A new syntax for multiple *wh*-questions:
Evidence from real time sentence processing

Hadas Kotek and Martin Hackl

Massachusetts Institute of Technology

Abstract

Current theories of interrogative syntax/semantics adopt two strategies for the interpretation of in-situ *wh*-phrases: *covert movement* and *in-situ interpretation*. The covert movement strategy is traditionally assumed to be all-or-nothing: the in-situ *wh*-phrase covertly moves to C or else stays in-situ and is interpreted in its base-generated position at LF. In this paper we argue that both traditional approaches to *wh*-in-situ cannot be maintained. We present evidence from real-time processing of English multiple *wh*-questions that in-situ *wh*-phrases require <u>both</u> covert movement and in-situ interpretation for their computation: an "in-situ" *wh*-phrase undergoes a short movement step, parallel to the behavior of traditional quantifiers, immediately upon integration into the structure. Following that step, the *wh*-phrase *can* but *need not* move any further: it can be interpreted in its landing site and make its contribution to interrogative semantics without any further movement. We propose the *partial movement* approach to *wh*-in-situ: A *wh*-phrase must be interpreted at a position with propositional type at LF—for *wh*-phrases in object position, this position is never the base-generated position and instead some covert movement is always required.

Keywords: multiple *wh*-questions, covert movement, in-situ *wh*-phrases, intervention effects

1 Introduction

In a language like English, the formation of a question involves at least two steps. First, a structure is formed in which a *wh-phrase* is produced in an argument position of a verb (1a). Second, that *wh*-phrase is *overtly moved* to the left edge of the sentence (1b). In a *multiple question* only one *wh*-phrase is pronounced at the left edge of the sentence, while the remaining *wh*-phrase(s) are pronounced *in-situ*, in, what appears to be, their base-generated positions (1c).

(1)   **Single and multiple questions**
        a. Fred introduced *which student* to Mary?
        b. *Which student* did Fred introduce ___ to Mary?
        c. *Which student* did Fred introduce ___ to *which professor*?

In the literature on the syntax and semantics of *wh*-questions, there are two schools of thought about the analysis of in-situ *wh*-phrases. Both approaches assume that *wh*-phrases are interpreted by an interrogative head C occurring at the left edge of the question at LF, but they differ in the mechanisms by which C is assumed to create interrogative meanings. Under the first of these

approaches, all *wh*-phrases must be structurally adjacent to the C head that interprets them. This approach thus predicts covert movement of in-situ *wh*-phrases as a prerequisite for their interpretation. The second approach invokes mechanisms that interpret *wh*-phrases without any movement.

Although these approaches differ substantially in their treatment of interrogative syntax and semantics, traditional means of investigation have not been successful in choosing a preferred account. In particular, these approaches differ in (a) how much covert movement they predict in the derivation of a multiple question, (b) what they assume could cause covert movement, and (c) what syntactic position is targeted by covert movement.

The goal of this paper is to investigate the possible position(s) of in-situ *wh*-phrases at LF. We will argue that both current approaches to *wh*-in-situ are insufficient to derive the correct syntax-semantics of *wh*-questions. We present the results of three self-paced reading experiments that show that the in-situ *wh* in English multiple questions minimally requires a short movement step, comparable to the behavior of quantifiers like *every*, but that from that point on *wh*-phrases *can* but *need not* move any further. These results are unpredicted by both approaches to in-situ *wh*-phrases, motivating a new proposal for interrogative syntax/semantics. Specifically, we argue that the results can be explained by a *partial movement* approach to *wh*-in-situ, but not under the traditional approaches.

The remainder of the paper is structured as follows. In section 2, we present the two traditional approaches to the interpretation of *wh*-in-situ and discuss the predictions these theories make regarding the source and extent of covert *wh*-movement in a multiple question. We also present the phenomenon of *Antecedent Contained Deletion* (ACD), which has been argued to require covert movement in its derivation, and the methodology that will be used in the experiments in this paper. In section 3, we present three real-time sentence processing experiments that test the differing predictions made by the two approaches to *wh*-in-situ using the environment of ACD and intervention effects. We propose a new approach to *wh*-in-situ and show that it makes correct predictions about the processing of multiple *wh*-questions. Section 4 offers a general discussion of the experimental findings and the architecture of the theory that they entail.

2 Theoretical background

In this section we introduce the background for the theory we develop in this paper. Section 2.1 presents the two traditional approaches to *wh*-in-situ, covert movement and in-situ interpretation, and discusses the predictions these theories make regarding the source and extent of covert *wh*-movement in multiple questions. Section 2.2 presents the phenomenon of Antecedent Contained Deletion (ACD)*,* which will be central to our experiments. Section 2.3 presents the methodology that will be used in the experiments, and section 2.4 spells out the predictions made by the two theories for real-time sentence processing.

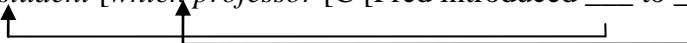## 2.1 Two approaches to interrogative syntax/semantics

The literature on the syntax/semantics of *wh*-questions provides two approaches to the interpretation of in-situ *wh*-phrases: *covert movement* and *in-situ* interpretation. Under the covert movement approach, all *wh*-phrases must be structurally adjacent to the head that interprets them in the CP periphery. This approach thus predicts covert movement of in-situ *wh*-phrases as a prerequisite for their interpretation. The in-situ approach invokes a mechanism that interprets *wh*-phrases without any movement. Hence, if covert movement does occur in a given question, it must be triggered by other factors. Below is a brief description of the two approaches.

2.1.1 The covert movement approach

Under the covert movement approach to questions, *wh*-phrases must be adjacent to C in order to be able to make their contribution to the meaning of the question. Consequently, no *wh*-phrase may remain in situ at LF; instead, all *wh*-phrases occur syntactically next to the complementizer, regardless of where they are pronounced (cf. Cable 2007; 2010, Hagstrom 1998, Hornstein 1995, Huang 1982, Karttunen 1977, Lasnik and Saito 1992, Nishigauchi 1986, Pesetsky 2000, Richards 2001).[1]

(2)   **The covert movement approach to *wh*-in-situ**
       [*Which student* [*which professor* [C [Fred introduced ___ to ___]]]]

The covert movement approach thus predicts pervasive covert movement in multiple *wh*-questions. Moreover, movement is always triggered for one and the same reason – the semantic needs of the *wh*-phrases themselves – and it always targets the same syntactic position at LF: C. This approach makes the prediction in (3).

(3)   **A prediction of the covert movement approach**
       All *wh*-phrases in a question must (overtly or covertly) move to C for interpretation.
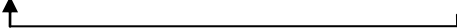
2.1.2 The in-situ approach

Under the in-situ approach to questions, no (overt or covert) movement is required in order to assign interrogative meaning to a structure containing *wh*-elements (cf. Cheng 1991, Chomsky 1995, Hamblin 1973, Kratzer and Shimoyama 2002, É. Kiss 1986, Reinhart 1998, Shimoyama 2006, Tsai 1994). The meaning of a question like (4) 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 (Rooth 1985, 1992). From this perspective, there is no reason to expect any instances of *wh*-movement that are caused by the semantic needs of the *wh*-words themselves. Even the fact that English questions require overt fronting of one *wh*-phrase is unexpected. To explain this

---

[1] Other theories argue that at least in some cases, it is not full *wh*-phrases but rather operators that must move at LF (Aoun and Li 1993, Hagstrom 1998, Kishimoto 2005, Pesetsky 2000, Watanabe 2001, a.o.). Since the target position of movement and the reasons for this movement are the same as in the approach introduced in above, we classify these theories here as consistent with the covert movement approach. However, as we will see, even if these theories are classified as *in-situ* theories, they will be unable to derive the experimental results we present in section 3.

fact, a purely syntactic mechanism must be invoked, unrelated to interrogative semantics, for example an 'EPP' feature requiring C to have a filled specifier (Chomsky 1981).

(4) **The in-situ approach to *wh*-in-situ**
    [*Which student* [C$_{+EPP}$ [Fred introduced ___ to *which professor*]]]

The in-situ approach to questions thus makes no 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 1991; 1993; 1995; 2000, Collins 2001, Fox 2000, Epstein 1992, Kitahara 1997, Reinhart 2006, Richards 2001), it is predicted that *wh*-phrases occupy the position at which they were merged into the syntactic structure. No covert movement occurs for semantic interpretation of *wh*-phrases, (5).

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

2.2 Antecedent Contained Deletion

In this section we discuss Antecedent Contained Deletion (ACD), which will be central to the experimental methodology introduced in section 2.3 and the experiments presented in section 3.

ACD is a phenomenon found in certain VP ellipsis contexts. For VP-ellipsis to be licensed, an *antecedent VP* in the sentence must exist that is identical to the *missing VP*.[2] This is straightforward in examples like (6), where the only VP in the sentence – *read a book* – can serve as an antecedent for the missing VP. In ACD cases like (7), 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.

(6) **VP-ellipsis and its resolution**
    a. John read a book and Mary did ⟨missing VP⟩, too.
    b. John read a book and Mary did ⟨read a book⟩, too.
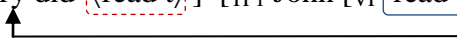
(7) **ACD and the containment problem**
    a. John read every book that Mary did ⟨missing VP⟩ .
    b. John read every book that Mary did ⟨??⟩

To solve the containment problem in (8a), 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 (8b). The resulting VP, containing only the Verb and the trace of covert movement, can then be used

---

[2] 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. Kehler 2002, Arregui et al. 2006, Kertz 2010, Pietro et al. 2012, a.o.). It has also been argued that such mismatches are only possible in VP ellipsis but not in sluicing (Merchant 2013). We will not attempt to contribute to the definition of identity here. We assume that traces count as identical for the purpose of ellipsis parallelism if they bound from parallel positions (e.g. Fox 2002).
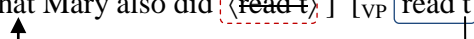
as an antecedent for the missing VP (cf. Fox 2003, Heim and Kratzer 1998, Johnson 1996, Larson and May 1990, May 1985, Sag 1976, Williams 1974; 1977).

(8) **Resolution of ACD using covert movement**

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

    b. [$_{TP2}$ [every book that Mary did ⟨read t⟩] [$_{TP1}$ John [$_{VP}$ read t]]] .

ACD can also occur in sentences with a *wh*-phrase complement of the verb. Example (9a) illustrates this for a sentence with an embedded multiple question, where the in-situ *wh*-phrase hosts an ACD site. The containment problem for this sentence is illustrated in (9b-c). To undo containment and allow for ACD resolution, the in-situ *wh*-phrase must undergo covert movement. Once covert movement has occurred, an appropriate antecedent can be found and used for the missing VP, (9d).

(9)   **ACD in multiple *wh*-questions and its resolution**

    a. John knows [$_{TP1}$ *which student* [$_{VP}$ read [*which book* that Mary also did ⟨missing VP⟩]]].

    b. John knows [$_{TP1}$ *which student* [$_{VP}$ read [*which book* that Mary also did ⟨~~read [*which book* that Mary also did ⟨missing VP⟩~~]]].

    c. John knows [$_{TP1}$ *which student* [$_{VP}$ read [*which book* that Mary also did ⟨~~read [*which book* that Mary also did read [*which book* that Mary also did ⟨missing VP⟩~~]]].

    d. John knows [$_{TP1}$ *which student* [$_{TP2}$ *which book* that Mary also did ⟨read t⟩] [$_{VP}$ read t]]

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

In this section we present the methodology that will be used in the experiments in section 3. The goal of the experiments is to investigate the presence and extent to which in-situ *wh*-phrases undergo covert movement in multiple *wh*-questions. To this end, the experiments will 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. In particular, this 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 (cf. Bever 1970, Frazier and Rayner 1982, Phillips 2003), and (b) structures without covert movement are simpler than structures with covert movement (cf. Anderson 2004, Fox 1995; 2000, Frazier 1999, Tunstall 1998). Given these assumptions, we conclude that the parser does not 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) test the covert movement theory of ACD by comparing the processing of sentences that contain non-quantificational objects (**the**, 10a-c) to sentences with quantificational

objects (**every**, 10d-f) — where this factor is crossed with three different gap sizes inside an attached relative clause: (a) *no ellipsis* using a lexical *verb*, (b) *small ellipsis* marked by *did,* where the antecedent of the ACD site is the embedded VP headed by *treat*, and (c) *large ellipsis* marked by *was*, where the antecedent of the ACD site is the matrix VP headed by *reluctant*.

(10) **The paradigm in Hackl et al. (2012)**
        The doctor was [$_{VP1}$ reluctant to [$_{VP2}$ treat ...

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

                        ... after looking over the test results.

Very specific predictions for language processing in real time are made under the assumptions (a) that ACD requires covert movement for its resolution (cf. section 2.2), and (b) that quantificational objects (but not non-quantificational objects) require covert movement 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 (10a-c), no covert movement is predicted to take place when the definite article is processed. The parser only assumes covert movement after it has been determined that the sentence contains an instance of ACD. This happens after encountering the auxiliaries *did* and *was* in (10b-c), which, together with the immediately following word, signal the presence of an ACD site and thus trigger reanalysis in order to resolve ACD. As discussed above, this reanalysis should incur some processing cost detectable in RTs following the ellipsis site, compared to the baseline with no ellipsis (10a). 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 (10c) 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 (10b) 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, (10c) is expected to produce longer RTs than (10b).

For the *every* conditions, different predictions are made. In all of (10d-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*. Crucially, this position is high enough to preemptively undo antecedent containment in the case of *small ellipsis* (10e) but not in the case of the *large ellipsis* (10f). Hence, ACD resolution should be easier compared to (10b) since part of the work necessary to resolve ACD has already happened prior to encountering the ACD marker *did.* In

the case of the *large ellipsis* in (10f), by contrast, the covert movement step that was 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 (10f) even though the host DP is quantificational in nature.

Hackl et al. (2012) used sentences in a paradigm as in (10) in a self-paced reading study (Just, Carpenter, and Woolley 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. The results presented in Figure 1, showing RRTs two words after the ellipsis site, 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

Figure 1: Hackl et al. (2012)

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 assumption that covert movement is required to accommodate both a quantificational object and an ACD site.
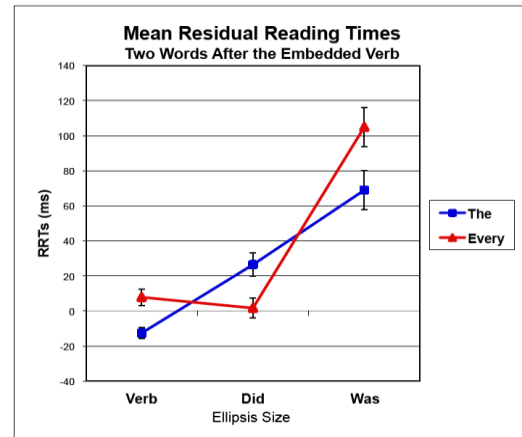
2.4 Predictions of the theories of *wh*-in-situ for the Hackl et al. paradigm

The goal of the present paper is to provide experimental evidence to distinguish between the covert movement approach and the in-situ approach to interrogative syntax/semantics. As we have seen, the two approaches make different predictions regarding the availability of covert movement for in-situ *wh*-phrases in multiple questions, and in particular regarding (a) how much covert movement is predicted to be involved in the derivation of a multiple question, (b) what causes covert movement, and (c) what syntactic position is targeted by covert movement. The predictions of the two approaches are summarized in (11).

(11) **Predictions of the two approaches with regard to covert *wh*-movement**
   a. The covert movement approach:
      All *wh*-phrases in a question must (overtly or covertly) move to C for interpretation.

   b. The in-situ approach:
      Wh-phrases in a question can be interpreted *in situ* and do not require any movement.

Within the Hackl et al. (2012) paradigm, the covert movement approach and the in-situ approach to the interpretation of in-situ *wh*-phrases make different predictions with regard to facilitation of ACD resolution in sentences in which an ACD gap site is hosted inside a *wh*-phrase, e.g. (12a-b):

(12) **ACD hosted by *wh*-phrase in the Hackl et al. paradigm**
        The conductor asked [CP which soloist was [VP1 willing to [VP2 perform…
            a. **which** concerto that the brilliant protégé <u>*did*</u> …   (*small ellipsis*)
            b. **which** concerto that the brilliant protégé <u>*was*</u> …   (*large ellipsis*)
                                                                …and restructured the rehearsal accordingly.

In the covert movement approach, all *wh*-phrases must move *non-locally* to C for interpretation. As a result, we predict that both *small* and *large* ellipsis 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.

In the in-situ approach, *wh*-phrases can be interpreted *in-situ* and hence the parser will not move the in-situ *wh*-phrase to C when it is encountered. Upon reaching the ellipsis site, the parser must reanalyze the structure and covertly move the *wh*-object to a position above the missing VP, 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 ellipsis* condition and in the *large ellipsis* condition. The predictions of the two approaches are summarized in (13), and will be tested in Experiment 1 below.

(13) **Predictions of the two approaches with regard to ACD resolution**
        a. <u>The covert movement approach</u>:
          The resolution of both *small* and *large ellipsis* is facilitated.

        b. <u>The in-situ approach</u>:
          The resolution of both *small* and *large ellipsis* is <u>not</u> facilitated.

3 Distinguishing between the approaches to *wh*-in-situ: Experimental evidence

In the following sections we present the results of three experiments that test the predictions of the covert movement approach and the in-situ approach to *wh*-in-situ. All three experiments use the paradigm in Hackl et al. (2012) as a tool for detecting covert movement in multiple *wh*-questions.

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 et al. to appear, Clifton et al. 2006, Fanselow et al. 2008, Featherstone 2005a,b, Frazier and Clifton 2002, Hofmeister et al. 2007), and furthermore that *superiority-obeying* questions are easier to process than *superiority-violating* questions. 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 in a superiority-obeying word-order.

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

This experiment tests the core predictions of the covert movement approach and in-situ approach to in-situ *wh*-phrases. In particular, we compare the behavior of questions with an ACD gap hosted by a *wh*-phrase with the behavior of questions with an ACD gap hosted by the quantificational determiner *every*.

3.1.1 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 <u>*did*</u>, where the antecedent of the ACD site is the embedded VP, VP_2, or *large ellipsis* marked by <u>*was*</u>, where the antecedent of the ACD site is the matrix VP, VP_1. A sample item is given in (14) below:[3]

(14) **Sample target item in Experiment 1**
 The conductor asked [$_{CP}$ which soloist was [$_{VP1}$ willing to [$_{VP2}$ perform…

  a. **every** concerto that the brilliant protégé <u>*did*</u> …   (*small ellipsis*)
  b. **which** concerto that the brilliant protégé <u>*did*</u> …   (*small ellipsis*)
  c. **every** concerto that the brilliant protégé <u>*was*</u> …   (*large ellipsis*)
  d. **which** concerto that the brilliant protégé <u>*was*</u> …   (*large ellipsis*)
                                      …and restructured the rehearsal accordingly.

There were 28 sentence templates following the sample paradigm in (14). 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 <u>*did*</u> marked a *small ellipsis* corresponding to the embedded predicate, VP_2, and the auxiliary <u>*was*</u> 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.

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

---

[3] To simplify the discussion, in what follows we ignore the highest predicate embedding the question. We 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."

(that is, the first line of the sentence was the first line in (14), and the second line contained the text in lines a-d and the continuation following these lines).[4]

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*). 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. 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*.

3.1.2 Methods

Experiment 1 used the moving window self-paced reading methodology and was hosted on Ibex Farm.[5] Participants were presented with sentences that appeared on the screen one word at a time in a moving window display. Each trial begins with a series of dashes marking the length of the sentence. Participants press the spacebar to reveal the next word of the sentence. Each press of the spacebar reveals a new word while the previous word is again replaced with dashes. The amount of time a participant spends reading each word is recorded (RT). After the final word of each sentence, a yes/no comprehension question appears, asking about information contained in the sentence. 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. Each experiment took approximately 30 minutes to complete.

Subjects for the experiments were recruited through Amazon Mechanical Turk and they were paid $1.5 for their participation. The 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

---

[4] 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 options in a pilot study.
[5] Ibex: Internet Based Experiments, created and maintained by Alex Drummond, accessible at http://spellout.net/ibexfarm/.

to the US using Amazon Mechanical Turk's user interface. Participation was further restricted so that only Turk Workers with a overall approval rate of over 95% of all their submissions were allowed to participate in the experiments.

3.1.3 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 (14)) as soon as the quantifier is encountered, we expect antecedent containment to be preemptively undone in the case of the *small ellipsis* (marked with <u>*did*</u>) (14a), leading to facilitation of ACD resolution in this case. However, since this movement does not target a position high enough to undo antecedent containment in the case of the *large ellipsis* (marked with <u>*was*</u>) (14c), we expect ACD resolution to be relatively more difficult in this case: once the auxiliary <u>*was*</u> 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 (14)), 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 (14b,d).

For the *which* conditions, the two approaches to in-situ *wh*-phrases make different predictions:

In the 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. When the parser encounters the ellipsis site—regardless of whether the ellipsis is *small* or *large*—all that is left to do is find an antecedent for the missing VP. We thus expect an interaction, such that the *which-did* and *which-was* conditions both pattern with the *every-did* condition and exhibit facilitation effects, compared to the remaining *every-was* condition where we predict participants to exhibit increased difficulties with ACD resolution.

According to the second approach, the in-situ *wh*-phrase can be interpreted *in-situ.* Hence, encountering the in-situ wh-phrase will not trigger reanalysis that could facilitate down-stream ACD resolution. Only upon reaching the ellipsis site itself 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. Under this approach we thus expect to find a main effect of the object type, such that sentences with a relative clause headed by *which* are more difficult to process than sentences with a relative clause headed by *every*. This main effect may be accompanied by an interaction, such that the two *which* conditions pattern with *every-was* and are <u>more difficult</u> to process than *every-did*, or they may be even more difficult than the *every-was* condition.

(15) **Predictions of the two approaches with regard to ACD resolution in Experiment 1**

    a. The covert movement approach:

      The resolution of both *small* and *large ellipsis* is facilitated.

      Which-did and *which-was* pattern with *every-did* and are <u>easier</u> to process than *every-was*.

    b. The in-situ approach:

      The resolution of both *small* and *large ellipsis* is <u>not</u> facilitated.

      Which-did and *which-was* are <u>more difficult</u> to process than *every-did*; they are at least as difficult to process as *every-was*.

3.1.4 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,[6] 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. In addition, two target sentences and no filler sentences were excluded from the analysis because of low accuracy (<60% across participants).

Questions across the full experiment (targets and fillers) were answered correctly 87.5% of the time across participants; questions for experimental items were answered correctly on 83.3% of trials. 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. Figure 2 shows the mean residual reading times (RRTs)[7] for the two regions of interest for the four target conditions.

---

[6] We note that 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.

[7] 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.
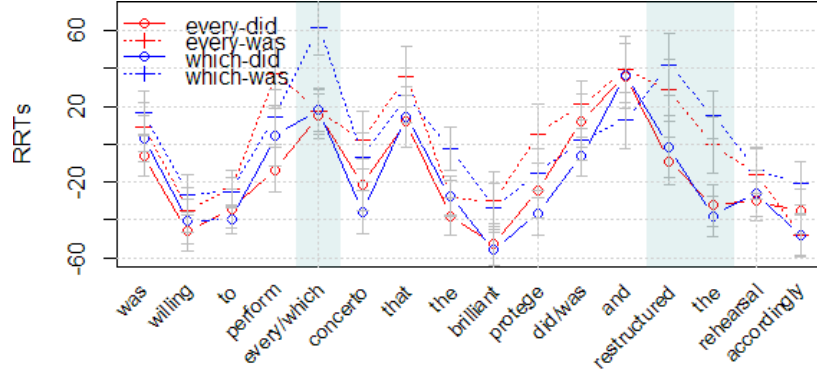
Figure 2: Residual reading times in target items in Experiment 1

A linear mixed effects model (Baayen 2004, Barr et al. 2012, Bates and Sarkar 2006) with random intercepts for subjects and items was fit to the data using R and the R package lme4 (Bates and Maechler 2009). The model predicted *logRT* from the two factors of interest: *determiner* (every vs. which) and *ellipsis size* (small ellipsis marked by *did*, vs. large ellipsis marked by *was*).[8] Results show a main effect of *determiner* at the slot at which the determiner appeared in the sentence (log likelihood tests, p<0.05). This result is driven by the fact that reaction times in the *which* condition were slower than the reaction times in the *every* condition, across both *ellipsis* conditions.[9] The results additionally show a main effect of *ellipsis size* two words and three words after the auxiliary site (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*. There were no differences between the two determiners at these slots, and there were no other significant effects in the results.

3.1.5 Discussion

We find two effects 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. When the parser encounters the *in-situ wh*-phrase in the *which* conditions, it realizes that the sentence will be paired with a more complex semantics than in the case of the simplex question with *every*. Regardless of the cause of this effect, it shows that our participants were processing the sentences at least at a depth sufficient for detecting the difference in determiner.

Second, we find 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 determiners: *every* and *which*. That is, we see the same effect of our *ellipsis size* manipulation affecting the two determiners equally. This result is not predicted by either traditional approach to *wh*-in-situ. Recall that the covert movement approach

---

[8] LogRTs in all the experiments in section 3 were manually inspected to ensure that they exhibited a normal distribution. Raw RTs as well as RRTs were not normally distributed and hence not used with the linear models.

[9] 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).

predicts facilitation of both *small* and *large* ellipsis, because the in-situ *wh*-phrase must move to C for interpretation, preemptively undoing antecedent containment in both ellipsis conditions. We would thus expect that RTs for both *which* conditions, irrespective of whether they involve local or non-local ACD, would be lower than those for the *every* conditions. The in-situ approach, on the other hand, assumes that the *wh*-phrase can be interpreted without any movement at all, and hence predicts <u>no</u> facilitation by an upstream *which* for either the *small* or *large* ellipsis conditions. However, the results of Experiment 1 show an effect of ellipsis size but no effect of determiner, which is not explained under both approaches to *wh*-in-situ.

Based on the results of Hackl et al. (2012), we may take the fact that there is no difference between *every-did* and *which-did* to indicate that local ACD resolution is facilitated in both cases. More specifically, if *which* where a determiner that could be interpreted in situ like *the* we should have seen relatively longer RTs for *which-did,* just like Hackl et al. (2012) did for *every-did.* The fact that we didn't shows that *wh*-phrases behave like traditional quantificational determiners such as *every* with regard to ACD resolution in real time.

This is compatible with a view under which *wh*-phrases, too, are quantifiers that must covertly move (via QR) to the nearest propositional node for interpretation.[10] Similar proposals have been previously made by several researchers for a variety of languages (cf. Baker 1970, Dornish 1998 for Polish, Huang 1995, Kim 1991 for Korean, Rullmann and Beck 1998). We propose that once the *wh*-phrase has been integrated into the structure, it does not require any further movement: it may be interpreted *in-situ* at the first position in which it can be interpreted in the structure. This proposal can be viewed as a modified version of the in-situ approach: the movement of *wh*-phrases is always forced by external considerations such as the interpretability of the *wh*-phrase, but it is not the case that the *wh* can always be interpreted in the position where it occurs in the overt syntax: some movement step is required whenever a *wh*-phrase occurs in object position. We propose that the movement step assumed by the parser is the smallest that can produce an interpretable structure. The approach we propose is summarized in (16)-(17):

(16) **The status of *wh*-phrases**
   Wh-phrases are existential quantifiers.

(17) **The *partial movement* approach**
   A *wh*-phrase must be interpreted at a position with propositional type at LF.

To substantiate this approach and its predictions, below we present Experiments 2-3. Experiment 3 will test a prediction of the *partial movement* approach that if the partial movement step is forced to be long-distance via external interpretability considerations, we expect to detect an increased domain in which ACD resolution is facilitated. As we will see, this prediction is borne out. We will return to a detailed discussion of this prediction in section 3.3, but first we present Experiment 2, whose aim is to deal with an alternative interpretation of the results of Experiment 1. One possible alternative explanation for the results of Experiment 1 is that the experimental

---

[10] See Heim and Kratzer (1998) for a discussion of the type-mismatch problem with quantifiers in object position.

manipulation did not in fact succeed as intended. In particular, we must entertain and reject the possibility that participants failed to process the sentences at a deep level, and that instead of seeing any facilitation effects in our sentences we are only seeing a *complexity of antecedent effect*.

Long distance dependencies, such as those required for the resolution of ellipsis, are constructed by integrating temporally and structurally distant linguistic material. As such, they require the support of working memory resources for their completion (Foraker and McElree 2011). In order to establish a dependency, the parser must retrieve previously processed material from working memory at the gap position, and find an appropriate antecedent for the ellipsis. The distance between the gap site and the antecedent affects the retrieval process, resulting in longer reading times and decreased acceptability for longer dependencies (cf. Gibson 1998, van Dyke and Lewis 2003, Lewis et al. 2006, Warren and Gibson 2002). Relatedly, we might expect semantically more complex antecedents (e.g. *willing to perform*) to be harder to retrieve from memory than simpler antecedents (e.g. *perform*).

One possible interpretation of Experiment 1, then, is that the Hackl et al. (2012) paradigm was not successfully replicated. Instead, participants were not reading the sentences carefully and were not affected by the *determiner* manipulation at all. A similar concern is that the experimental paradigm is not sensitive enough to detect differences between the processing of different quantifiers, perhaps because an embedding inside a question presents too much of a challenge to the participants. Hence, Experiment 1 only exhibits an effect of complexity of the antecedent, where the processing of long-distance ellipsis with a more complex antecedent is more costly than that of short-distance ellipsis with a simpler antecedent, and nothing more. To address this concern, first note that some comprehension questions directly targeted the relative clause, to ensure that it was processed at a deep level. Those questions did not suffer from lower accuracy rates than other comprehension questions. We furthermore detect a main effect of *determiner* once *which* and *every* are read, ensuring that participants did indeed pay attention to the determiners in the sentences.

However, to more directly address the concern that the Hackl et al. (2012) paradigm might not have worked as expected, Experiment 2a-b will attempt to replicate the Hackl et al. results—comparing the behavior of *every* and *the*—in the context of an embedded question, and then additionally compare the behavior of *which* and *the*. As we will see below, we find that ACD resolution with the quantificational determiner *every* is facilitated compared to ACD resolution in sentences with the non-quantificational definite article *the*, as predicted by the results of Hackl et al. (2012). We furthermore find that ACD resolution in sentences with *which* is facilitated compared to ACD resolution in sentences with the definite article *the*. This result will eliminate the concern that the results of Experiment 1 do not show any facilitation effects with *which*.

3.2. Experiment 2a-b: *the* **vs.** *every* **and** *the* **vs.** *which* **in an embedded question context**

The goal of Experiment 2 is to ensure that the paradigm used in Experiment 1 is in principle sensitive enough to detect differences in the behavior of different determiners. Experiment 2a

will attempt to replicate the Hackl et al. (2012) paradigm in a context similar to that of Experiment 1, using an embedded question instead of a simple declarative sentence. Experiment 2b will then compare the behavior of multiple questions with *which* with the behavior of sentences with the non-quantificational definite article *the*. To allow for a direct comparison with Experiment 1, Experiment 2 uses the same materials as in Experiment 1, with minor changes to accommodate the experimental manipulation of this experiment.

3.2.1 Design

Like Experiment 1, Experiment 2 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*, the wh-word *which*, or the definite article; and (b) *ellipsis size*: whether the sentence contained a *small ellipsis* marked by <u>*did*</u>, where the antecedent of the ACD site is the embedded VP_2, or *large ellipsis* marked by <u>*was*</u>, where the antecedent of the ACD site is the matrix VP_1. 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*, has already been done in Experiment 1. Sample items are given in (18)-(19):

(18) **Sample target item in Experiment 2a**
The conductor asked [$_{CP}$ which soloist was [$_{VP1}$ willing to [$_{VP2}$ perform…
    a. **every** concerto that the brilliant protégé <u>*did*</u> …   (*small ellipsis*)
    b. **the** concerto that the brilliant protégé <u>*did*</u> …   (*small ellipsis*)
    c. **every** concerto that the brilliant protégé <u>*was*</u> …   (*large ellipsis*)
    d. **the** concerto that the brilliant protégé <u>*was*</u> …   (*large ellipsis*)
                …and restructured the rehearsal accordingly.

(19) **Sample target item in Experiment 2b**
The conductor asked [$_{CP}$ which soloist was [$_{VP1}$ willing to [$_{VP2}$ perform…
    a. **which** concerto that the brilliant protégé <u>*did*</u> …   (*small ellipsis*)
    b. **the** concerto that the brilliant protégé <u>*did*</u> …   (*small ellipsis*)
    c. **which** concerto that the brilliant protégé <u>*was*</u> …   (*large ellipsis*)
    d. **the** concerto that the brilliant protégé <u>*was*</u> …   (*large ellipsis*)
                …and restructured the rehearsal accordingly.

The same 28 target sentences from Experiment 1 were used, along with the same 48 filler items. The only change introduced to the sentences was to the determiners, as indicated above. The comprehension questions to some items were also minimally changed to accommodate this manipulation. Experiment 2 used the same methods as described above for Experiment 1. See the Appendix for a full list of the materials.

3.2.2 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*.

In the *every* conditions (18a,c), the parser must assume covert movement as soon as it encounters *every*. This movement is expected to be *local* and to target a position above the embedded VP_2 headed by *perform* in (18), a position that is high enough to preemptively undo antecedent containment in the case of *small ellipsis* (18a) but not in the case of the *large ellipsis* (18c). Hence, the processing costs at the ellipsis site in the *small ellipsis* condition (19a) should be lessened since part of the work to resolve ACD, specifically reanalysis to undo antecedent containment, has already happened prior to encountering the ACD site. When the parser encounters <u>*did*</u> all it needs to do is find an antecedent for the missing VP and nothing more. In the case of the *large ellipsis* in (18c), the covert movement that was assumed following the processing of *every* is not sufficient: after local movement of the object DP the missing VP is still contained inside its antecedent. Hence, when the auxiliary <u>*was*</u> is reached, the parser must again reanalyze the structure, covertly moving the object a second time, from its position above the embedded VP_2 to a position higher than the matrix VP_1. Hence, a high processing cost should be associated with the resolution of the *large ellipsis* in (18c).

In the *definite* conditions in (18b,d), no covert movement is predicted to take place when the definite article is encountered. The parser only assumes a structure with covert movement of the object when it encounters the auxiliaries <u>*did*</u> and <u>*was*</u>, as the ACD marker is the earliest point at which the need for moving the object DP is detectable. The fact that processing the ACD site requires both reanalysis and retrieval of an antecedent is expected to result in relatively high processing costs following the ellipsis site. Moreover, just like in the case of *every*, we expect that the processing costs for the *large* ellipsis condition should be higher than those associated with the *small* ellipsis, since retrieving a more complex antecedent VP is more difficult and requires longer distance covert movement than in the case of the *small* ellipsis.

We furthermore expect to find the same behavior pattern with *which* in Experiment 2b, (19): if *which*, like *every*, is a quantifier in object position that must QR to the first propositional node in the structure for interpretation, we expect to find that *which*, like *every*, facilitates the resolution of *small ellipsis* with *did*, but not of *large ellipsis* with *was*, 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*.

3.2.3 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 sentence 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 across participants 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 3 shows the mean residual reading times for the region of interest for the four target conditions in Experiment 2a, comparing the processing of sentences with *the* and *every*.
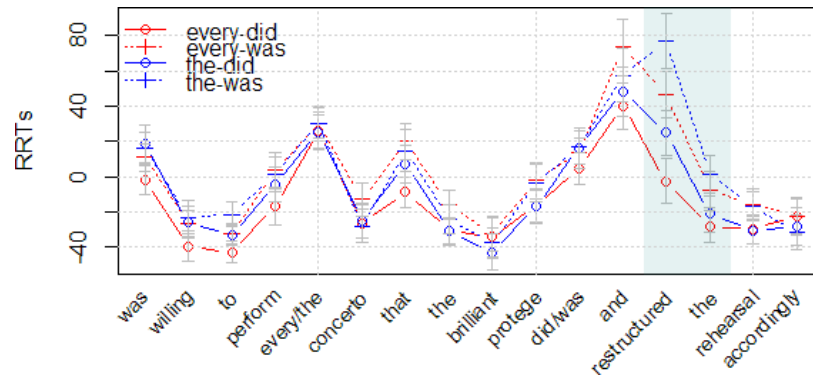

Figure 3: Residual reading times in target items in Experiment 2a

A linear mixed effects model with random intercepts for subjects and items was fit to the data. The model predicted *logRT* from the two factors of interest: *determiner* (every vs. the) and *ellipsis size* (small ellipsis marked by _did_, vs. large ellipsis marked by _was_). Results show a main effect of *ellipsis size* two and three words after the auxiliary verb site and a main effect of *determiner* two words after the auxiliary verb site (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 *every* and *the*, and furthermore that the resolution of ACD in sentences with a relative clause headed by *every* is faster than the resolution of ACD in sentences with a relative clause headed by *the*. There were no other significant effects in the data.

Next, we examine the results of Experiment 2b, comparing the determiners *the* and *which*. Figure 4 shows the mean RRTs for the region of interest for the four target conditions.
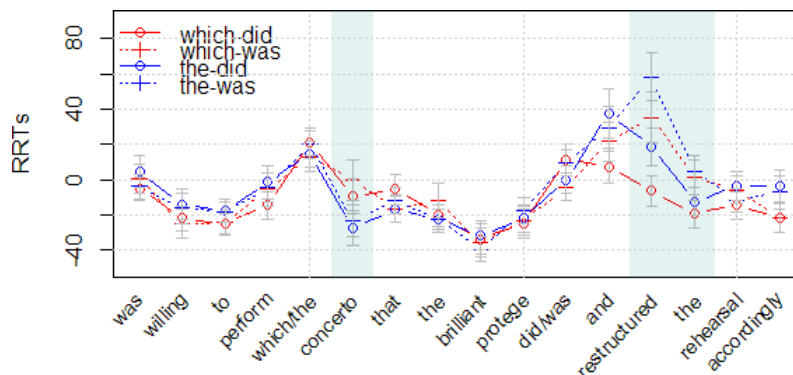

Figure 4: Residual reading times in target items in Experiment 2b

A linear mixed effects model with random intercepts for subjects and items was fit to the data. The model predicted *logRT* from the two factors of interest: *determiner* (which vs. the) and *ellipsis size* (small ellipsis marked by *did*, vs. large ellipsis marked by *was*). Results show a main effect of *ellipsis size* two and three words after the auxiliary verb site and a main effect of *determiner* two words after the auxiliary verb site (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 *which* and *the*, and furthermore that the resolution of ACD in sentences with a relative clause headed by *which* is faster than the resolution of ACD in sentences with a relative clause headed by *the*. There were no other significant effects in the data.

3.2.4 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* also a main effect of *determiner*, such that sentences with *every* are processed faster than sentences with *the*, and sentences with *which* are also 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 5, which compares the reading times of *the*, *every* and *which* two words after the auxiliary verb site in Experiments 2a-b.
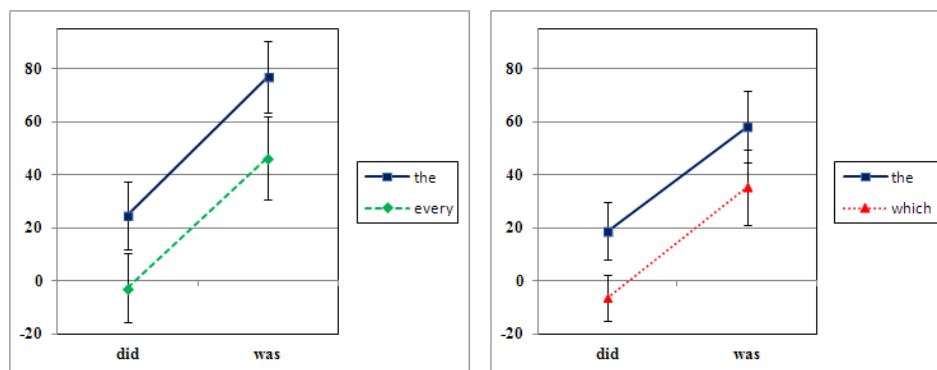

Figure 5: Residual reading times two word after the auxiliary verb in Experiments 2a-b

For the results of Experiment 2a, we adopt Hackl et al.'s explanation and assume that this is so 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*.[11] As a result, covert movement must always be assumed to take place at the point of ACD resolution in the case of *the* but not in the case of *every*: when a small ellipsis (marked with *did*) is encountered, the parser must assume no additional covert movement with *every* since that has happened earlier in the parse. To explain the fact that non-local ACD resolution is less difficult with *every* as well we hypothesize

---

[11] However, an important difference between our 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 an important role in Hackl et al.'s argument that only local QR facilitates ACD resolution, which we are unable to replicate here. However, the results of Experiment 3 below will provide us with the missing evidence to make this claim in this paper here as well, and we will return to this point after introducing that experiment.

that the second step of QR to a position above the matrix VP_1 is easier in the case of *every*, perhaps because the object needs to be moved less far than in the case of *the*.

The results of Experiment 2b are similarly explained if we assume that *wh*-phrases are quantifiers that must QR to the nearest propositional node for interpretation as soon as they are encountered by the parser, but that no such movement is assumed in the case of the definite article *the*. As a result, covert movement must always be assumed to take place at the point of ACD resolution in the case of *the* but not in the case of *every*: when a small ellipsis (marked with *did*) is encountered, the parser must assume no additional covert movement with *every* since that has happened earlier in the parse. As in the case of *every* in Experiment 2a, we find that non-local ACD resolution is less difficult with *which* than with *the*. This is explained in the same way as for *every*: we hypothesize that the second step of QR to a position above the matrix VP_1 is easier in the case of *which* because the object needs to be moved less far than in the case of *the*.

The results of Experiment 2 thus confirm that Experiment 1 is sensitive enough to detect potential differences between determiners with regard to their interaction with ACD, and that there is no reason to suspect that participants were not processing the experimental materials with sufficient depth. Here again, we find similar accuracy rates on target and filler sentences as in Experiment 1, and the overall reaction times in Experiment 2 are comparable to those found in Experiment 1 (see Figure 2).

The facilitation effect found with *every* compared to the definite article *the* and the fact that parallel behavior is observed with *which* support the interpretation we have given to the results of Experiment 1. Specifically, they show that local ACD resolution with *which* is facilitated just as it is for *every*. The facilitation effect with *which* is not compatible with the *wh*-phrase having been processed *in-situ* when it was encountered by the parser, as we would then predict no effect of *determiner* in Experiment 2b. Instead, the results are consistent with *which* being QRed *locally* to a position above the embedded VP_2, just like traditional quantificational determiners. Under this approach, *which* is expected to facilitate ACD resolution to a similar extent as *every*, and this is indeed what we see in Experiment 2b. This leads us to adopt a new approach to the interpretation of *wh*-in-situ, repeated in (20).

(20) **The *partial movement* approach**
    A *wh*-phrase must be interpreted at a position with propositional type at LF.

There is, however, an alternative interpretation of the results of Experiments 1-2: it is possible that Experiments 1-2 do not show lack of sufficient covert movement with *every* and *which* in the *large ellipsis* (*was*) condition, but instead only an effect of the *complexity of the antecedent*. If the difference between *the* on the one hand and *every* and *which* on the other hand is not the result of a difference in QR assumed for the interpretation of these determiners, but instead is contributed by some other source,[12] then we can no longer infer from the difference between the

---

[12] For example, if – for whatever reasons – the processing of ACD hosted by *every* and *which* is easier than the processing of ACD hosted by *the*.

behavior of *every* (and *which*) and *the* that upstream covert movement has occurred with *every* (and *which*) but not with *the*. The main effect of *ellipsis size* could be explained as an effect of the complexity of the antecedent: integrating a smaller, simpler antecedent into the structure is easier than integrating a larger, more complex antecedent. The main effect of *determiner* would be a consequence of the as yet to be identified property of *the* that makes ACD resolution with it difficult.

Experiment 3 will address this possible alternative interpretation of Experiments 1-2 by considering a prediction of the *partial movement* approach: that although *wh*-phrases need not move any further once they have been integrated into the structure—that is, following an initial QR step for *wh*-phrases in object position—they may be forced to move higher on independent grounds. This prediction allows us to distinguish between the behavior of *every* and the behavior of *which* in certain environments, where long-distance *wh*-movement may be necessary, leading us to expect a larger region in the question in which ACD facilitation effects are expected with *which* but not with *every*. If the no-movement approach sketched above is on the right track— that is, if all Experiments 1-2 are showing is a complexity effect, then we should not see an increased domain of ACD facilitation effects, because no long-distance covert movement is assumed when the parser reaches *which*. On the other hand, if the *partial movement* approach is on the right track, then additional covert movement must occur in the question and we expect to find facilitation effects of ACD resolution for any ellipsis that is smaller than the landing site of this extended covert movement step. Below we test this prediction and show that it is indeed borne out.

(21) **Predictions of the *partial movement* approach**
   Wh-phrases are quantifiers that QR to the nearest propositional node for interpretation. No additional movement is required for the interpretation of a *wh*-phrase, but movement may be forced by external interpretability considerations.
   Wh-movement may target other positions beside C.

3.3. Experiments 3a-b: *every* vs. *which* with interveners

As noted above, the *partial movement* approach to *wh*-in-situ predicts that, although *wh*-phrases in object position can be interpreted in-situ following an initial QR step, other factors may force further *wh*-movement. If long-distance movement can be forced, we expect to find facilitation effects of ACD resolution in a larger portion of the sentence, up to the position targeted by covert movement. One phenomenon that has been argued to force covert movement of in-situ *wh*-phrases is *intervention effects*. Importantly, this phenomenon has not been argued to affect quantificational determiners like *every*, and so intervention effects allow us to make diverging predictions about the behavior of *every* and *which* and to experimentally test these predictions.

Below we give a brief introduction to the theory of intervention effects that will be assumed here. We show how intervention effects can be used in the experimental paradigm to test the predictions of the *partial movement* approach, and then present Experiments 3a-b, which use
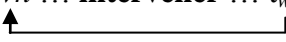
21

intervention effects in the paradigm of Experiment 1 and corroborate the predictions of the *partial movement* approach to *wh*-in-situ.

3.3.1 The theory of intervention effects

The term *intervention effect* describes a situation in which a question is rendered ungrammatical because an in-situ *wh*-phrase is c-commanded at LF by an *intervener,* for instance a focus sensitive operator such as *only* or negation. Cross-linguistically, intervention effects have been found in *wh*-fronting languages as well as in *wh*-in-situ languages, and several competing theories have been proposed to explain the phenomenon (cf. Beck 1996; 2006, Beck and Kim 1997, Cable 2010, Kim 1997; 2002, Mayr 2013, Pesetsky 2000, Tomioka 2007a,b, see also Hagstrom 1998, Hoji 1985, Soh 2005).

In this paper we adopt Beck's (2006) approach to intervention effects, who following Pesetsky (2000), proposes that these effects are caused whenever a *wh*-phrase cannot covertly move above a potentially offending intervener and is instead forced to be interpreted in-situ using focus-alternatives. The intervener is argued to disrupt the transmission of alternatives from the *wh* to C, resulting in ungrammaticality, (22b). If, on the other hand, the *wh* is able to covertly move to a position above the intervener, the question can be assigned a convergent semantics, (22a). This state of affairs is summarized in the schema in (22), where covert movement is marked with an arrow and areas where alternatives are projected are marked with a squiggly arrow.[13]

(22) **The configuration of an intervention effect (Beck 2006)**
   a. ✓ [$_{CP}$ C … *wh* … **intervener** … t$_{wh}$ …]

   b. * [$_{CP}$ C … **intervener** … *wh* …]

In English, superiority-obeying questions are immune to intervention effects, but superiority-violating questions are ungrammatical whenever an intervener occurs above an in-situ *wh*-phrase (É. Kiss 1986, Pesetsky 2000). This is illustrated in (23a-b), which employ the focus-sensitive operator *only* as intervener.

(23) **Intervention effects only target superiority-violating questions**
   a. *Which student* did **only** Fred introduce ___ to *which professor*?   Superiority-obeying
   b. **Which professor* did **only** Fred introduce *which student* to ___?   Superiority-violating

Pesetsky (2000) and Beck (2006) argue that in superiority-obeying questions, the in-situ *wh* is able to covertly move to C at LF. As a result, it is above the offending intervener and hence the question is grammatical, (23a). In superiority-violating questions, on the other hand, the in-situ *wh* cannot move to C and must stay in-situ at LF and be interpreted via focus-alternatives. As a result, if an intervener occurs above the in-situ *wh*-phrase, the result is ungrammaticality, (23b).

---

[13] For Beck (2006), covert movement necessarily targets the interrogative complementizer. That is, Beck adopts the traditional all-or-nothing view of covert movement. However, in principle all that is necessary for the *wh*-phrase to be interpretable in Beck's theory is for it to move *above the intervener*. This is compatible with a smaller movement step under the partial movement approach, as schematized in (22a).

The grammaticality of example (23a) is explained because the in-situ *wh*-phrase *which professor* is able to covertly move above the intervener at LF. Note that this movement of *which professor* targets a higher position than what is expected if no intervener is present. This is shown in (24)-(25) below, where a question with and without an intervener are contrasted. We can see that longer movement has occurred in the question with an intervener (25) than in the question that lacks an intervener, (24): In example (24), *which professor* has covertly moved above the predicate *introduce* to satisfy the interpretability requirement which forces it to QR to a propositional node for interpretation. In example (25), *which professor* undergoes covert movement to C in order to avoid the illicit intervention effect configuration in (22b). If *which professor* were not covertly raised above the intervener, the result would be the ungrammatical (25b) (which is identical to (24) except for the presence of the intervener). As before, a dashed line represents the area in which focus-alternatives are computed.

(24) **LF of multiple question under the *smallest movement* approach**
Which student did Fred introduce ___ to *which professor*?
[[*which student*]$_1$ [C [$_{TP}$ … t$_1$ [*which professor*]$_2$ [$_{VP}$ introduce … t$_2$]]]]

(25) **LF of multiple question with intervener under the *smallest movement* approach**
Which student did Fred introduce ___ to *which professor*?
a.   [[*which student*]$_1$ [*which professor*]$_2$ [C [$_{TP}$ … **intervener** t$_1$ t$_2$ [$_{VP}$ introduce … t$_2$]]]]
b. * [[*which student*]$_1$ [C [$_{TP}$ … **intervener** t$_1$ [*which professor*]$_2$ [$_{VP}$ introduce … t$_2$]]]]

We thus predict that the presence of an intervener in a question can force additional covert movement of *wh*-phrases that could otherwise be interpreted without movement following their QR step. The movement must be at least as high as the intervener, but need not target C. Hence, we expect the *presence* and *placement* of an intervener in a question to affect ACD resolution in that question. The intervener marks the scope of the necessary covert movement in the question: the *wh*-phrase must move at least as high as the intervener in order to escape intervention. We thus expect to find facilitation of ACD resolution if the ellipsis is smaller than the movement required by the intervener. On the other hand, no additional movement step is predicted to occur in the case of the quantificational determiner *every*, because interveners do not interact with traditional quantifiers. Consequently, no additional facilitation of ACD resolution is expected when an intervener occurs in the context of *every*. Thus, we predict no effect of the presence and placement of an intervener in sentences with ACD hosted by *every*. Testing this prediction will be the goal of Experiments 3a-b.

Experiment 3a-b will use the focus-sensitive intervener *also*. Although previous work on English has not used this intervener, it is known to be focus-sensitive (Beaver and Clark 2008, a.o.) and has been argued to act as an intervener in other languages (Beck 2006, Kim 2002, a.o.). We argue that it acts as an intervener in English as well, exhibiting the same grammaticality paradigm as known interveners such as *only* and *even*, where a superiority-obeying question with

also above the in-situ *wh*-phrase is grammatical, but a superiority-violating question with *also* above the in-situ *wh*-phrase is ungrammatical (cf. example (23)).

(26) ***Also* is an intervener in English**
   I know that the teacher gave some punishments to some students on Thursday. Tell me,
   a.   Which punishment did the teacher **also** give ___ to which student on Friday?
   b. *Which student did the teacher **also** give which punishment to ___ on Friday?

3.3.2 Design

Experiment 3 is an expansion of Experiment 1, adding interveners to the questions. Three factors will be 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; (b) *ellipsis size*: whether the sentence contained a *small ellipsis* marked by <u>did</u>, where the antecedent of the ACD site is the embedded VP_2, or *large ellipsis* marked by <u>was</u>, where the antecedent of the ACD site is the matrix VP_1; and (c) *position of also*: whether the intervener *also* is *low* and occurs above the embedded VP_2, or *high* and occurs above the matrix VP_1.

To simplify the design and the analysis of the results, Experiment 3 will be divided into two sub-experiments, each using just one determiner, *every* or *which*. Sample items for Experiments 3a-b are given in (27)-(28). See the Appendix for a full list of the materials.

(27) **Sample target item in Experiment 3a (*every*)**[14]
   The conductor asked [$_{CP}$ which soloist was …
   a. [$_{VP1}$ **also** willing to [$_{VP2}$ perform every concerto that the brilliant protégé <u>did</u>…   (*high also*)
   b. [$_{VP1}$ willing to [$_{VP2}$ **also** perform every concerto that the brilliant protégé <u>did</u>…   (*low also*)
   c. [$_{VP1}$ **also** willing to [$_{VP2}$ perform every concerto that the brilliant protégé <u>was</u>…   (*high also*)
   d. [$_{VP1}$ willing to [$_{VP2}$ **also** perform every concerto that the brilliant protégé <u>was</u>…   (*low also*)
                              …and restructured the rehearsal accordingly.

(28) **Sample target item in Experiment 3b (*which*)**
   The conductor asked [$_{CP}$ which soloist was …
   a. [$_{VP1}$ **also** willing to [$_{VP2}$ perform which concerto that the brilliant protégé <u>did</u>…   (*high also*)
   b. [$_{VP1}$ willing to [$_{VP2}$ **also** perform which concerto that the brilliant protégé <u>did</u>…   (*low also*)
   c. [$_{VP1}$ **also** willing to [$_{VP2}$ perform which concerto that the brilliant protégé <u>was</u>…   (*high also*)
   d. [$_{VP1}$ willing to [$_{VP2}$ **also** perform which concerto that the brilliant protégé <u>was</u>…   (*low also*)
                              …and restructured the rehearsal accordingly.

Experiment 3 contained the same 28 target sentences from Experiment 1, along with the same 48 filler items. The items were minimally changed to add the intervener *also* in the appropriate places. No changes were made to the comprehension questions from Experiment 1.

---

[14] We note that there are several choices for what element in (27)-(28) 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 us are possible choices are *concerto* and the verb *perform*. The actual choice made by readers is immaterial to our 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*.

3.3.3 Predictions

As sketched above, we expect interveners to interact with *wh*-phrases and force movement of the in-situ *wh*-phrase to a position above the intervener. Following our assumptions about the parser, we expect it to perform the smallest movement possible, that is to target the first interpretable position above the intervener and no higher. Hence, in the *low also* condition, we expect *wh*-movement to target a position above *also* but below VP_1, while in the *high also* condition we expect *long distance wh*-movement above *also* and thus above VP_1. We do not expect interveners to interact with *every*, as conventional quantifiers do not interact with focus interveners, and hence we expect *every* to undergo a small QR step and not to move any further until the need for long-distance movement is apparent after reaching the ACD site marked by *was*.

For Experiment 3a we expect 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 interaction with the placement of *also*. The parser must assume covert movement as soon as it encounters *every* to a *local* position just above the embedded VP_2. This position is high enough to preemptively undo antecedent containment in the case of *small ellipsis* (27a-b) but not in the case of the *large ellipsis* (27c-d). Hence, no particular processing cost should be detected at the ellipsis site in the *small ellipsis* conditions, because no reanalysis is required at the gap site in order to resolve ACD in this sentence. In the case of the *large ellipsis*, the covert movement that was assumed following the processing of *every* is not sufficient: 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_2 to a position higher than the matrix VP_1. Hence, a high processing cost should be associated with the resolution of the *large ellipsis* in (27c-d). Since the presence and position of *also* should not affect the movement of *every*, we predict no differences between the *low also* conditions (27b,d) and *high also* conditions (27a,c).

Different predictions are made in the case of Experiment 3b. Here, we expect the position of *also* to affect the amount of covert movement that the in-situ *wh*-phrase must undergo irrespective of the size of ACD.

In the *low also* conditions (28b,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 *local* movement, targeting 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* (28b) but not in the case of the *large ellipsis* (28d). Consequently, we expect ACD resolution to be facilitated in the case of *small ellipsis* but not in the case of the *large ellipsis*, resulting in larger RRTs for the *was* conditions compared to the *did* condition.

In the *high also* conditions (28a,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, in the *high also* conditions 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 <u>both</u> in the case of *small ellipsis* (28a) and *large ellipsis* (28c). As a result, when the parser reaches the gap site—in both ellipsis conditions—no reanalysis is necessary in order to construct an appropriate antecedent for the ellipsis. All the parser must do is find an antecedent for the ellipsis in the already constructed structure, but nothing more. We thus expect ACD resolution to be facilitated in both the *small ellipsis* and the *large ellipsis* conditions.

In summary, while we expect a main effect of *ellipsis size* in the case of Experiment 3a (with *every*), we expect an interaction between *ellipsis size* and *position of also* in Experiment 3b (with *which*), such that the processing of the *which-was* condition is made easier with *high also* compared to *low also*. Both experiments contain conditions that will serve to replicate the results of the previous experiments. Furthermore this experiment will present novel data on the online processing of questions with interveners.

3.3.4 Results

122 native speakers of English participated in this study: 64 subjects participated in Experiment 3a and 58 subjects participated in Experiment 3b. 20 subjects were excluded from the analysis of Experiment 3a and 13 subjects were excluded from the analysis of Experiment 3b using the same exclusion criteria specified in Experiment 1. Two target sentences and two filler sentences were excluded from the analysis of Experiment 3a and two target sentences and one filler sentence were excluded from the analysis of Experiment 3b because of low accuracy (<60% across all participants).

Questions across the full experiment (targets and fillers) were answered correctly 85.7% of the time across participants in Experiment 3a and 86.4% of the time across participants in Experiment 3b; questions for experimental items were answered correctly on 83.4% of trials in Experiment 3a and on 83% of 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 we present first the results of Experiment 3a and then those of Experiment 3b. Figure 6 shows the mean RRTs for the region of interest for the four target conditions in Experiment 3a. Recall that Experiment 3a contained *every* as the determiner heading the object relative clause in all target sentences.
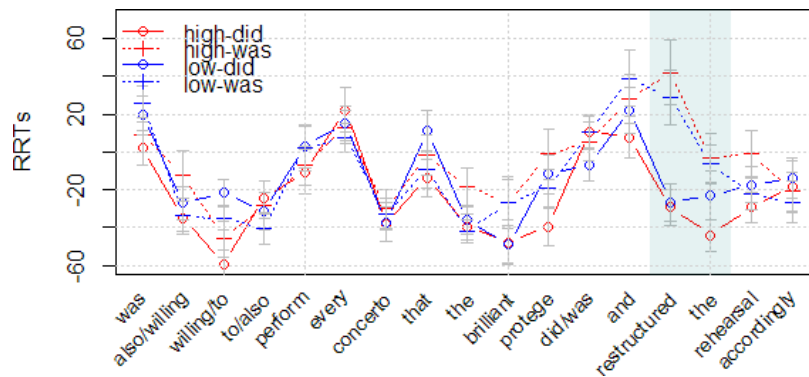

Figure 6: Residual reading times in target items in Experiment 3a

26

A linear mixed effects model with random intercepts for subjects and items was fit to the data. The model predicted *logRT* from the two factors of interest: *position of also* (low, above the embedded verb vs. high, above the matrix verb) and *ellipsis size* (small ellipsis, marked by *did*, vs. large ellipsis, marked by *was*). Results show a main effect of *ellipsis size* two and three words after the auxiliary site (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 the *low also* and *high also* conditions. There were no other significant effects in the data.

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 7 shows the mean RRTs for the region of interest for the four target conditions.
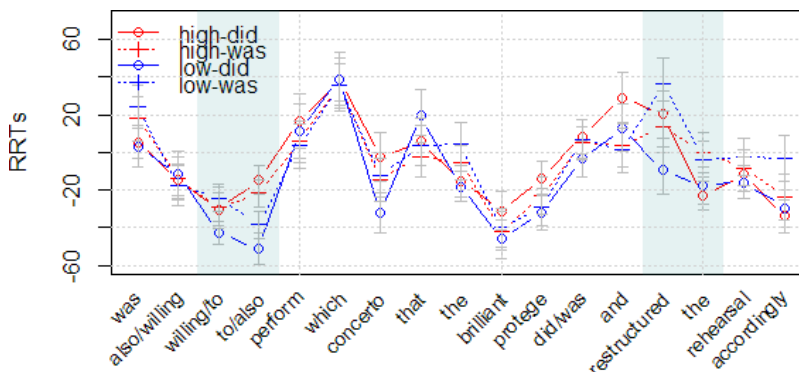

Figure 7: Residual reading times in target items in Experiment 3b

A linear mixed effects model with random intercepts for subjects and items was fit to the data. The model predicted *logRT* from the two factors of interest: *position of also* (low, above the embedded verb vs. high, above the matrix predicate) and *ellipsis size* (small ellipsis, marked by *did*, vs. large ellipsis, marked by *was*). The results show a main effect of *position of also* at the third and fourth word in the region of interest (where *also* occurs in the sentences), such that the processing of sentences with *low also* is faster than sentence with *high also*. In addition, there is a *position of also×ellipsis size* interaction at the fourth word in the region of interest (log likelihood tests, p's<0.05). The results further show a main effect of *ellipsis size* two and three words after the auxiliary verb site and an interaction between *position of also×ellipsis size* two words after the auxiliary site (log likelihood tests, p's<0.05). The main effect of *ellipsis size* is driven by the fact that the overall resolution of *small ellipsis* is faster than the resolution of *large ellipsis* for both the *low also* and *high also* conditions, while the interaction is caused because with *high also*, the two *ellipsis* conditions are processed at the same rate, but with *low also*, the *small ellipsis* is faster than the *large ellipsis*. There were no other significant effects in the data.

3.3.5 Discussion

Two effects distinguish between Experiments 3a-b. First, we observe that the presence of *also* affects participants' behavior in Experiment 3b but not in Experiment 3a. This can be explained if the presence (and position) of *also* affects participants' behavior only if it influences the

structure-building choices made by the parser, as in the case of in-situ *wh*-phrases and not in the case of *every*.[15]

More importantly, we find that the *also* manipulation affects the resolution of ACD differently in Experiments 3a-b. Specifically, *also* appears not to affect the results of Experiment 3a, where we replicate the results of Experiment 1 with *every*: we observe a main effect of *ellipsis size* and nothing else. In the case of Experiment 3b, on the other hand, we find a *position of also×ellipsis size* interaction. The *low also* condition replicates the results of Experiment 1 with *which*: we observe an effect of *ellipsis size* and nothing else. In the *high also* condition, the two *ellipsis* conditions are processed at the same speed, faster than the *large ellipsis* and slower than the *small ellipsis* with a *low also*.

To see this more closely, consider Figure 8, which compares the behavior of the four target conditions in Experiments 3a-b two words after the auxiliary verb. The graph on the left shows RRTs for sentences with an object relative clause hosted by *every* (Experiment 3a) and the graph on the right shows RRTs for sentences with an object relative clause hosted by *which* (Experiment 3b).
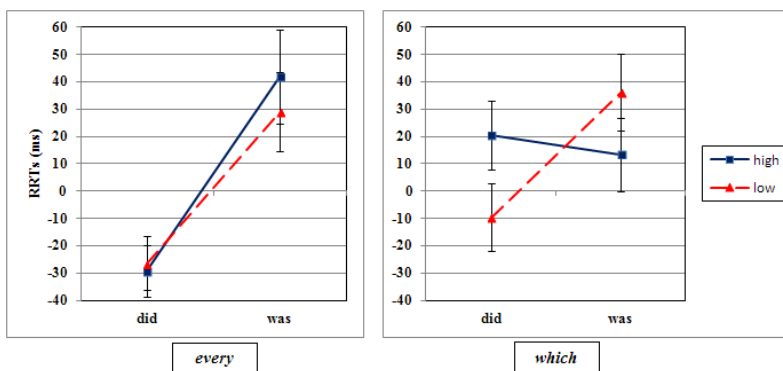

Figure 8: Residual reading times two word after the auxiliary verb in Experiments 3a-b

The results of Experiment 3a exhibit parallel behavior for the two *also* conditions: the only effect here is that of the *ellipsis size*: *small ellipsis* is processed faster than *large ellipsis*. In Experiment 3b, on the other hand, the two *also* conditions exhibit a crossing pattern: with *low also*, the *small ellipsis* is processed faster than *large ellipsis*, just as in Experiment 3a. With *high also*, we find similar RTs for both ellipsis conditions. Comparing RTs across the two *also* conditions, *small ellipsis* is parsed <u>slower</u> with a *high also* than with a *low also*, and *large ellipsis* is parsed <u>faster</u> with a *high also* than with a *low also*.

We also find a main effect of *ellipsis size* in Experiment 3b, occurring two and three words after the gap site. This represents the complexity of the antecedent effect, which is found here as well as in all the other experiments presented in this paper: although an intervener may force movement that preemptively undoes antecedent containment, resolving the ellipsis to a larger verbal complex is nonetheless more difficult than resolving it to a smaller, more local one. The

---

[15] We have no explanation for the interaction that we observe on word 4 in Experiment 3b.

main effect thus reflects the fact that composing a smaller more local VP (e.g., *perform*) into the structure is easier than composing a larger, more complex VP (e.g. *willing to perform*).

The *position of also×ellipsis size* interaction in Experiment 3b, exhibiting a facilitation effect of non-local ACD resolution with *high also*, is predicted by the *smallest movement* approach. Specifically, if interpretability requirements force long-distance covert *wh*-movement, we expect ACD to be facilitated in the entire movement domain. In Experiment 3b, the *high also* forces *wh*-movement above it and as a result, antecedent containment is preemptively undone not only in the case of *small ellipsis* but also in the case of *large ellipsis*. When the parser reaches the gap site in both conditions, all it has to do is find an appropriate antecedent for the ellipsis and nothing more. This should translate into a facilitation effect both for the *small ellipsis* and for the *large ellipsis*.

Finally, note that the results of Experiment 3b are not fully predicted by what we have said above: the RTs for both *small ellipsis* and *large ellipsis* with a *high also* do not pattern with the RTs of the *small ellipsis* with a *low also*. Recall that our previous discussion concluded that if covert movement preemptively undoes antecedent containment, we expect ACD resolution to be facilitated in the entire domain of movement. In the case of *small ellipsis*, sufficient covert movement occurs to preemptively undo containment in both *also* conditions. However, the resolution of the *did-high also* condition is unexpectedly slower than the *did-low also* condition.

We propose that this effect in Experiment 3b is the result of a scope-matching preference effect observed in both off-line judgments (Hardt and Romero 2004) and real-time sentence processing (Breakstone et al. 2011): there is a preference for the scope of an expression that hosts an ACD gap to match the size of an elided constituent in the same sentence.

(29) **Scope-matching preference principle (Breakstone et al. 2011)**
      The scope of a phrase hosting an ACD site is the lowest interpretable position that allows ACD resolution.

Experiment 3b contains two *harmonious* conditions and two *mismatched* conditions. In the harmonious conditions, the size of covert movement forced by the intervener matches the size of the antecedent required for ACD resolution. These conditions include the *small ellipsis* with *low also* and the *large ellipsis* with *high also*. There were two mismatched conditions: in the case of *large ellipsis* with *low also*, the movement forced by the intervener is not sufficient to undo antecedent containment and hence when the parser reaches the gap site, it must reanalyze the structure by covertly moving the *which*-relative clause from its position above the intervener to a position above the matrix VP_1. This reanalysis causes difficulty, reflected in reading times, which we have also seen in other experiments. The other mismatched condition in Experiment 3b is *small ellipsis* with *high also*: in this case, the movement of the *which*-relative clause is sufficiently high to undo antecedent containment, and in fact it is to a position that is larger than necessary for ACD resolution: the movement targets the larger VP_1, but ACD is resolved to the smaller VP_2.

The mismatch of the kind observed in the *small ellipsis* with *high also* condition has been shown to cause processing difficulties in a previous sentence processing study. Breakstone et al. (2011) find that the online processing of sentences with ACD is slowed when the size of QR in the sentence is larger than what is required for ACD resolution. Breakstone et al. find that the processing of a *small ellipsis* condition is more difficult than the processing of a parallel *large ellipsis* condition, when the quantifier heading the DP (in their case, a comparative quantifier) is interpreted as taking scope in the matrix clause but ACD is resolved to an embedded clause. A similar effect has been observed for children in a study on the acquisition of ACD in English (Sugawara et al. 2013) and by Hardt & Romero (2004).

Experiment 3b shows a similar effect: in the *high also* condition, the intervener forces *non-local* movement above VP_1. In the *large ellipsis* condition, the antecedent is of the same size as the movement step forced by the intervener and hence we observe a facilitation effect – relatively lower RRTs for the *was* condition when *also* was attached high compared to when it was attached low. In the *small ellipsis* condition, on the other hand, the movement step is larger than the size of structure required to resolve the ACD and we observe increased RRTs for the *did* condition when *also* was attached high compared to when *also* was attached low, indicating a temporary increase in processing difficulty. The combined effects of these processes is that the two *high also* conditions are processed at a similar speed.

4 General discussion and conclusion

The results of Experiments 1-3 shed light on the interaction between properties of in-situ *wh*-phrases and ACD resolution. More specifically, our findings suggest the following distribution at LF: in-situ *wh*-phrases cannot be interpreted in their base position, but also do not necessarily move to C for interpretation. Instead, they can be interpreted at any propositional node in the structure, just like regular quantifiers. Unlike regular quantifiers, however, in-situ *wh*-phrases are subject to intervention effects. Thus, the presence of an element like *also*, which projects a domain of intervention, can force in-situ *wh*-phrases to move higher than a regular quantifier in order to escape the intervention effect.

The evidence that supports these claims comes in the form of effects that in-situ *wh*-phrases have on ACD resolutions in various environments. In neutral environments, we observe that in-situ *wh*-phrases interact with ACD resolution in the same way that *every* does: relative to a baseline provided by the definite article *the*, ACD resolution is facilitated. In intervention environments, in-situ *wh*-phrases can facilitate non-local ACD resolution, while *every* can't.

To see how facilitation of ACD resolution might arise as a consequence of the position of the in-situ *wh*-phrase at LF, observe that ACD resolution involves at least three steps: (a) creating a structure in which antecedent containment is undone; (b) identifying an appropriate antecedent for the ellipsis; and (c) filling the antecedent into the gap and computation of the resulting meaning. The latter two steps are required for all cases of VP ellipsis, and we will refer to them

here as the *ellipsis resolution* step. The former *movement step*—the reanalysis of the structure so as to undo antecedent containment—is only required in the case of ACD. The facilitation effect that we find in our experiments is the result of movement that preemptively undoes antecedent containment in the structure, so that step (a) need not be performed when the presence of an ACD site is detected by the parser.

Experiments 1-2 show parallel behavior of *every* and *which*: both determiners facilitate the resolution of *small ellipsis* but not of *large ellipsis* (relative to our baseline provided by *the*): resolving a larger, more complex ellipsis that takes a non-local antecedent is more difficult than resolving a smaller, less complex one, that takes a local antecedent. The facilitation effect is explained if the parser must assume a small QR step of the object DP to the nearest propositional node in order to integrate both *every* and *which* into the structure, because they are quantifiers that cannot semantically compose with the verb in-situ. This early QR step preemptively undoes antecedent containment in the case of *small ellipsis*: step (a) of ACD resolution can be avoided in this case, resulting in faster reading times. This step cannot be avoided for the larger ellipsis, leading to higher processing costs. For *the*, no covert movement is assumed when it is encountered by the parser: *the* can be composed as the sister of a verb in-situ. The parser assumes covert movement in the structure only when it reaches the gap site of ACD, resulting in slower reading times for both *small ellipsis* and *large ellipsis* with *the*.

The scope-matching effect in Experiment 3b again highlights that step (a) of ACD resolution can be avoided whenever antecedent containment is preemptively undone: an intervener can force long-distance movement early in the parse of the sentence, so that when the gap site is reached no reanalysis is needed in order to find an appropriate antecedent for long distance ACD, resulting in faster reading times in those cases. In harmonious conditions, the size of movement matches the size of the antecedent, facilitating the *ellipsis resolution* step of ACD resolution by making the most recently constructed VP the appropriate antecedent for ellipsis resolution.[16] Mismatched conditions were difficult to parse, because more movement was assumed by the parser than necessary for ellipsis resolution: the VP that was constructed most recently is not a good match for the ellipsis, and instead memory must be searched for a previously constructed VP. As a result, *ellipsis resolution* is made more difficult.

As we have seen, the results of Experiments 1-3 are unpredicted by the traditional approaches to the interpretation of *wh*-in-situ. Both theories incorrectly predict invariability with regard to ACD resolution, such that ellipsis resolution should either be facilitated always (the *covert movement* approach) or never (the *in-situ* approach). However, the resolution of *small ellipsis* in Experiment 1—targeting the smaller VP in the sentence—was facilitated, but the *large ellipsis*— targeting a higher VP—incurred a high processing cost. The fact that ellipsis whose resolution requires *non-local* movement of the in-situ *wh*-phrase was not facilitated in Experiment 1 shows that the *wh*-phrase *need not* and in fact *does not* move to C immediately upon integration into the

---

[16] The VP that is most recently constructed is the sister of the head of the movement chain binding the trace position, see fn 2.

structure. The fact that we see an increased effect of ACD facilitation with *which* compared to *the* in Experiment 2 confirms that *which* is not *in situ* at LF but instead patterning with *every* and undergoing a small QR step. This shows that the in-situ approach to *wh*-in-situ also cannot be correct.

We have proposed that *wh*-phrases are quantifiers that minimally must QR covertly to the first propositional node in the structure to resolve a type-mismatch, much like conventional quantifiers such as *every*. We furthermore proposed the *partial movement* approach. This approach can be seen as an amendment to the *in-situ* approach: *wh*-phrases must move for basic structure-building reasons and interpretability, but they do not move any more than is needed for these external reasons. Once a *wh*-phrase is in an interpretable position, it can make its contribution to the interrogative semantics without moving to C. As a consequence, we require a theory of interrogative syntax/semantics that allows in-situ *wh*-phrases to be interpreted in positions other than C, and immediately below C as well.[17]

Experiment 3 corroborated an important prediction of the *partial movement* approach: that it should be possible to distinguish the behavior of *which* from that of *every*, and in particular that if a *wh*-phrase is forced to move even higher than its QRed position for interpretation, we should find facilitation effects in the entire domain of movement. Experiment 3 tested this prediction by using *intervention effects*, which have been argued to require movement of *wh*-phrases above an intervener for interpretation, but do not affect the behavior of other quantifiers such as *every*. We argued that *also* is an intervener in English and showed that, indeed, when *also* forces *non-local* movement of an in-situ *wh*-phrase, we see facilitation effects in the resolution of non-local ACD. The sensitivity of *which*, and the insensitivity of *every*, to the same manipulation show that *also* does not simply act as a scope marker for covert movement: such a proposal would incorrectly predict similar effects on *every* and *which*. Instead, we see *also* affecting only the *wh*-phrase but not the quantifier, consistent with the predictions of the intervention effect theory (Beck 2006) and the predictions of the *partial movement* approach.[18]

If the *partial movement* approach to *wh*-in-situ is on the right track, we hope to have shown that it is not captured by the traditional approaches to interrogative syntax/semantics. *Wh*-phrases cannot be interpreted where they are pronounced, but they also need not move to C for their interpretation. Instead, *wh*-phrases are able remain in their lowest interpretable position, and from there make their contribution to interrogative semantics without any further movement.

---

[17] The development of such a theory goes beyond the scope of the present paper. See Cable (2007; 2010) for a proposal that contains the ingredients to interpret structures like we propose.

[18] We note that as a consequence, any theory of intervention effects that derives intervention without any covert movement faces difficulties in explaining the results of Experiment 3.

References

Anderson, C. 2004. The structure and real-time comprehension of quantifier scope ambiguity, PhD dissertation, Northwestern University.

Arnon, I., Snider, N., Hofmeister, P., Jaeger, T.F, Sag, I. Cross-linguistic variation in a processing account: The case of multiple wh-questions. To appear in Proceedings of BLS 32.

Arregui, A., C. Clifton, L. Frazier, and K. Moulton. 2006. Processing elided verb phrases with flawed antecedents: The recycling hypothesis. Journal of Memory and Language 55:232–246.

Aoun, J. and Li A. Y-H. 1993. Syntax of Scope. Cambridge, Mass.: MIT Press.

Bates, D.M. and Maechler, M. 2009. lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-32.

Bates, D., and Sarkar, D. 2006. lme4: Linear mixed-effects models using S4 classes. R package version 0.995-2.

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

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

Barr, D. Levy, R., Scheepers, C., and Tily, H.J. 2012. Random effects structure in mixed-effects models: Keep it maximal. Ms.

Beaver, D. and Clark. B. 2008. Sense and sensitivity: How focus determines meaning. Wiley-Blackwell.

Beck, S. and Kim, S.-S. 1997. On wh- and operator scope in Korean. Journal of East Asian Linguistics 6: 339-384.

Beck, S. 1996. Quantified structures as barriers for LF movement. Natural Language Semantics 4:1-56.

Beck, S. 2006. Intervention effects follow from focus interpretation. Natural Language Semantics 14:1-56.

Bever, T. 1970. The cognitive basis for linguistic structures. In Cognition and the Development of Language, J.R. Hayes, ed., Wiley, New York: 279-362.

Breakstone, M. Y., Cremers, A., Fox, D., and Hackl, M. 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, ed. by Ashton N., Chereches A., and Lutz D., 712-731.

Cable, S. 2007. The grammar of Q: Q-particles and the nature of *wh*-fronting, as revealed by the *wh* questions of Tlingit. PhD dissertation, MIT.

Cable, S. 2010. The Grammar of Q: Q-Particles, Wh-Movement and Pied-Piping. Oxford: Oxford University Press.

Cheng, L. L-S. 1991. On the typology of wh-questions. PhD dissertation, MIT.

Chomsky, N. 1981. Lectures on government and binding, Dordrecht (Holland): Foris Publications.

Chomsky, N. 1991. Some notes on economy of derivation and representation. In Principles and Parameters in Comparative grammar. R. Freidin, ed., Cambridge: MIT Press.

Chomsky, N. 1993. A minimalist program for linguistic theory. In: The View from Building 20: Essays in the Honor of Sylvain Bromberger, K. Hale and S.J. Keyser, eds., Cambridge: MIT press.

Chomsky, N. 1995. The Minimalist Program. Cambridge, Mass.: MIT Press.

Chomsky, N. 2000. Minimalist inquiries: the framework. In Step by Step: Essays on Minimalist Syntax in Honour of Howard Lasnik, R. Martin et al., eds., Cambridge, Massachusetts: MIT press.

Clifton, C. J., Fanselow, G., and Frazier, L. 2006. Amnestying superiority violations: processing multiple questions. Linguistic Inquiry 37: 51-68.

Collins, C. 2001. Economy conditions in syntax. In Handbook of Syntactic Theory, C. Collins and M. Baltin eds., Blackwell publishing.

Dornish, E. 1998. Overt quantifier raising in Polish. Unpublished Manuscript.

Van Dyke, J. A., and Lewis, R. L. 2003. Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities. Journal of Memory and Language, 49, 3: 285-316.

Epstein, S., 1992. Derivational constraints on A' chain formation. Linguistic Inquiry 23(2).

Fanselow, G., Lenertovà D., and Weskott, T. 2008. Studies on the acceptability of object movement to Spec,CP. In The Discourse Potential of Underspecified Structures, A. Steube ed., Berlin: Mouton de Gruyter. 413-437

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

Featherston S. 2005b. Universals and grammaticality: Wh-constraints in German and English. Linguistics 43:4.

Fox, D. 1995. Economy and scope. Natural Language Semantics, 3 (3): 283-341.

Fox, D. 2000: Economy and Semantic Interpretation. MIT Press: 727-796.

Fox, D. 2002. Antecedent Contained Deletion and the Copy Theory of Movement. Linguistic Inquiry 33(1):63-96.

Fox, D. 2003. On logical form. In Minimalist Syntax, R. Hendrick ed., Blackwell publishers, Oxford, 82–123.

Frazier, L.,and Clifton, C., Jr. 2002. Processing D-linked phrases. Journal of Psycholinguistic Research 31(6):633-659.

Frazier, L., and Rayner, K. 1982. Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. Cognitive Psychology 14:178-210.

Frazier, L. 1999. On Sentence Interpretation. Dordrecht: Kluwer.

Gibson, E. 1998. Linguistic complexity: Locality of syntactic dependencies. Cognition, 68, 1–78.

Hagstrom, P. 1998. Decomposing questions. PhD dissertation, MIT.

Hamblin, C. L. 1973. Questions in Montague English. Foundations of Language 10(1):41-53.

Hackl M., Koster-Hale, J. and Varvoutis, J. 2012. Quantification and ACD: Evidence from real time sentence processing. Journal of Semantics 29:145-206.

Hardt, D. and Romero, M. 2004. Ellipsis and the Structure of Discourse, Journal of Semantics 21, 375-414.

Heim I. and Kratzer, A. 1998. Semantics in generative grammar. Blackwell publishers, Malden, MA.

Hofmeister, P., Jaeger, T.F., Sag, I., Arnon, I., and Snider, N. 2007. Locality and accessibility in wh-questions. In Roots: Linguistics in Search of its Evidential Base, S. Featherston and W. Sternefeld eds., Berlin: Mouton de Gruyter.

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

Hornstein, N. 1995. Logical Form: from GB to Minimalism, Cambridge, Mass: Blackwell publishers.

Huang, J. C.-T. 1982. Logical relations in Chinese and the theory of grammar, PhD Dissertation, MIT.

Huang, J. C.-T. 1995. Logical Form. In: Government and Binding Theory and the Minimalist Program, G. Webelhuth ed., Cambridge, Massachusetts: Blackwell, 127-173.

Johnson, K. 1996. When verb phrases go missing, GLOT International 2.5:3–9.

Just, M. A., Carpenter, P. A., and Woolley, J. D. 1982. Paradigms and processes and in reading comprehension, Journal of Experimental Psychology: General 111: 228-238.

Karttunen, L. 1977. Syntax and semantics of questions. Linguistics and Philosophy 1(1):3-44.

Kehler, A. 2002. Coherence, reference, and the theory of grammar. Center for the Study of Language and Information.

Kertz, L. 2010. Ellipsis reconsidered. Doctoral Dissertation, UC San Diego

Kitahara, H. 1997. Elementary Operations and Optimal Derivations. Cambridge, Mass. MIT.

Kim, S. W. 1991. Chain scope and quantification structure. PhD dissertation, Brandies University.

Kim, S.-S. 2002. Focus matters: two types of intervention effects. Paper presented at WCCFL 21, Santa Cruz.

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

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

Kratzer, A. and Shimoyama, J. 2002. Indeterminate pronouns: The view from Japanese. In Y. Otsu. ed. The Proceedings of the Third Tokyo Conference on Psycholinguistics (TCP 2002), 1-25. Tokyo: Hituzi Syobo.

Larson, R. and May, R. 1990. Antecedent containment or vacuous movement: reply to Baltin. Linguistic Inquiry 21:103-122.

Lasnik, H., and Saito, M. 1992. Move α, Conditions on its Application and Output. Cambridge, Mass: MIT Press.

Lewis, R. L., Vasishth, S., and Van Dyke, J. A. 2006. Computational principles of working memory in sentence comprehension. Trends in Cognitive Science, 10, 10: 447-454

May, R. 1985. Logical Form: Its structure and derivation. MIT Press. Cambridge, MA.

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

Merchant, Jason. 2013. Voice and ellipsis. Linguistic Inquiry 44.

Nishigauchi, T. 1986. Quantification in syntax. PhD dissertation, University of Massachusetts, Amherst.

Pesetsky, D. 1987. Wh-in-situ: Movement and unselective binding. In The representation of (in)definiteness, E. Reuland and A. ter Meulen, eds., Cambridge, MA.: MIT Press.

Pesetsky, D. 2000. Phrasal Movement and its Kin. Cambridge, MA: MIT Press.

Phillips, C. 2003. Parsing: Psycholinguistic aspects. In International Encyclopedia of Linguistics, 2nd edition. Oxford University Press.

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

Reinhart, T. 2006. Interface Strategies. Cambridge, MA: MIT Press.

Richards, N. 2001. Movement in Language: Interactions and Architectures, Oxford University Press.

Rooth, M. 1985. Association with focus. PhD dissertation, University of Massachusetts, Amherst.

Rooth, M. 1992. A theory of focus interpretation. Natural Language Semantics 1:75-116.

Rullmann, H. and Beck, S. 1998. Reconstruction and the interpretation of which phrases. In Reconstruction: Proceedings of the 1997 Tübingen Workshop, G. Katz, S.-S. Kim and H. Winhart, eds., Universität Stuttgart and Universität Tübingen, 233-256.

Sag, I. 1976. Deletion and Logical Form. PhD dissertation, MIT.

Shimoyama, J. 2006: Indeterminate phrase quantification in Japanese. Natural Language Semantics 14: 139-173.

Soh, H. 2005. Wh-in-situ in Mandarin Chinese. Linguistic Inquiry 36:143-155.

Sugawara, A., Kotek, H., Hackl, M. and Wexler, K. 2013. Long vs. short QR: Evidence from the acquisition of ACD. Proceedings of Boston University Conference on Language Development (BUCLD) 37. Somerville, MA: Cascadilla Press.

Tomioka, S. 2007a. Intervention effects in focus: From a Japanese point of view. In Working Papers of the SFB632, Interdisciplinary Studies on Information Structure (ISIS) 9, S. Ishihara, S. Jannedy and A. Schwarz eds., Potsdam: Universitätsverlag Potsdam, 97-117.

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

Tsai. W, -T.D. 1994. On nominal islands and LF extractions in Chinese. Natural Language and Linguistic Theory 12:121-175.

Tunstall, S. 1998. The interpretation of quantifiers: Semantics and processing. PhD dissertation, University of Massachusetts, Amherst.

Warren, T. and Gibson, E. 2002. The influence of referential processing on sentence complexity. Cognition 85: 79-112.

Williams, E. 1974. Rule Ordering in Grammar. PhD dissertation, MIT.

Williams, E. 1977. Discourse and logical form. Linguistic Inquiry 8:101-139.

Appendix: Materials for Experiments 1-3

Below is the full of items used in Experiments 1-3. The same item paradigms were used in all three experiments, with different choices of determiner and placement/position of *also* as described in the body of the paper. In short: Experiments 1 and 3 compared the determiners *every* and *which*, while Experiment 2 compared the determiners *the* and *every*. Experiment 3 compared the behavior of a *high* and a *low* placement of *also*, while Experiments 1-2 did not contain *also*.

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.