachusetts Institute of Technology.

Cheng, Lisa Lai-Shen, and Hamida Demirdache. 2010. Trapped at the edge: On long-distance pair-list readings. *Lingua* 120:463–480.

Chierchia, Gennaro. 1993. Questions with quantifiers. *Natural Language Semantics* 1:181–234.

Chomsky, Noam. 1973. Conditions on transformations. In *A festschrift for Morris Halle*. New York: Holt, Reinhart, and Winston.

Chomsky, Noam. 1981. *Lectures on government and binding*. Foris.

Chomsky, Noam. 1992. A minimalist program for linguistic theory. In *MIT Working Papers in Linguistics*. Cambridge, Mass.

Chomsky, Noam. 1993. A minimalist program for linguistic inquiry. In *The view from Building 20*, ed. Kenneth Hale and Samuel Jay Keyser. MIT Press.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1998. Minimalist inquiries: the framework. *MIT Occasional Papers in Linguistics* 15.

Chomsky, Noam. 1999. Derivation by phase. *MIT Occasional Papers in Linguistics* 18.

Chomsky, Noam. 2000. Minimalist inquiries: the framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–155. MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: a life in language*. MIT Press.

Chomsky, Noam. 2008. On phases. In *Foundational issues in linguistic theory*, ed. Robert Freidin, Carlos Otero, and Maria Luisa Zubizarreta, 133–166. MIT Press.

Ciardelli, Ivano, Jeroen Groenendijk, and Floris Roelofsen. 2017a. Inquisitive semantics: A new notion of meaning. Book manuscript.

Ciardelli, Ivano, Floris Roelofsen, and Nadine Theiler. 2017b. Composing alternatives. *Linguistics and Philosophy* 40:1–36.

Clifton, Charles Jr., Gisbert Fanselow, and Lyn Frazier. 2006. Amnestying superiority violations: processing multiple questions. *Linguistic Inquiry* 37:51–68.

Clifton, Charles Jr., Lyn Frazier, and Patricia Deevy. 1999. Feature manipulation in sentence comprehentsion. *Rivista di Linguistica* 11:11–39.

Cole, Peter, and Gabriella Hermon. 1998. The typology of *wh*-movement. *Wh*-questions in Malay. *Syntax* 1:221–258.

Collins, Chris. 1996. *Local economy*. Cambridge, MA: MIT Press.

Collins, Chris. 2001. Economy conditions in syntax. In *Handbook of syntactic theory*, ed. Chris Collins and Mark R. Baltin. Blackwell publishing.

Comorovski, Ileana. 1989. Discourse and the syntax of multiple constituent questions. Doctoral Dissertation, Cornell University.

Constant, Noah. 2014. Contrastive topic: meanings and realizations. Doctoral Dissertation, University of Massachusetts Amherst.

Culicover, Peter W. 1992. Topicalization, inversion, and complementizers in English. In *Going romance and beyond*. OTS Working Papers in Linguistics 91.

Dayal, Veneeta. 1996. *Locality in wh quantification*. Kluwer Academic Publishers.

Dayal, Veneeta. 2002. Single-pair versus multiple-pair answers: *wh*-in-situ and scope. *Linguistic Inquiry* 33:512–520.

Dayal, Veneeta. 2016. *Questions*. Oxford, UK: Oxford University Press.

Diesing, Molly. 1992. *Indefinites*. Cambridge, MA: MIT Press.

Dornisch, Ewa. 2000. Overt quantifier raising in Polish. In *Proceedings of Generative Linguistics in Poland GLiP*, volume 1, 47–58.

van Dyke, Julie A., and Brian McElree. 2006. Retrieval interference in sentence comprehension. *Journal of Memory and Language* 55:157–166.

É Kiss, Katalin. 1986. Against the LF-movement of *wh*-phrases. Ms., Hungarian Academy of Sciences, Budapest.

É Kiss, Katalin. 1987. *Configurationality in hungarian*. Dordrecht: Reidel.

É Kiss, Katalin. 2002. *The syntax of Hungarian*. Cambridge University Press.

Eilam, Aviad. 2008. Intervention effects: Why Amharic patterns differently. In *Proceedings of WCCFL 27*, ed. Natasha Abner and Jason Bishop, 141–149. Somerville, MA: Cascadilla Proceedings Project.

Elliott, Patrick. 2015. Nested *wh*-questions and the locality of scope-taking. Slides from the workshop "Questions at the Syntax-Semantics Interface", University College London.

Engdahl, Elisabet. 1986. *Constituent questions: With special reference to swedish*. Dordrecht: Reidel.

Epstein, Sam. 1992. Derivational constraints on A' chain formation. *Linguistic Inquiry* 23:235–259.

Erlewine, Michael Yoshitaka. 2014. Movement Out of Focus. Doctoral Dissertation, Massachusetts Institute of Technology.

Erlewine, Michael Yoshitaka, and Hadas Kotek. 2014. Intervention in focus pied-piping. In *Proceedings of NELS 43*, ed. Hsin-Lun Huang, Ethan Poole, and Amanda Rysling, volume 1, 117–130. Amherst: GLSA.

Erlewine, Michael Yoshitaka, and Hadas Kotek. 2016. Even-NPIs in Dharamsala Tibetan. *Linguistic Analysis* 40. Special issue on formal approaches to South Asian languages.

Erlewine, Michael Yoshitaka, and Hadas Kotek. to appear. Focus association through movement: Evidence from Tanglewood. *Linguistic Inquiry* .

Erteschik-Shir, Nomi. 1973. On the nature of island constraints. Doctoral Dissertation, Massachusetts Institute of Technology.

Fanselow, Gisbert, Denisa Lenertova, and Thomas Weskott. 2008. Studies on the acceptability of object movement to Spec,CP. In *Language content & cognition: The discourse potential of underspecified structures*, ed. Anita Steube. Mouton de Gruyter.

Featherston, Sam. 2005a. Magnitude estimation and what it can do for your syntax: some *wh*-constraints in German. *Lingua* 115:1525–1550.

Featherston, Sam. 2005b. Universals and grammaticality: *Wh*-constraints in German and English. *Linguistics* 43:667–711.

von Fintel, Kai, Danny Fox, and Sabine Iatridou. 2014. Definiteness as maximal informativeness. In *The art and craft of semantics: A festschrift for Irene Heim*, ed. Luka Crnič and Uli Sauerland, volume 1, 165–174. MITWPL.

Ford, Marylin, Joan Bresnan, and Ron Kaplan. 1982. A competence-based theory of syntactic closure. In *The mental representation of grammatical relations*, ed. Joan Bresnan, 727–796. Cambridge, Mass.: MIT Press.

Fox, Danny. 1995. Economy and scope. *Natural Language Semantics* 3:283–341.

Fox, Danny. 2000. *Economy and semantic interpretation*. Cambridge, Mass.: MIT Press.

Fox, Danny. 2002. Antecedent-contained deletion and the copy theory of movement. *Linguistic Inquiry* 33:63–96.

Fox, Danny. 2003. *On Logical Form*, chapter 2, 82–123. Oxford, UK: Blackwell Publishers.

Fox, Danny. 2012. More on questions. Class notes, MIT seminar.

Fox, Danny, and Martin Hackl. 2006. The universal density of measurement. *Linguistics and Philosophy* 29:537—586.

Fox, Danny, and Jon Nissenbaum. 1999. Extraposition and the nature of covert movement. In *Proceedings of WCCFL 18*, ed. Sonya Bird, Andrew Carnie, Jason D. Haugen, and Peter Norquest, 132–144. Somerville, MA: Cascadilla Press.

Frampton, John. 1991. Relativized minimality: A review. *The Linguistic Review* 8:1–46.

Frampton, John, and Sam Gutman. 1999. Cyclic computation. *Syntax* 2:1–27.

Frazier, Lyn. 1999. *On sentence interpretation*. Dordrecht: Kluwer.

Frazier, Lyn, and Charles Jr. Clifton. 2000. On bound variable interpretations: The LF-only hypothesis. *Journal of Psycholinguistic Research* 29:125–139.

Frazier, Lyn, and Charles Jr. Clifton. 2001. Parsing coordinates and ellipsis: Copy α. *Syntax* 4:1–22.

Frazier, Lyn, and Charles Jr. Clifton. 2002. Processing D-linked phrases. *Journal of Psycholinguistic Research* 31:633–660.

Frazier, Lyn, and D. Fodor, Janet. 1978. The sausage machine: A new two-stage parsing model. *Cognition* 6:291–325.

Frazier, Lyn, and Keith Rayner. 1982. Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive psychology* 14:178–210.

Gärtner, Hans-Martin. 2002. *Generalized transformations and beyond: Reflections on minimalist syntax*. Berlin: Akademie Verlag.

George, Benjamin R. 2011. Question embedding and the semantics of answers. Doctoral Dissertation, University of California at Los Angeles.

Giannakidou, Anastasia. 1997. The landscape of polarity items. Doctoral Dissertation, University of Groningen.

Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68:1–76.

Golan, Yael. 1993. Node crossing economy, superiority and D-linking. Ms., Tel-Aviv University.

Gordon, Peter C., Randall Hendrick, and Marcus Johnson. 2004. Effects of noun phrase type on sentence complexity. *Journal of Memory and Language* 51:97–114.

Graf, Thomas. 2013. Local and transderivational constraints in syntax and semantics. Doctoral Dissertation, University of California Los Angeles.

Groat, Erich, and John O'Neil. 1996. Spell-out at the LF-interface. In *Minimal ideas*, ed. Werner Abraham, Sam Epstein, Höskuldur Thráinsson, and Jan-Wouter Zwart, 113–139. Amsterdam: John Benjamins.

Groenendijk, Jeroen. 2009. Inquisitive semantics: Two possibilities for disjunction. In *Seventh International Tbilisi Symposium on Language, Logic, and Computation*, ed. P. Bosch, D. Gabelaia, and J. Lang. Springer-Verlag.

Groenendijk, Jeroen, and Floris Roelofsen. 2009. Inquisitive semantics and pragmatics. In *Proceedings of SPR 09*.

Groenendijk, Jeroen, and Martin Stokhof. 1984. Studies on the semantics of questions and the pragmatics of answers. Doctoral Dissertation, University of Amsterdam.

Hackl, Martin, Jorie Koster-Hale, and Jason Varvoutis. 2012. Quantification and ACD: Evidence from real-time sentence processing. *Journal of Semantics* 29:145–206.

Hagstrom, Paul. 1998. Decomposing questions. Doctoral Dissertation, Massachusetts Institute of Technology.

Haida, Andreas, and Sophie Repp. 2013. Focus alternatives or indefinite alternatives? experimental evidence. In *Proceedings of the 19th Amsterdam Colloquium*, ed. Maria Aloni, Michael Franke, and Floris Roelofsen, 131–138.

Hallman, Peter J. 1997. Reiterative syntax. Master's thesis, University of California at Los Angeles.

Hamblin, Charles. 1973. Questions in Montague English. *Foundations of Language* 10:41–53.

Heck, Fabian. 2008. *On pied-piping: wh-movement and beyond*. Berlin: Mouton de Gruyter.

Heck, Fabian, and Gereon Mueller. 2000. Repair-driven movement and the local optimization of derivations. Ms., University of Stuttgart and IDS Mannheim.

Heim, Irene. 1994. Interrogative semantics and Karttunen's semantics for *know*. In *Proceedings of IATL 1*.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Blackwell.

Hofmeister, Philip, Florian T. Jaeger, Ivan A. Sag, Inbal Arnon, and Neal Snider. 2007. Locality and accessibility in *wh*-questions. In *Roots: Linguistics in search of its evidential base*, ed. Sam Featherston and Wolfgang Sternefeld. Mouton de Gruyter.

Hoji, Hajime. 1985. Logical Form constraints and configurational structures in Japanese. Doctoral Dissertation, University of Washington.

Hornstein, Norbert. 1995. *Logical Form: from GB to Minimalism*. Cambridge, Mass.: Blackwell Publishers.

Horvath, Julia. 1986. *FOCUS in the theory of grammar and the syntax of Hungarian*. Foris, Dordrecht.

Huang, C. T. James. 1995. Logical Form. In *Government and binding theory and the minimalist program*, ed. G. Webelhuth, 127–173. Blackwell.

Huang, Cheng-Teh James. 1982a. Logical relations in Chinese and the theory of grammar. Doctoral Dissertation, Massachusetts Institute of Technology.

Huang, Cheng-Teh James. 1982b. Move *wh* in a language without *wh* movement. *The Linguistic Review* 1.

Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. MIT Press.

Johnson, David, and Shalom Lappin. 1999. *Local constraints vs. economy*. Stanford, CA: CSLI Publications.

Johnson, Kyle. 2000a. How far will quantifiers go? In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 187–210. MIT Press.

Johnson, Kyle. 2000b. When verb phrases go missing. In *The first glot international state-of-the-article book*, ed. Lisa Lai-Shen Cheng and Rint P. E. Sybesma.

Johnson, Kyle, and Satoshi Tomioka. 1997. Lowering and mid-size clauses. In *Proceedings of the 1997 Tübingen Workshop on Reconstruction*, ed. Graham Katz, Shin-Sook Kim, and Winhart Haike, 185–206.

Just, Marcel A., Patricia A. Carpenter, and Jacqueline D. Wooley. 1982. Paradigms and processes and in reading comprehension. *Journal of Experimental Psychology: General* 111:228–238.

Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* 1:3–44.

Kazanina, Nina, and Colin Phillips. 2000. Differential effects of constraints in the processing of russian cataphora. *The Quarterly Journal of Experimental Psychology* 63:371–400.

Kehler, Andrew. 2001. *Coherence, reference, and the theory of grammar. center for the study of language and information*. CSLI Publications.

Kennedy, Christopher. 1997. Antecedent-contained deletion and the syntax of quantification. *Linguistic Inquiry* 4:662–688.

Kertz, Laura. 2010. Ellipsis reconsidered. Doctoral Dissertation, University of California Davis.

Kim, Shin-Sook. 2002a. Focus matters: Two types of intervention effects. In *Proceedings of WCCFL 21*.

Kim, Shin-Sook. 2002b. Intervention effects are focus effects. In *Proceedings of Japanese/Korean Linguistics 10*, 615–628.

Kim, Soowon. 1991. Chain scope and quantification structure. Doctoral Dissertation, Brandeis University.

Kishimoto, Hideki. 2005. *Wh*-in-situ and movement in Sinhala questions. *Natural Language & Linguistic Theory* 23:1–51.

Kitahara, Hisatsugu. 1996. Raising Q without QR. In *Minimal ideas*, ed. Werner Abraham, Sam Epstein, Höskuldur Thráinsson, and Jan-Wouter Zwart. Amsterdam: John Benjamins.

Kitahara, Hisatsugu. 1997. *Elementary operations and optimal derivations*. Cambridge, Mass.: MIT Press.

Kobele, Gregory, and Harold Torrence. 2006. Intervention and focus in Asante Twi. In *ZAS papers in Linguistics*, volume 46, 161–184.

Kotek, Hadas. 2012. Readings of Hebrew multiple questions. In *Proceedings of WCCFL 30*, ed. Nathan Arnett and Ryan Bennett, 216–225. Somerville, MA, USA: Cascadilla Proceedings Project.

Kotek, Hadas. 2014a. Composing questions. Doctoral Dissertation, Massachusetts Institute of Technology.

Kotek, Hadas. 2014b. *Wh*-fronting in a two-probe system. *Natural Language & Linguistic Theory* 32:1105–1143.

Kotek, Hadas. 2016. Covert partial *wh*-movement and the nature of derivations. *Glossa* 1(1):1–19.

Kotek, Hadas. 2017a. Dissociating intervention effects from superiority in English *wh*-questions. *The Linguistic Review* 34:397–417.

Kotek, Hadas. 2017b. Intervention effects arise from scope-taking over alternatives. In *Proceedings of NELS 47*, ed. Andrew Lamont and Katerina Tetzloff, volume 2, 153–166. Amherst, MA: GLSA.

Kotek, Hadas. 2017c. Questioning superiority. In *A pesky set: Papers for David Pesetsky*, ed. Claire Halpert, Hadas Kotek, and Coppe van Urk, volume 80, 457–466. MITWPL.

Kotek, Hadas, and Michael Yoshitaka Erlewine. 2016. Covert pied-piping in English multiple *wh*-questions. *Linguistic Inquiry* 47:669–693.

Kotek, Hadas, and Martin Hackl. 2013. An experimental investigation of interrogative syntax/semantics. In *Proceedings of the Amsterdam Colloquium*, ed. Maria Aloni, Michael Franke, and Floris Roelofsen.

Kratzer, Angelika. 1991. The representation of focus. In *Semantik/semantics: An international handbook of contemporary research*, ed. Arnim von Stechow and Dieter Wunderlich, 825–834. Berlin: Mouton de Gruyter.

Kratzer, Angelika. 1995. Stage-level vs. individual level predicates. In *The generic book*, ed. Gregory Carlson and Francis Pelletier, 125–175. Chicago: Chicago University Press.

Kratzer, Angelika, and Junko Shimoyama. 2002. Indeterminate pronouns: the view from Japanese. In *Proceedings of the 3rd Tokyo conference on psycholinguistics*, 1–25.

Krifka, Manfred. 2001. Quantifying into question acts. *Natural Language Semantics* 9:1–40.

Krifka, Manfred. 2006. Association with focus phrases. In *The architecture of focus*, 105–136. Mouton de Gruyter.

Kuno, Susumu, and Jane J. Robinson. 1972. Multiple *wh* questions. *Linguistic Inquiry* 3.

Kuroda, Sige-Yuki. 1965. Generative grammatical studies in the japanese language. Doctoral Dissertation, Massachusetts Institute of Technology.

Ladusaw, William A. 1980. *Polarity sensitivity as inherent scope relations*. Garland Publishing.

Larson, Richard K., and Robert Carlen May. 1990. Antecedent containment or vacuous movement: Reply to Baltin. *Linguistic Inquiry* 21:103–122.

Lasnik, Howard, and Mamoru Saito. 1984. On the nature of proper government. *Linguistic Inquiry* 15:235–290.

Lasnik, Howard, and Mamoru Saito. 1992. *Move α, conditions on its application and output*. Cambridge, Mass.: MIT Press.

Lebeaux, David. 2009. *Where does binding theory apply*. Linguistic Inquiry Monographs. MIT Press.

Lechner, Winfred. 2013. The syntax-semantics interface. In *Syntax, an international handbook*, ed. Tibor Kiss and Artemis Alexiadou. Mouton de Gruyter, 2nd edition.

Li, Haoze, and Jess Law. 2014a. Focus intervention: A quantificational domain approach. In *Proceedings of NELS 44*.

Li, Haoze, and Jess Law. 2014b. Generalized focus intervention. In *Proceedings of SALT 24*, 473–493.

Li, Haoze, and Jess Law. 2016. Alternatives in different dimensions: A case study of focus intervention. *Linguistics and Philosophy* 39:201–245.

Liptak, Aniko. 2000. On the difference between focus movement and wh-movement. *GLOW Newsletter* 44:78–79.

MacDonald, Maryellen C., Neal J. Perlmutter, and Mark S. Seidenberg. 1994. Lexical nature of syntactic ambiguity resolution. *Psychological review* 101:676–703.

Mahajan, Anoop K. 1990. The A/A-bar distinction and movement theory. Doctoral Dissertation, Massachusetts Institute of Technology.

Martin, Andrea E., and Brian McElree. 2008. A content-addressable pointer underlies comprehension of verb–phrase ellipsis. *Journal of Memory and Language* 58:879–906.

Martin, Andrea E., and Brian McElree. 2009. Memory operations that support language comprehension: Evidence from verb–phrase ellipsis. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 35:1231–1239.

Martin, Andrea E., and Brian McElree. 2011. Direct-access retrieval during sentence comprehension: Evidence from sluicing. *Journal of Memory and Language* 64:327–343.

Mathieu, Eric. 1999. *Wh* in situ and the intervention effect. In *UCL Working Papers in Linguistics*, ed. C. Iten and Ad Neeleman, 441–472.

Mathieu, Eric. 2002. The syntax of non-canonical quantification: A comparative study. Doctoral Dissertation, University College London.

May, Robert Carlen. 1985. *Logical Form: Its structure and derivation*. Linguistic Inquiry Monographs. MIT Press.

Mayr, Clemens. 2010. The role of alternatives and strength in grammar. Doctoral Dissertation, Harvard University.

Mayr, Clemens. 2014. Intervention effects and additivity. *Journal of Semantics* 31:513–554.

Merchant, Jason. 2001. *The syntax of silence*. Oxford University Press.

Merchant, Jason. 2013. Voice and ellipsis. *Linguistic Inquiry* 44:77–108.

Miyagawa, Shigeru. 2001. EPP, scrambling, and *wh*-in-situ. In *Ken Hale: a life in language*, ed. Michael Kenstowicz. Cambridge, Mass.: MIT Press.

Miyagawa, Shigeru. 2011. Optionality. In *The Oxford handbook of linguistic minimalism*, ed. Cedric Boeckx, 354–376. Oxford University Press.

Moltmann, Friederike, and Anna Szabolcsi. 1994. Scope interaction with pair-list quantifiers. In *Proceedings of NELS 24*, ed. Merce Gonzalez, 381–395. University of Massachusetts, Amherst: GLSA.

Murphy, Gregory. 1985. Processes of understanding anaphora. *Journal of Memory and Language* 24:290–303.

Nicolae, Andreea. 2013. Any questions? Doctoral Dissertation, Harvard University.

Nishigauchi, Taisuke. 1986. Quantification in syntax. Doctoral Dissertation, University of Massachusetts Amherst.

Nishigauchi, Taisuke. 1990. *Quantification in the theory of grammar*. Kluwer.

Nissenbaum, Jon. 2000. Investigations of covert phrase movement. Doctoral Dissertation, Massachusetts Institute of Technology.

Novel, Marc, and Maribel Romero. 2009. Movement, variables, and Hamblin alternatives. In *Proceedings of Sinn und Bedeutung 14*.

Partee, Barbara Hall. 1986. Noun phrase interpretation and type-shifting principles. In *Studies in discourse representation theory and the theory of generalized quantifiers*, ed. Jeroen Groenendijk, Dick de Jongh, and Martin Stokhof, 115–143. Dordrecht: Foris.

Partee, Barbara Hall, and Mats Rooth. 1983. Generalized conjunction and type ambiguity. In *Meaning, use and interpretation*. de Gruyter.

Pesetsky, David. 1987. *Wh*-in-situ: movement and unselective binding. In *The representation of (in)definiteness*. MIT Press.

Pesetsky, David. 1998. Some optimality principles of sentence pronunciation. In *Is the best good enough*, ed. Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky, 337–383. MIT Press.

Pesetsky, David. 2000. *Phrasal movement and its kin*. Cambridge, MA: MIT Press.

Phillips, Colin. 1996. Order and structure. Doctoral Dissertation, Massachusetts Institute of Technology.

Phillips, Colin. 2003. Linear order and constituency. *Linguistic Inquiry* 34:37–90.

Pica, Pierre, and William Snyder. 1995. Weak crossover, scope, and agreement in a minimalist framework. In *Proceedings of WCCFL 13*, ed. R. et al. Aranovich, 334–349. Stanford, CA: CSLI Publications.

Polinsky, Maria, Carlos Gomez Gallo, Peter Graff, and Ekaterina Kravtchenko. 2012. Subject preference and ergativity. *Lingua* 122:267–277.

Polinsky, Maria, and Eric Potsdam. 2014. Left edge topics in Russian and the processing of anaphoric dependencies. *Journal of Linguistics* 50:629–669.

Pope, Emily N. 1976. *Questions and answers in English*. The Hague: Mouton de Gruyter.

Potts, Christopher. 2001. Three kinds of transderivational constraints. In *Syntax at Santa Cruz*, ed. Séamas Mac Bhloscaidh, volume 3, 21–40. Santa Cruz, CA: Linguistics Department, UC Santa Cruz.

Preminger, Omer. 2011. Agreement as a fallible operation. Doctoral Dissertation, Massachusetts Institute of Technology.

Pritchett, Bradley L. 1992. *Grammatical competence and parsing performance*. The University of Chicago Press.

Ramchand, Gillian. 1997. Questions, polarity and alternative semantics. In *Proceedings of NELS 27*, 383–396.

Ratiu, Dafina. 2005. Questions multiples en Roumain. Master's thesis, Université de Nantes.

Ratiu, Dafina. 2007. Evidence for (counter) cyclic movement from Romanian. Paper presented at the 37th Linguistics Symposium on Romance Languages.

Reinhart, Tanya. 1997. Quantifier scope: How labour is divided between QR and choice functions. *Linguistics and Philosophy* 20:335–397.

Reinhart, Tanya. 1998. *Wh*-in-situ in the framework of the Minimalist Program. *Natural Language Semantics* 6:29–56.

Reinhart, Tanya. 2006. *Interface strategies: Optimal and costly computations*. Cambridge, Mass.: MIT Press.

Richards, Marc D. 2004. Object shift and scrambling in north and west Germanic: A case study in symmetrical syntax. Doctoral Dissertation, University of Cambridge.

Richards, Norvin. 1997. What moves where when in which language? Doctoral Dissertation, Massachusetts Institute of Technology.

Richards, Norvin. 2000. An island effect in Japanese. *Journal of East Asian Linguistics* 9:187–205.

Richards, Norvin. 2001. *Movement in language: Interactions and architectures*. Oxford University Press.

Richards, Norvin. 2010. *Uttering trees*. MIT Press.

Richards, Norvin. 2016. *Contiguity theory*. Cambridge, MA: MIT Press.

Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, MA: MIT Press.

Rizzi, Luigi. 2001. Relativized minimality effects. In *The handbook of contemporary syntactic theory*, ed. Mark R. Baltin and Chris Collins, 89–110. Blackwell, Oxford.

Roberts, Craige. 1996. Information structure: Towards an integrated formal theory of pragmatics. In *OSU working papers in linguistics: Papers in semantics*, ed. J.H Yoon and A. Kathol, volume 49, 91–136. The Ohio State University Department of Linguistics.

Roelofsen, Floris. 2011. Algebraic foundations for inquisitive semantics. In *Proceedings of the Third International Conference on Logic, Rationality, and Interaction*, ed. H. van Ditmarsch, J. Lang, and J. Shier, 233–243. Springer-Verlag.

Rooth, Mats. 1985. Association with focus. Doctoral Dissertation, University of Massachusetts, Amherst.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1:75–116.

Rooth, Mats, and Barbara Hall Partee. 1982. Conjunction, type ambiguity, and wide scope *or*. In *Proceedings of WCCFL 1*, ed. D. Flickenger, M. Macken, and N. Wiegand.

Ross, John Robert. 1967. Constraints on variables in syntax. Doctoral Dissertation, Massachusetts Institute of Technology.

Ross, John Robert. 1969. Guess who? In *Proceedings of CLS 5*, ed. Robert I. Binnick, Alice Davison, Georgia M. Green, and Jerry L. Morgan, 252–286. Chicago: Chicago Linguistic Society.

Rudin, Catherine. 1985. *Aspects of Bulgarian syntax: Complementizers and* wh *constructions*. Columbus, Ohio: Slavica.

Rudin, Catherine. 1988. On multiple questions and multiple *wh* fronting. *Natural Language and Linguistic Theory* 6:445–501.

Rullmann, Hotze, and Sigrid Beck. 1998. Reconstruction and the interpretation of *which*-phrases. In *Proceedings of the 1997 Tübingen Workshop*, ed. G. Katz, S.-S Kim, and H. Winhart.

Sabel, Joachim. 2003. *Malagasy as an optional multiple wh-fronting language*, 229–254. John Benjamins.

Sag, Ivan Andrew. 1976. Deletion and Logical Form. Doctoral Dissertation, Massachusetts Institute of Technology.

San Pietro, Steven A., Ming Xiang, and Jason Merchant. 2012. Accounting for voice mismatches in ellipsis. In *Proceedings of WCCFL 30*, ed. Nathan Arnett and Ryan Bennett, 303–312. Somerville, MA: Cascadilla Proceedings Project.

Shan, Chung-chieh. 2004. Binding alongside Hamblin alternatives calls for variable-free semantics. In *Proceedings of SALT 16*.

Shimoyama, Junko. 2001. *Wh*-constructions in Japanese. Doctoral Dissertation, University of Massachusetts Amherst.

Shimoyama, Junko. 2006. Indeterminate quantification in Japanese. *Natural Language Semantics* 14:139–173.

Simpson, Andrew. 2000. *Wh-movement and the theory of feature-checking*. John Benjamins.

Simpson, Andrew, and Tanmoy Bhattacharya. 2003. Obligatory overt *wh*-movement in a *wh*-in-situ language. *Linguistic Inquiry* 34:127–142.

Soh, Hooi Ling. 2005. *Wh*-in-situ in Mandarin Chinese. *Linguistic Inquiry* 36:143–155.

Sternefeld, Wolfgang. 1996. Comparing reference-sets. In *The role of economy principles in linguistic theory*, ed. Chris Wilder, Hans-Martin Gärtner, and Manfred Bierwisch. Berlin: Akademie Verlag.

Sulemana, Abdul-Razak. 2017. Q-particles and the nature of covert movement: evidence from Buli. Manuscript, MIT, April 2017.

Szabolcsi, Anna. 1981. The semantics of topic-focus articulation. In *Formal methods in the study of language*. Mathematisch Centrum, Amsterdam.

Szendrői, Kriszta. 2003. A stress-based approach to the syntax of Hungarian focus. *The Linguistic Review* 20:37–78.

Takano, Yuji. 1998. Object shift and scrambling. *Natural Language & Linguistic Theory* 16:817–889.

Tancredi, Chris. 1990. Not only EVEN, but even ONLY. Manuscript, MIT.

Tancredi, Chris. 1992. Deletion, deaccenting and presupposition. Doctoral Dissertation, Massachusetts Institute of Technology.

Tanenhaus, Michael K., and John C. Trueswell. 1995. Sentence comprehension. In *Speech, language, and communication. handbook of perception and cognition*, ed. Joanne L. Miller and Peter D. Eimas, volume 11, 217–262. San Diego, CA, US: Academic Press, 2nd edition.

Tomioka, Satoshi. 2007a. Intervention effects in focus: From a Japanese point of view. In *Proceedings of the 2nd Workshop on Prosody, Syntax, and Information Structure (WPSI 2)*, ed. Shinichiro Ishihara, 97–117. Potsdam: Potsdam University Press.

Tomioka, Satoshi. 2007b. Pragmatics of LF intervention effects: Japanese and Korean interrogatives. *Journal of Pragmatics* 39:1570–1590.

Toosarvandani, Maziar. 2008. *Wh*-movement and the syntax of sluicing. *Journal of Linguistics* 44:677–722.

Toosarvandani, Maziar. 2009. Ellipsis in Farsi complex predicates. *Syntax* 12:60–92.

Truckenbrodt, Hubert. 2013. An analysis of prosodic F-effects in interrogatives: prosody, syntax, and semantics. *Lingua* 124:131–175.

Tsai, Wei-tien Dylan. 1994. On nominal islands and LF extractions in Chinese. *Natural Language and Linguistic Theory* 12:121–175.

Tunstall, Susanne L. 1998. The interpretation of quantifiers: Semantics and processing. Doctoral Dissertation, University of Massachusetts Amherst.

Uegaki, Wataru. 2015. Interpreting questions under attitudes. Doctoral Dissertation, Massachusetts Institute of Technology.

Uegaki, Wataru. 2018. A unified semantics for the Japanese Q-particle *ka* in indefinites, questions and disjunctions. *Glossa: a journal of general linguistics* 3(1).

Wachowicz, Krystyna. 1974. On the syntax and semantics of multiple questions. Doctoral Dissertation, University of Texas.

Wagers, Matthew W., and Colin Phillips. 2009. Multiple dependencies and the role of the grammar in real-time comprehension. *Journal of Linguistics* 45:395–433.

Warren, Tessa, and Edward Gibson. 2002. The influence of referential processing on sentence complexity. *Cognition* 85:79–112.

Wexler, Ken, and Peter Culicover. 1980. *Formal principles of language acquisition*. Cambridge, MA: MIT Press.

Williams, Edwin. 1974. Rule ordering in grammar. Doctoral Dissertation, Massachusetts Institute of Technology.

Williams, Edwin. 1977. Discourse and Logical Form. *Linguistic Inquiry* 8:101–139.

Willis, Paul. 2008. The role of topic-hood in multiple-*wh* question semantics. In *Proceedings of WCCFL 27*, 87–95.

Wiltschko, Martina. 1997. D-linking, scrambling and superiority in German. In *Groninger Arbeiten zur germanistischen Linguistik*, ed. Werner Abraham, 107–142. Groningen: Germanistisch Institut.

Wold, Dag. 1996. Long distance selective binding: the case of focus. In *Proceedings of SALT 6*, ed. Teresa Galloway and Justin Spence, 311–328.

Wurmbrand, Susi. To appear. The cost of raising quantifiers. *Glossa* .

Xiang, Ming, Boris Harizanov, Maria Polinsky, and Ekaterina Kravtchenko. 2011. Processing morphological ambiguity: An experimental investigation of russian numerical phrases. *Lingua* 121:548–560.

Yang, Barry Chung-Yu. 2008. Intervention effects and the covert component of grammar. Doctoral Dissertation, National Tsing Hua University.

Yang, Barry Chung-Yu. 2011. Intervention effects and *wh*-construals. *Journal of East Asian Linguistics* 21.

Zentz, Jason. 2016. Forming *wh*-questions in Shona: A comparative Bantu perspective. Doctoral Dissertation, Yale University.

Zwicky, Arnold M. 1971. In a manner of speaking. *Linguistic Inquiry* 2:223–233.