Direct-access retrieval during sentence comprehension: Evidence from Sluicing

Andrea E. Martin[1] and **Brian McElree**[2]

[1] Basque Center on Cognition, Brain, and Language (BCBL)

[2] New York University

Abstract

Language comprehension requires recovering meaning from linguistic form, even when the mapping between the two is indirect. A canonical example is ellipsis, the omission of information that is subsequently understood without being overtly pronounced. Comprehension of ellipsis requires retrieval of an antecedent from memory, without prior prediction, a property which enables the study of retrieval *in situ* (Martin & McElree, 2008, 2009). Sluicing, or inflectional phrase ellipsis, in the presence of a conjunction, presents a test case where a competing antecedent position is syntactically licensed, in contrast with most cases of nonadjacent dependency, including verb phrase ellipsis. We present speed-accuracy tradeoff and eye-movement data inconsistent with the hypothesis that retrieval is accomplished via a syntactically guided search, a particular variant of search not examined in past research. The observed timecourse profiles are consistent with the hypothesis that antecedents are retrieved via a cue-dependent direct-access mechanism susceptible to general memory variables.

Keywords

sentence processing; ellipsis; sluicing; speed-accuracy tradeoff; serial position; recency; retrieval interference

Language comprehension relies on a correspondence between *linguistic form* and *meaning*. As such, the primary goals of linguistic and psycholinguistic theory have been to understand the nature of this correspondence and to uncover the psychological processes that compute it. However, the exact relationship between the overt linguistic signal (e.g., speech or written text) and its meaning or message (what the producer intends or what the comprehender understands) is not always straightforward. From perceiving reduced vowels to binding pronouns and moved traces, resolving anaphoric expressions or interpreting pragmatic implication, language users must compute meanings when some of the "ingredients" of the meaning have gone missing. Nowhere else does the correspondence between form and meaning fail as dramatically as during ellipsis, or omission of a phrase that is understood *without being pronounced*. Via this omission, ellipsis functions as a silent "natural compression algorithm," where silent meaning appears to function in the same complex way that overt structure does (Merchant, 2001).

Comprehension of nonadjacent dependencies, such as ellipsis, requires, at minimum, the recovery of previously processed information from memory. The nature of the retrieval process that subserves dependency processing has been studied in several linguistic structures, notably, filler/gap dependencies, and recently, verb-phrase ellipsis (Martin & McElree, 2008, 2009; McElree, 2000; McElree, Foraker, & Dyer, 2003; Van Dyke & McElree, 2006; Van Dyke, 2007). These studies have consistently indicated that the retrieval process is a *direct-access* operation over *content-addressable* representations, without a search through irrelevant representations.

The primary evidence for direct access is that speed of processing, and therefore retrieval speed, is unaffected by increasing distance between target and retrieval site, nor affected by the number of competing representations in memory (see McElree, 2006 for a review). In contrast, the aforementioned research has found substantial decrements in accuracy, or the likelihood of successful comprehension, resulting from increases in distance between antecedent and retrieval site, as well as from such other manipulations as increasing the number and similarity of intervening elements and the number of referential entities. For example, the probability of computing an acceptable interpretation decreases in filler-gap dependencies such as *This was the book that the editor admired*, where *the book* must be associated as the direct object of the final verb, *admired* (from McElree et al., 2003) when material must be interpreted between the filler *the book* and the gap in the direct object position. However, increasing the amount of interpolated material does not impact processing speed. These findings are at odds with a search process being operative during comprehension. Rather, they suggest that linguistic representations are content-addressable, i.e., elicited from memory via their content, and directly accessible via the cues provided by the verb at the dependency site. The key principle of direct-access retrieval from memory with a content-addressable architecture is that representations with varying degrees of distinctiveness can be recovered in equal time, without a search.

A potential limitation of studies using filler-gap constructions, or related studies using subject-verb dependencies (McElree et al., 2003), is that the grammar marks the dependent constituent, the subject or the filler, as having a future role in the sentence— subjects must be unified with a verb and fillers must receive a grammatical and semantic role from a gap position. It is possible, then, that the target constituents in question were assigned some special status in memory in these constructions. In fact, models of parsing often assume that these constituents in these types of dependencies are held in specialized stacks or buffers, which can mimic direct-access (McElree et al., 2003).

To address this concern, Martin and McElree (2008, 2009) investigated the processing of verb-phrase ellipsis (VPE). Crucially, antecedents of VPE are fully integrated in local sentential context, and there is no overt grammatical marking signaling a further role downstream, although antecedents of VPE must meet certain licensing requirements, such as being given or previously in focus. VPE engenders retrieval *in situ*, where recovery of the antecedent is triggered only by cues at the ellipsis site. These studies also varied distance between antecedent and VPE, as well as antecedent complexity. Neither factor affected processing speed, but as with subject-verb and filler-gap dependencies, interpreting more material within the dependency decreased accuracy. Again, these studies provide support for the hypothesis that, in general, the representations formed during comprehension are content-addressable and retrieved with a cue-dependent direct-access operation, not an iterative search process.

Yet another limitation of previous studies is that the competing constituents that were used to detect a search process were never in syntactically licensed positions. These competitors might have been semantically or pragmatically appropriate matches to the target item, but

they were not syntactically available. A search process can be defined using varying amounts and kinds of content to form the search set. Depending on how the search is defined, syntactic information may gate the search set to include only items that are in licensed positions. If this were so, previous studies would not have been able to detect a search process: the search set would only include syntactically available items and thus the *de facto* search size in most prior research would be one unit, the target constituent. This situation would yield the same pattern of results (e.g., speed insensitivity to set size) that has been seen in previous research and used to argue for direct-access retrieval.

In this paper, we extend this research by investigating the retrieval mechanism underlying resolution of *sluicing*, a construction first described by Ross (1969), and also notably by Chung, Ladusaw, and McCloskey (1995). Sluicing has most recently been analyzed by Merchant (2001) as a form of inflectional phrase (IP) ellipsis, where an interrogative clause is reduced to a *wh*-phrase. Sluicing is more widespread than VPE in the world's languages (Merchant, 2001), and like VPE, allows the study of retrieval *in situ*. However, sluicing differs from VPE in at least one important way: it reduces to a *wh*-phrase at its ellipsis site. In order for the *wh*-phrase to have a meaning, the missing IP that the *wh*-element moved out of must be interpreted at the sluicing ellipsis site (SES). In effect, the *wh*-item sits in specifier position of the complementizer phrase (CP) that dominates the missing IP. Thus, the "sluice" itself is a CP whose sentential domain or IP is elided:

1)   Michael$_1$ [$_{IP}$[$_{VP}$ typed something]] but he$_1$ didn't tell me [$_{CP}$ what$_2$ [$_{IP1}$ $_{VP2}$]].[1]

The meaning of the sentence can be paraphrased as "Michael typed something but he didn't tell me what he typed." That the *wh*-item need not correlate to an overt form, although it also may, presents a formal puzzle for structural isomorphism accounts of sluicing and of ellipsis more broadly. Merchant (2001) shows how this fact might be better captured by the deletion of the IP being triggered by a different feature – not by morphosyntactic identity with the antecedent IP *alone*, but by the e-GIVENness[2] of the deleted IP. In other words, what can be sluiced is what is entailed by the prior context. In (1), [$_{IP}$ he [$_{VP}$ typed [t]]] can be elided because it has mutual entailment with [$_{IP}$ Michael [$_{VP}$ typed something]], and the *what* remains as the result of regular *wh*-movement into the specifier position of CP for an interrogative.

These properties, the overt *wh*-remnant and the lack of reliance only on strict morphosyntactic isomorphism, allows for the IP that must be interpreted at the SES to occur in a conjunction and not be ambiguous, as would be the case in VPE, since the remnant *wh*-phrase carries further information about the identity of the missing clause:

2)   Michael$_1$ [$_{IP}$[$_{VP1}$ drank coffee] and [$_{VP2}$ typed something]] but he$_1$ didn't tell me [$_{CP}$ what$_2$ [$_{IP1}$[$_{VP2}$]]].

This fact, in combination with a conjunction in the sentence context, allows us to manipulate the position of the target constituent in memory, or its *serial position*, in a way that could not be done in VPE. This in turn provides a potentially strong test of whether the same type of retrieval operation is used in the recovery of all types of dependent constituents, and especially a case where syntactic information does not exclude competing items.

Serial position in memory

Timecourse investigations of several basic memory tasks indicate that there are two distinct ways that memories can be retrieved (see McElree, 2006, for a review). One mechanism

---

[1]The examples used in the text are derived from Frazier & Clifton (2005); we further augmented their items for use in our studies.
[2]Or mutual semantic entailment between the antecedent and elided phrases, modulo *E*-type shifting.

involves a search operation where memory representations are sequentially sampled, either forward or backward, until the required information is recovered. This type of serial operation appears to be the primary means by which relational (temporal and spatial order) information is recovered (Gronlund, Edwards, & Ohrt, 1997, McElree, 2001, 2006; McElree & Dosher, 1993). The signature pattern of this type of operation is that retrieval speed varies with recency or with the serial position of an item in a study list, either increasing with serial position if a forward search is used or decreasing if a backward search is used (McElree & Dosher, 1993; Neath, 1993; Neath & Knoedler, 1994; Öztekin, McElree, Staresina, & Davachi, 2008).

The second retrieval mechanism presupposes content-addressable memory representations; information (cues) in the retrieval context provides direct access to relevant memory representations, without the necessity of a search through irrelevant representations (McElree & Dosher, 1989). Cue-dependent direct-access operations can be implemented in memory models with diverse storage architectures (Clark & Gronlund, 1996), and have been found to be the primary means through which item representations are accessed (e.g., McElree & Dosher, 1989, 1993; McElree, 2000, 2006). Unlike under a search process, the serial position of a to-be-retrieved item would have no direct effect on the speed of a cue-dependent direct-access retrieval operation.

With one exception (McElree et al., 2003, Exp. 3), all prior comprehension studies have manipulated the amount of material between dependent constituents but not necessarily the serial position of the to-be-retrieved constituent with respect to other *available* positions in a sentence. That is, when the distance between a filler and gap is increased in a sentence such as *This was the book that the editor admired* by adding a relative clause, as in *This was the book that the editor who signed the author admired*, the additional noun phrase, although pragmatically appropriate as the direct object of the final verb, is not in a syntactic position that could serve that role. That this type of manipulation does not affect retrieval speed is consistent with claims that information at the retrieval site, including grammatical constraints, functions as cues for direct-access retrieval. However, it is also consistent with a sophisticated search process in which grammatical information is used to constrain the search set. In the latter case, such a search operation would render irrelevant all other noun phrases in the sentence, including the additional one, thereby reducing the search set to an effective set of one.

Because sluicing appears to be tied to isomorphism of entailment, it offers a mirror image, in some ways, of the previous test cases, where competitors were in syntactically unavailable positions. As such, sluicing is well suited to detecting a sophisticated search process that makes use of different kinds of information at retrieval. Sluicing with a conjoined antecedent provides a context where there are two syntactically licensed antecedent positions, one where the VP is ultimately excluded by virtue of its mismatching semantic entailment. If a search for the antecedent that is guided by syntactic constraints is operative, then in this case, we would expect to see the following search pattern: longer processing time when there is an additional licensed antecedent position in the previous sentence context. To illustrate:

2) Michael$_1$ [$_{IP}$[$_{VP1}$ drank coffee] and [$_{VP2}$ typed something]] but he$_1$ didn't tell me [$_{CP}$ what$_2$ [$_{IP1}$ [$_{VP2}$]]].

Sluicing in the presence of a conjunction represents a rare case where a constituent at another position in the sentence is syntactically available as an antecedent. In search operations restricted to syntactically licensed positions, both VPs, as well as their conjunction, would be included in the search set, and their order can be varied:

3) Michael$_1$ [$_{IP}$[$_{VP1}$ typed something] and [$_{VP2}$ drank coffee]] but he$_1$ didn't tell me [$_{CP}$ what$_2$ [$_{IP1}$ [$_{VP2}$]]].

If a forward search is used, (3) should be processed faster than (2), since *typed something* is closer to the beginning of the sentence. If a backward search from the SES is used, then (2) should be faster than (3). A direct-access process predicts no difference in retrieval speed, and therefore in processing speed, as a function of serial position. Under direct-access, the cues at the SES elicit the antecedent VP from memory by virtue of its match to those cues. In this way, only the VP(s) with matching entailment can be elicited as the antecedent.

We are not the first to examine this property of sluicing in a processing context. Frazier and Clifton (2005) conducted self-paced reading and questionnaire studies on sluicing that focused on this very property, albeit with different problems in mind. They asserted the *conjunction domain hypothesis* (CDH), where sluicing from the first conjunct in a conjunction is dispreferred and costly due to syntactic domain restrictions[3]. On a related point, a sluicing antecedent VP being a conjunct poses a problem for syntactic isomorphism accounts of sluicing, but a conjoined antecedent does not seem as problematic for Merchant's entailment-based approach[4]. Frazier and Clifton (2005) found that when the sluicing antecedent verb occurred in a conjunction with an intransitive verb that could not be sluiced from, it was rated as more natural, and was read faster, when it occurred in the most recent or second position, where (2) was preferred over (3). They argued that this result indicated the presence of syntactic structure at the SES, because a purely semantic or discourse-based anaphoric resolution of sluicing would not be sensitive to the syntactic domain of the conjunction.

However, other factors may contribute to the processing asymmetry between (2) and (3). First, presence of the second VP may induce interference at retrieval, due to overlap with cues at the SES. In (2), the intervening *drank coffee* may interfere with the interpretation of the SES, even though it is ultimately unable to bind *what*, which in turn may impact processing time or comprehension accuracy, because it is a partial match to the information needed at the SES. Second, there is greater distance between antecedent and SES in (3) as compared to (2), and the intervening VP is also more recent in that it is interpolated between the antecedent and SES. Thus, the correct VP is closer to the SES in (2) as compared to (3). Both of these factors were present in Frazier and Clifton (2005), where the serial position of the antecedent verb varied. In order to address the roles of serial position and recency, we manipulated both the distance between the antecedent and the SES, and the number of VPs present in the sentence context, via the construction of two Single VP conditions:

4) In the morning, Michael typed something, but he didn't tell me what.

5) Michael typed something in the morning, but he didn't tell me what.

We hypothesized that the presence of an additional VP in the conjoined cases might impact the time it takes to resolve the dependency, especially if sluicing invokes a sophisticated search process. Given that the presence of additional material also increases the distance between antecedent verb and SES, it also seems likely that accuracy could be affected as well, as distance within a dependency has been shown to decrease comprehension accuracy (Foraker & McElree, 2007; Martin & McElree, 2008, 2009; McElree, 2000; McElree, et al., 2003). If sluicing from a distant conjunct does slow processing speed, then it is likely that

---

[3]It seems that part of the motivation for this hypothesis about domain restriction may come from a structural isomorphism view of sluicing and ellipsis, although it is not stated as such in Frazier & Clifton (2005).
[4]As we understand it, entailment-based accounts of sluicing and other forms of ellipsis are not claims whether or not there is syntactic structure at the ellipsis site, rather these accounts are claims about how ellipsis is licensed. Merchant (2001) presents evidence for and explicitly assumes syntactic structure at the ellipsis site.

structural factors, as argued by Frazier and Clifton (2005) are at play, where order information resolves the dependency. If sluicing from a distant conjunct results in lower likelihood of successful interpretation but no differences in processing speed, then the processing asymmetry between (2) and (3), as observed by Frazier and Clifton (2005), may in fact stem from other factors related to the recovery of the sluicing antecedent representation. Indeed, if sluicing antecedents are recovered directly via their content, then the retrieval process is rendered susceptible to general memory variables, one such being retrieval interference from similar representations in memory.

Retrieval interference

Much evidence implicates interference during retrieval as a determinant of processing difficulty in sentence comprehension (Gordon, Hendrick, & Johnson, 2001; Gordon, Hendrick, & Levine, 2002; Gordon, et al., 2004; Lewis, 1996; Lewis & Vasishth, 2005; Lewis, Vasishth, & Van Dyke, 2006; McElree, 2006; McElree, et al., 2003; Van Dyke, 2007; Van Dyke & Lewis, 2003; Van Dyke & McElree, 2006). Similarly, retrieval interference is a major determinant of performance in working memory tasks (Anderson & Neely, 1996; Crowder, 1976; Keppel & Underwood, 1962; see Nairne, 2002 for a review; Waugh & Norman, 1965). A notable example comes from Keppel and Underwood (1962), where the Brown-Peterson procedure was used to show that recall accuracy in a list-learning paradigm decreased as a function of the number of items in the study list, even when retention time was held constant. This result demonstrated that interference, not delay, decreased accuracy, and therefore, that time-based variables like decay or displacement alone do not determine forgetting. Accordingly, within the memory literature, there has been a substantial shift away from decay or displacement accounts towards accounts that emphasize interference at retrieval as the source of forgetting. Despite this, there are few accounts of dependency processing that specifically characterize exactly how retrieval interference accrues (cf. Lewis et al., 2006), but it seems likely that some aspect of processing additional interpolated material between a dependency and its antecedent creates interference. Van Dyke and McElree (2006) argue that interpreting additional material decreases the specificity of retrieval cues to competing constituent representations outside the focus of attention. This type of interference results from so-called *cue overload* where the cues that make contact with the correct constituent in memory are insufficient for successful retrieval (Nairne, 2002; Öztekin & McElree, 2007; Watkins & Watkins, 1975). When more material is interpreted between the target and the retrieval site, this additional processing results in the cues at retrieval being less diagnostic or specific to the target.

Given the relationship between cue overload and the position of the target, what is the source of the advantage for more recent positions over earlier ones? One explanation is that serial position is co-extended with *recency*, a factor already implicated in dependency resolution within the sentence processing literature (Frazier, 1978; Gibson, Pearlmutter, Canseco-Gonzalez, & Hickok, 1996; Kimball, 1973; Neath, 1993; Neath & Knoedler, 1994; Pearlmutter & Gibson, 2001; Sturt, Scheepers, & Pickering, 2002). Although recency is not the only dimension along which one position in a string may differ from another, later positions' relative recency likely gives them an advantage not only in representational strength of the target item, but more crucially in the specificity of the retrieval cues to the target (i.e., incurring less cue overload, and therefore less interference, when a target is recent). Higher cue specificity may lead to greater distinctiveness for more recent positions and thus higher performance (Neath & Knoedler, 1994). In this way, increasing cue overload, which includes a contribution from decreased recency, may create interference during retrieval that contributes to classic serial position effects found in the memory literature. The impact of both distance of integration and position of the target antecedent on the processing of sluiced sentences is the focus of Experiment 1.

Experiment 1

In the context of sluicing, cue overload could manifest itself in two ways: (1) accrual of interfering representations or similar information, where match between multiple items in memory and the retrieval cue prevent elicitation of the correct target, and (2) degradation of extant representations in short-term memory, where insufficient strength of the antecedent representation prevents cues at the dependency site from eliciting the antecedent verb phrase adequately enough for successful interpretation, in both cases lowering accuracy of interpretation.

In order to measure differences in processing time orthogonally from processing accuracy, we used the response-signal speed-accuracy tradeoff procedure (SAT) to examine the effects of distance and number of antecedent positions on sluicing interpretation. The primary benefits of this procedure are that the speed and the accuracy of processing can be measured conjointly within a single task (e.g., Dosher, 1979; Wickelgren, 1977; Reed, 1973, 1976) and that SAT forces participants to fully resolve the sluice in order to make their judgment. Since retrieval of order information has been known to impact processing speed, we were especially interested in measuring processing speed uncontaminated by differences in representational strength or response bias. We had participants read sentences presented phrase-by-phrase and, at designated points, decide (yes/no) whether the passage was sensible. We used a multiple response variant of the SAT procedure that has been used in several investigations of language processing (e.g., Bornkessel, McElree, Schlesewsky, & Friederici, 2004; Foraker & McElree, 2007; Martin & McElree, 2008, 2009; McElree, 1993; McElree, Pylkkänen, Pickering, & Traxler, 2006): Participants were trained to respond to an auditory response signal presented multiple times after the onset of the critical word, here a *wh*-element. Crucially, the first response signal onset occurred 300 ms before the onset of the *wh*-element, and thus participants were required to respond before processing of the critical word had begun. The subsequent sampled times (0–5250 ms) enabled us to fully measure how the interpretation of the sluice unfolded over time. For each sampled point, we constructed a d' measure of accuracy by scaling correct responses to sensible sluiced expressions (hits) against incorrect responses to control expressions with nonsensical sluiced interpretations (false alarms). This scaling provided a measure of the ability of participants to discriminate acceptable from unacceptable interpretations.

Figure 1 presents illustrative SAT functions—d' accuracy versus processing time—for two hypothetical conditions. Characteristically, the functions show a period of chance performance ($d' = 0$), a period of increasing accuracy, and an asymptotic period during which further processing does not improve performance. In our studies, the timecourse functions for each participant were fit with an exponential approach to a limit, which enabled us to quantify how the interpretation of sluicing under different sentential contexts unfolded over time:

$$d' = \lambda(1 - e^{-\beta(t-\delta)}) \text{ for } t > \delta; \text{otherwise } t = 0.$$

<div align="right">Equation 1</div>

The parameter λ, which estimates the asymptote of the function, measures the highest level of discrimination reached with maximal processing time, and hence yields a basic measure of processing accuracy. Differences in asymptote alone are illustrated in Figure 1. Conditions that vary in asymptote differ in the likelihood that a meaningful interpretation can be assigned to each type of expression or that the interpretations of the expressions differ in their overall degree of acceptability. Here, the asymptotes index how successful participants were at retrieving an antecedent for the sluice and interpreting it locally. Increasing distance or complexity should lower asymptotic accuracy if these factors

decrease the quality of the antecedent's representation in memory, making the antecedent less likely to be retrieved from memory or reducing the quality of the retrieved information.

The principal advantage of SAT is that it enables one to measure and compare the speed of interpretation of conditions that may also differ in overall accuracy. Thus, we can determine the relative speed of interpreting an expression on the respective proportion of trials in which readers succeed in computing a sensible interpretation. The intercept (δ) and rate (β) of the function provide joint measures of the speed of processing, indexing how quickly accuracy accrues to its asymptotic level. The parameter δ estimates the intercept of the function, or the point at which participants are first sensitive to the information necessary to discriminate acceptable from unacceptable sluicing (i.e. d' departs from 0, chance performance). The parameter β estimates the rate at which accuracy grows from chance to asymptote. Figure 1 illustrates two conditions that differ in rate and intercept. If one stimulus can be interpreted more quickly than another, the SAT functions will differ in rate, intercept, or some combination of the two parameters (e.g., Bornkessel et al., 2004;McElree, 1993;McElree & Nordlie, 1999;McElree, Pylkkänen, Pickering, & Traxler, 2006). Whether speed differences are expressed in rate or intercept depends on the mean and variance of the time it takes to compute the different interpretations. In some contexts, the locus of the effect can be theoretically important (e.g., McElree & Dosher, 1993). However, the predictions we tested are based on general differences in speed of processing, which can be assessed by effects on either parameter. Importantly, whether differences are expressed in rate, in intercept, or in both parameters, the associated functions will display disproportional dynamics, reaching a given proportion of their respective asymptote at different times. This is illustrated by the intersection of the horizontal and vertical lines in Figure 1, which shows the point in time (abscissa) when the functions reach two-thirds of their respective asymptote (ordinate). When processing speed is identical, as in Panel A, the functions reach this point at the same time, shown by the vertical line. When processing speed varies, as in Panel B, the functions reach a given proportion of their respective asymptotes at different times.

Method

Participants—Seventeen native speakers of American English from the New York University community were paid to participate in the study. They participated in four 75-minute sessions, and a 45-minute practice session for familiarization with the SAT procedure. All participants were between the ages of 18 and 26.

Materials—Thirty-one sets of 16 sentences of the form illustrated in Table 1 were created. The main contrasts concerned sluicing with a recent antecedent, such as the Recent Double VP (1a) and Recent Single VP (3a) conditions, sluicing with an intervening verb between the antecedent and the SES, such as the Distant Double VP (2a) and the Distant Single VP (4a) conditions, and contrasts between the Double VP and Single VP conditions. Interference was increased by placing the antecedent verb in the first position of a conjunction, such that only the position of the antecedent verb varied across conditions. For each of these conditions, we created a matching unacceptable condition, (1b) and (2b), by replacing the stranded *wh*-element of the sluice (e.g., *what, who, which*) with an unacceptable *wh*-element (also *what, who, which*), which would create an unacceptable interpretation when interpreted elliptically (e.g., **but he didn't tell me which*). These unacceptable conditions were designed to encourage participants to fully process the sluice. We reasoned that, to discriminate acceptable from unacceptable sentences, participants would have to process the ellipsis at least to the point where they had retrieved the antecedent and interpreted it in the local context.

Additionally, we included an equal number of acceptable and unacceptable, recent and distant, conjoined and unconjoined control conditions, without sluicing in the final phrase, such as the (a) and (b) versions of (5) – (8). These sentences had the same lexical content as (1) – (4), except that a final noun phrase was added to the final clause to block an elliptical interpretation (e.g., *but he didn't tell me much/*must; but he didn't tell me what to do/*fall*, depending on the item). These sentences were included to reduce any tendency for participants to anticipate sluicing from the initial form of the sentence. Also with this concern in mind, we varied the contents of the conjoined verb phrases, so that a verb could not be ruled out as an antecedent based on, for example, transitivity alone. To block participants from predicting sluicing antecedents based on initial verb phrase forms, we included 8 "sprouting" sluices, as in *studied* in *slept and studied* or as in *married* in *graduated and married*, 12 "merger" sluices verb phrases such as *drank coffee and typed a paper*, and 11 "merger" sluices with verb phrases containing *someone* or *something* as in *went home and texted someone*. It should be noted that the "sprouting" sluices are sensitive to the Coordinate Structure Constraint, making the "sprouting" sluices island-sensitive while the other "merger" sluices are not. This difference could potentially affect their well-formedness. To prevent participants from predicting the acceptability of a sentence based on the *wh*-element, we alternated *wh*-elements over items such that the acceptability of the sentence was not predictable based on the particular *wh*-item. The word *what* appeared in acceptable conditions in 12 items and in unacceptable conditions in 13 items. The word *which* appeared in acceptable conditions in 8 items and in unacceptable conditions in 8 items. The word *who* appeared in acceptable conditions in 11 items and in unacceptable conditions in 10 items. The full list of the materials used in Experiment can be found in Appendix A.

Norming—We collected pre-test data on our items, including the unacceptable versions, from 20 native speakers of American English from the New York University community. Participants were asked to rate, on a scale from 1 to 5, the degree to which each sentence was an acceptable or well-formed sentence of English. On this acceptability scale, we informed participants that a rating of 1 indicated that the sentence was "totally unacceptable," 2 indicated "somewhat unacceptable," 3 indicated that "neither horrible nor natural," 4 indicated "acceptable" and 5 indicated "perfectly acceptable." Participants saw all 31 items in each of the 8 acceptable conditions. We included 15 of the 31 items' Unacceptable conditions as foils; participants rated 368 sentences each, 248 of which were "acceptable" and 120 of which were "unacceptable." The following are the mean ratings for the Sluicing conditions: the Recent Double VP condition had a mean rating of 3.54 ($SD = .18$), the Distant Double VP condition had a mean rating of 3.01 ($SD = .26$), the Recent Single VP condition had a mean rating of 3.48 ($SD = .19$), and the Distant Single VP condition had a mean rating of 3.42 ($SD = .27$). The foil Unacceptable conditions had an average rating of 1.16 on the 5-point scale, while the average rating across Acceptable conditions was 3.35. The average rating for Acceptable Sluicing conditions was 3.36, while the average rating for Acceptable Control conditions was 3.33. A repeated measures ANOVA on the Sluicing conditions showed a main effect of Number of VPs and a main effect of Recency, as well as an interaction. Pairwise comparisons showed that the Distant Double VP condition was rated as .53 units less acceptable than the Recent Double VP condition (95% CI = −.69 to −.38), which was a reliable difference. The Distant Double VP condition was rated as .41 units less acceptable than the Distant Single VP condition (95% CI = −.59 to −.23). No other pairwise comparisons reached significance; the Distant Double VP condition thus seems to be the source of the interaction effect. In sum, the ratings indicated that Acceptable conditions were consistently rated as more acceptable relative to the Unacceptable conditions, and that the Acceptable conditions had reliable differences

from each other consonant with Frazier and Clifton (2005). Importantly, there were no differences between "merger" and "sprouting" sluices.

Procedure—In each of the four sessions, participants read 124 experimental sentences, four conditions per item (two Sluiced and two Control conditions), counterbalanced within and across sessions. Participants saw every item in every condition, but at different points in the experiment. Conditions were counter-balanced across sessions such that participants saw an equal number of items in each condition in each session, though the item used to represent that condition varied. In order to vary which item was used to represent a given condition in a session systematically, two conditions within an item were yoked together and presented in the same session. These pairs were then shuffled through the 31 items. Conditions 1a and 2b of a given item appeared together in the same session, as did conditions 1b and 3a, conditions 2a and 4b, and 3b and 4a of the same item. Conditions 5a and 6b of a given item appeared together in the same session, as did conditions 6a and 8b, conditions 7a and 5b, and 8a and 7b of the same item. Critical trials, including unelided controls, constituted 32% of each session, and were presented randomly among the remaining 68%, none of which was elided. The fillers were multi-clause sentences, with equal numbers of acceptable and unacceptable (underlined) versions: *The writer who had asked if the editor was sympathetic heard that the vice-president of marketing quit/ populated.*

Stimulus presentation, timing, and response collection were all carried out on a personal computer using software with millisecond timing. A trial began with a 500 ms fixation point presented at the center of the screen. Sentences were presented in a phrase-by-phrase controlled presentation manner, 335 ms per number of words in the phrase. A 50 ms, 1000 Hz tone served as the response signal. The first response signal occurred 300 ms before the onset of the sentence final phrase (which included the sluiced phrase in the experimental conditions). After the onset of the final phrase, 14 more response signals occurred, 350 ms apart, while the final phrase remained on the screen. The response signals continued until 5 seconds after the onset of the final phrase, for a total of 15 response signals. Participants were trained to synchronize their responses to the tones, responding within 200 milliseconds of each tone. They were instructed to simultaneously press both the "yes" and "no" keys as an initial (undecided) response, and then to select, and continue to press, only one of the two keys (with "yes" indicating a judgment that the sentence was acceptable, and "no" indicating that the sentence had been judged unacceptable) when information on the acceptability of the sentence became available. They were also encouraged to modulate their responses if their judgment changed during the trial.

Participants first completed a 45-minute practice session in order to familiarize themselves with the task. They were trained on pressing and switching responses rhythmically across the sampling period to ensure that they were practiced at modulating their responses, and training continued until they became comfortable with the response requirements and could make a response within 200 milliseconds. Between-trial intervals were participant controlled, and there were two mandatory breaks each session.

Data Analysis—Comprehension accuracy was calculated using a standard d' measure, $d' = z(hits) – z(false\ alarms)$, where a "hit" was an "acceptable" response to an acceptable sentence and a "false alarm" was an "acceptable" response to an unacceptable sentence. The d' scores provide a measure of the participant's ability to discriminate acceptable from unacceptable structures, uncontaminated by response biases.

A hierarchical model-testing scheme was used to determine whether conditions differed in asymptote (λ), rate (β), or intercept (δ) in Equation (1), constrained by differences in

empirical d'. Exponential model fits of the data ranged from a null model in which all functions were fit with a single asymptote, rate, and intercept parameter (a 1λ-1β-1δ fit) to a fully saturated (a 4λ-4β-4δ fit) model in which each condition was fit with a unique asymptote, rate, and intercept. For each participant and the averaged data, separate parameters were allotted to the different conditions if they systematically improved the fit of the SAT function to the observed d' data. The exponential function in Equation (1) was fit to the data with an iterative hill-climbing algorithm (Reed, 1976), which minimized the squared deviations of predicted values from observed data. Fit quality was assessed by an adjusted-R^2 statistic—the proportion of variance accounted for by the fit, adjusted by the number of free parameters (Judd & McClelland, 1989)—and by an evaluation of the consistency of the parameter patterns across the individual participant fits. Additionally, we performed inferential tests of significance computed over individual participants' d' data, used to constrain the model selected, and tests on the fitted parameter estimates for each of the candidate models, detailed in the Results section. We report 95% confidence intervals (CIs) around the mean difference for paired comparisons of interest.

Results and discussion

Figure 2 presents the average (across participants) d' values as a function of processing time, along with the best-fitting exponential model described below. Parameter values of the best-fitting model can be seen in Table 3. Inspection of Figure 2 suggests that antecedents in a conjunction were less accurately processed than Single VP antecedents, and that Recency modulated this effect. As an initial means of determining whether there were reliable differences in asymptotic performance as a function of distance or of interference, we averaged the d' values for each participant (and, for an item analysis, by each item) in each condition from 3.5 to 5.25 seconds post-initial response cue in order to derive an empirical estimate of asymptotic accuracy. A repeated measures ANOVA on these values for the elided conditions revealed an interaction between Number of antecedent positions and Recency, $F_1 (1,16) = 20.19$, $p < .001$, $F_2 (1, 30) = 6.06$, $p < .05$. Main effects of Number of antecedent positions and Recency were also found, $F_1 (1,16) = 13.89$, $p < .01$, $F_2 (1, 30) = 5.40$, $p < .05$ for Number of antecedent positions and $F_1 (1,16) = 19.50$, $p < .001$, $F_2 (1, 30) = 10.17$, $p < .01$ for Recency, respectively.

Pairwise comparisons indicated reliable differences mostly driven by comparison with the Distant Double VP condition. Notably, responses to the Distant Double VP condition were on average .79 d' units lower in accuracy than responses to the Recent Double VP condition (95% CI = −1.10 to −.47 d' units), which was a significant difference, and .98 d' units lower than those to the Distant Single VP condition (95% CI = −1.44 to −.52 d' units), which was also significant. Responses to the Recent Single VP condition were on average .27 d' units higher in accuracy than those to the Recent Double VP condition (95% CI = −. 05 to .57 d' units), and .08 d' units higher than those to the Distant Single VP condition (95% CI = −.14 to .29 d' units); neither of these differences reached significance.

Control conditions showed a different pattern. A repeated measures ANOVA on the unelided controls showed no effects of Number of antecedent positions or Recency on empirical d' accuracy. Pairwise comparisons revealed that accuracy for unelided sentences in the Recent Single VP condition was on average .18 d' units lower than in the Recent Double VP condition, which was not significant (95% CI = −.40 to .06 d' units), and was on average .35 d' units lower than the Distant Single VP condition, which was also not a significant difference (95% CI = −.83 to .14 d' units). Accuracy for unelided sentences in the Recent Double VP condition was on average .19 d' units lower than accuracy for the Recent Single VP condition, which was not a significant difference (95% CI = −.44 to .05 d' units). The Recent Single VP condition was on average .02 d' units higher than the accuracy for

Distant Single VP unelided sentences. This difference was also not significant (95% CI = −. 18 to .22 d' units).

Competitive fits of the exponential equation to the empirical data also yielded clear evidence that both Number of antecedent positions and Recency modulated asymptotic performance in the sluiced conditions: Models that did not allocate separate asymptotes for the Double VP versus Single VP antecedent conditions produced poor fits to the empirical SAT data, and they left systematic residuals. In fits of the average data, allocating separate asymptotes to each sluiced condition increased the adjusted-R^2 from .851, observed with a null 1λ-1β-1δ model to .992 with a 4λ-1β-1δ model. This model improved the quality of the fits of the individual participants' data, systematically increasing the adjusted-R^2 values over what was observed with a 1λ-1β-1δ model (ranging from .667 to .961 as compared to .564 to .906). In the fit of the average data to the 4λ-1β-1δ model, the asymptote for sentences with recent conjoined antecedents was estimated to be 2.71, the estimate for the sentences with distant conjoined antecedents was 1.89, the estimate for the sentences with recent Single VP antecedents was 3.23, and the estimate for the sentences with distant Single VP antecedents was 3.03. We performed a repeated measures ANOVA on the asymptote parameter estimates for the sluiced conditions from the fits of the individual participants' data from this model. An interaction between Number of antecedent positions and Recency was found, F_1 (1,16) = 18.13, $p < .01$. Main effects of Number of antecedent positions and Recency were also found, F_1 (1,16) = 46.37, $p < .001$, for Number of antecedent positions and F_1 (1,16) = 22.47, $p < .001$, for Recency, respectively. Pairwise comparisons of the across participants asymptote parameter estimates showed that the Recent Double VP condition was .84 d' units higher in accuracy than the Recent Double VP condition, which was significant (95% CI = .52 to 1.17 d' units), and .54 d' units lower in accuracy than the Recent Single VP condition, which was also significant (95% CI = −.77 to −.31 d' units). The Distant Single VP condition was .19 d' units lower in accuracy than the Recent Single VP condition, which was marginal (95% CI = −.43 to .04 d' units), and 1.19 d' units higher in accuracy than the Distant Double VP condition, which was significant (95% CI = .81 to 1.57 d' units).

The differences in asymptote parameter estimates indicate that antecedent verbs in conjunctions were less likely to be successfully retrieved and interpreted than single antecedent verbs. This difference suggests that the quality of the retrieved information was poorer for distant conjoined antecedents, perhaps leading to a less acceptable interpretation. It was also the case that Recency modulated this effect under conjunction, and resulted in lower likelihood of successful retrieval even in the absence of the second VP. If Number of antecedent positions or Recency also affected the speed of processing the sluice, then either factor should have engendered differences in either rate (β) or intercept (δ). Allocating separate rate or intercept parameters to different conditions only slightly improved adjusted-R^2. In fits of average data, a 4λ-4β-1δ model resulted in an adjusted-R^2 of .993 and a 4λ-1β-4δ model resulted in an adjusted-R^2 of .992, as compared to the .991 values observed with the simpler 4λ-1β-1δ model. Importantly, there were no consistent trends across participants in either the rate or intercept parameters when they were allowed to vary (e.g., models with dynamics parameters such as 4β-4δ, 4β-2δ, and various 3β-1δ and 2β-1δ models) and crucially t-tests on the parameter estimates were not significant. This is also true for empirical d' calculated with a common false alarm rate, constructed from the average of the foil conditions. Hence, there was no evidence to suggest that distance of interpolation, serial position within the conjunction, or processing additional VPs affected processing speed, and therefore, there was no evidence that antecedents were retrieved more slowly under any of these circumstances.

For completeness, we also compared the functions for control conditions without ellipses in the final region. As inspection of Panel 2 in Figure 2 suggests, there were no differences evident in the control conditions. Consequently, the best fit for these functions was a simple 1λ-1β-1δ model, adjusted-$R^2 = 0.992$. All t-tests on the parameter estimates for models that varied one of the SAT parameters were not significant. This suggests that the distance effect evident in the ellipsis conditions is related to the availability of the antecedent, not due to general differences between the recent antecedent and distant antecedent sentence forms.

The timecourse profile and pattern of comprehension accuracy for sluiced sentences was identical to what has been found for VPE and other dependencies (Martin & McElree, 2008, 2009; McElree, 2000; McElree et al., 2003): Recency affected the likelihood that an appropriate antecedent can be recovered from memory, thereby lowering asymptotic accuracy, but did not affect the speed with which an antecedent representation can be accessed. Whether the antecedent VP occurred in a recent position in the conjunction also affected the likelihood of successful interpretation, but not processing speed. Similarly, presence of an additional VP decreased likelihood of successful interpretation, but again did not affect processing speed. Crucially, there were no timecourse differences between Double and Single VP conditions. This pattern is inconsistent with both a forward and a backward search, even if the search were only over syntactically licensed positions. That there were no differences in processing time, and therefore no differences in retrieval speed, as a function of number of syntactically available antecedent positions indicates that not only can we exclude content-free searches, but we can also exclude searches that are guided by syntactic content.

This pattern is consistent with a content-addressable direct-access process, which enables representations of differing quality to be recovered with comparable speed (McElree, 2006). Notably, the interaction between Number of antecedent positions and Recency suggests that not only does recency benefit comprehension, but that sluicing from the most recent serial position is more likely to be successfully interpreted than when another verb intervenes, even if it is not a possible antecedent in the end. Because there were no differences in processing speed, it is unlikely that the mismatching VP in the conjunction is considered as an antecedent or that accessing an antecedent within a conjunction impacts processing speed. Nonetheless, some aspect of this information affects the likelihood of successful interpretation, perhaps due to decreases in cue specificity or distinctiveness. That is, retrieving the antecedent in the presence of another verb or in the presence of a conjunction decreased processing accuracy as a function of recency. A strong indication of this comes from the Distant Double VP condition having reliably lower accuracy than both the Distant Single VP condition and the Recent Double VP condition. As only accuracy was affected, it seems likely that recency and similarity-based interference from similar items in memory contribute to the processing asymmetry between sluicing out of a recent or distant conjunct.

We found that sluicing from the distant conjunct negatively impacted processing, in effect replicating Experiments 1a and 1b from Frazier and Clifton (2005), and also enabling us to exclude accounts of their data that entail differences in processing speed. However, as recency interacted with number of antecedent positions, our data indicated that these factors contribute to the processing asymmetry between sluicing from a recent or from a distant conjunct. Besides manipulating the recency, resolving sluicing with a conjunction in the sentence context may have other consequences that highlight the dependency resolution mechanism at work. There are two explanations of our findings. First, as previously mentioned, interference could be higher in the Double VP case than the Single VP case, lowering accuracy. Alternatively, accessing or retrieving one conjunct of a conjoined constituent in memory, as needed in the Double VP cases to select one of the VPs as an antecedent, could be costly as well. In combination with a distance effect, if accessing a

conjunction were costly, we would see the pattern of results we saw in Experiment 1. Interestingly, although one might predict that accessing one conjunct within a conjunction would affect processing speed, this prediction is not born out, as we did not find any speed differences as a function of either factor. Although we cannot yet know precisely how the conjuncts and the conjunction in Experiment 1 were encoded into memory, there are indications from the formal linguistics literature, in the form of the Coordinate Structure Constraint - a probable impetus for Frazier & Clifton (2005)'s CDH, and from the memory literature that suggest that the conjunction may be encoded and represented as a single unit. Miller (1956) coined the term "chunking" to refer to the psychological process of encoding of individual units of information in groups or chunks. Further evidence for the existence of this process comes from McNamara, Hardy, and Hirtle (1989), who found an advantage in retrieval speed for items in the same chunk, thought to be retrieved together (~50 ms). Johnson (1970) found that elements of a chunk tend to be retrieved together or not at all, pointing to their storage as a chunk of individual inputs. This latter finding suggests the possibility of processing ramifications from retrieving an item or cueing for a target that is part of a larger chunk, then accessing it in memory. We will call this process *dechunking*; retrieving only one part of a chunk or retrieving the whole chunk and selecting one item in it for further processing. If the accuracy effects found result from "dechunking" costs, then dechunking would appear only to impact accuracy. Nonetheless, the decrements in accuracy could reflect either interference *or* dechunking costs. If lower accuracy is a reflection of interference at retrieval, then comparing a case where the conjunction itself is available as an antecedent with a case where one position is selected could indicate the source of difficulty. Given that we did not find differences in processing speed, a more naturalistic reading task is sufficient for testing the dechunking hypothesis. We next tested the processing of sluicing under conjunction, where the conjunction is also available as an antecedent, using eye-movement recordings.

Experiment 2

Our SAT findings suggest that recency and serial position effects on reading time measures might reflect the quality of the antecedent representation in memory but that iterative retrieval of order information is not necessary. Specifically, as distance increases and recency decreases, the availability of the antecedent representation in memory may decrease, either because representations no longer benefit in strength from recency of onset, or because the processing of interpolated material interferes with the storage or cue-dependent retrieval of the antecedent. In reaction time measures, this would mean elevated reading times or more regressions when the antecedent verb is distant, in the first conjunct position, or under the presence of a second VP.

Experiment 1 found that sluicing from the more recent conjunct resulted in greater likelihood of successful comprehension, but a question still remains as to whether the difficulty observed comes from accessing a single position in an antecedent that is part of a conjunction or from dechunking and accessing a conjunction as an antecedent. This is especially relevant because the Recent Single VP condition did not show a reliable difference from the Distant Single VP in the pairwise comparisons from Experiment 1, although a main effect of Recency was found. Frazier and Clifton's work also raises this issue; their CDH asserts that some structural component negatively affects access to the antecedent and subsequent processing in these cases.

Experiment 2 examined eye-movement patterns during the reading of variants of the (acceptable) materials used in Experiment 1, plus two additional conditions designed to test the role of conjunction in a different way. These had two main purposes. First, we wished to explore how the observed SAT differences are expressed in more natural reading situations,

and to determine how our timecourse findings would align with more conventional markers of difficulty in sentence processing and during comprehension of sluicing. For example, longer fixations could indicate greater initial processing difficulty while more regressions could indicate difficulty in later stages of processing, possibly regarding the quality of the information needed to resolve the dependency. Second, we wanted to further examine our SAT effects using a case where the whole conjunction is available as an antecedent, and where a conjunction is recently processed but no single position within it is selected as the antecedent. If some aspect of dechunking and then accessing a conjoined antecedent– rather than some aspect of selecting just one position – is costly, then availability of the full conjunction as the antecedent should produce similar or greater difficulty relative to the Double VP cases in Experiment 1. To test this, we constructed the following versions of our previous stimuli, so-called "Conjunction Available" conditions:

6) Michael drank coffee and typed something, but he didn't tell me where.

7) Michael typed something and drank coffee, but he didn't tell me where.

where (6) corresponds to the Recent conditions from Experiment 1, which were also included in Experiment 2, and (7) corresponds to the Distant conditions. However, because we wanted the conjunction to be available as an antecedent, the interpretation of these conditions is rendered ambiguous, in that *where* can come from *drank coffee,* from *typed something,* or from *drank coffee and typed something/typed something and drank coffee*. This construction also obscures the distinction between "Recent" and "Distant," because if the conjunction is interpreted as the sluicing antecedent, it is equidistant in both conditions. If sluicing out of a conjunction is more difficult than from out of an unconjoined antecedent due to structural constraints or domain restrictions on the conjunction as suggested by Frazier and Clifton (2005), then any conjoined antecedent should entail processing difficulty and not be different from the case where one position within the conjunction is selected.

Alternatively, if what is costly is selecting one position in the face of retrieval interference, then the Recent and Distant Double VP cases should be more difficult than the Conjunction Available cases. As for the predictions for the conditions from Experiment 1, if longer distances and additional VPs to sluice from increase interference at retrieval, then we might expect longer reading times on the antecedent under sluicing with distant as compared to recent antecedents. This prediction follows from the intuitive idea that participants may often need to reprocess a distant antecedent because the antecedent's representation in memory is too poor or the cues at retrieval are not diagnostic enough to support interpretation when initially encountering the SES. Accompanying longer and more frequent rereading of the antecedent verb should be increases in measures of regression on or immediately after the SES.

Method

Participants—Forty native speakers of American English from the New York University community were paid to participate in the study. They were between the ages of 18 and 26 and all had normal or corrected-to-normal vision.

Materials—The materials were variants of 31 items (acceptable versions) used in Experiment 1, plus five additional items, for a total of 36, so that we could Latin-square the items over eight lists. Example materials are illustrated in Table 2. A spillover region was added to each sentence, such as *because he was still tired*. Fillers constituted 82% of the stimuli presented in the session and were multi-clause sentences like the fillers used in Experiment 1. The 36 items in eight conditions (including two conditions not applicable to this study, which made up approximately 1% of the sentences each participant read) were broken up into eight lists, such that participants in four of the lists saw 32 sentences from

this experiment during the session, while participants in the other four saw 33 sentences, but never saw an item in more than one condition within the session. As a means to tease apart the roles of serial position, distance, and conjunction, we created two additional variants for Experiment 2: two Conjunction Available conditions (corresponding to Recent and Distant conditions), where the conjunction of the verb phrases is available as an antecedent, see Table 2 for example materials. These variants were constructed with *drank coffee and typed something* in the Recent condition and *typed something and drank coffee* in the Distant condition, and such that the conjunction is also available as sluicing antecedents. We included 9 "sprouting" sluices with unsprouted antecedents, as in *studied* in *slept and studied* or as in *married* in *graduated and married*, 15 "merger" sluices with verb phrases such as *drank coffee and typed a paper*, and 12 "merger" sluices with *someone* or *something* as in *went home and texted someone*. The word *what* appeared as the *wh*-item in 13 items. The word *which* appeared as the *wh*-item in 11 items. The word *who* appeared as the *wh*-item in 12 items. The word *where* appeared as the *wh*-item in 17 items, while the word *why* appeared as the *wh*-item in 19 items. The full list of the materials used in Experiment can be found in Appendix B. The sentences were presented in an individually randomized order to each participant.

Procedure—We monitored the movements of the participants' right eyes during reading using a SensoriMotor Instruments Eyelink I head-mounted eye-tracker (SensoMotoric Instruments GmbH, Teltow, Germany), sampling at 500Hz. Eye cameras were positioned under each eye, recording eye movements and fixations every 4 msec. Screen resolution was set at 1600×1200 pixels. Sentences were presented in a fixed font, with each letter 18 pixels wide and 33 pixels high. No more than 80 characters were presented on one line of text. The stimuli appeared on a CRT monitor approximately 71 cm from the participant's eyes, where $1°$ of visual angle corresponded to 2.7 characters. A chin rest was used to reduce head movement. Yes/no-comprehension questions were asked for 50% of the items. After determining fixation, an automatic procedure pooled short contiguous fixations. This procedure combined all fixations shorter than 80 ms and within one character of each other. Fixations that were shorter than 80 ms but not contiguous were excluded from the analyses, because presumably little information is extracted (Rayner & Pollatsek, 1989). Fixations longer than 1000 ms were excluded as well.

Participants read a general explanation of the eye-tracking procedure. They were encouraged to read at a normal pace for understanding during the experiment. A calibration procedure was carried out at the beginning of the experiment, and recalibration was performed whenever the experimenter felt it necessary due to difficulty during inter-trial fixation validation. Before a new trial was presented, participants first looked at a fixation box in the middle of the screen, and a drift correction was performed. They were then presented with a fixation box coinciding with the position of the first letter of the upcoming sentence. This box served as a trigger, with the sentence being displayed only if the fixation was judged to be close enough to the center of the box. Participants read sentences at their own pace and pressed a button on a hand-held button box to make the sentence disappear.

Reading times were analyzed by region. We defined our regions of interest in the following way, using as an example the sentence in the Recent Double VP condition, *Michael drank coffee and typed something but he didn't tell me what, because he was tired*: (1) sluicing antecedent region (*typed*), (2) sluicing ellipsis site region (*what*), (3) first spillover region (*because*), and (4) second spillover region (*he was*). The following eye-movement dependent measures were used: (1) *first-pass reading* time, or the time spent fixating a region initially, excluding re-fixations, (2) *first-pass regressions-out*, or the proportion of trials in which at least one regression out of a region occurred during first reading of the region, (3) *second-pass reading time*, or the time spent in that region only after the eyes

have moved out of that region to the right, (4) *total time*, or the sum of all fixations within a region, (5) *regression path duration,* or the sum of all fixations in a region from first entering that region until moving to the right of that region, and (6) *regressions-in*, or the percentage of trials where one or more fixations in the region are preceded by a fixation in a region later in the sentence.

Results and discussion

All participants scored well on comprehension questions, with an average accuracy of 90%. We excluded trials on which track loss occurred (~2%). Reading times over 3 standard deviations from the mean (participant x condition mean) were excluded from the analysis. The reading time measures at the antecedent regions were analyzed with two repeated measures 2×2 ANOVAs. The first was Recency x Number of VPs (where Recency: Recent or Distant, Number of VPs: Double or Single VP) and the second was Recency x Conjunction Availability (where Conjunction Availability: Available or Unavailable conjunction as sluicing antecedent; Conjunction Unavailable conditions were the same as the Double VP conditions above) treating both participants (F_1) and sentences (F_2) as random-factors. These ANOVAs were followed up by pairwise comparisons to locate the source of effects. Tables 4 and 5 report these *F*-values on the sluicing antecedent, sluiced *wh*-item, and spillover regions for relevant effects. In the text, we report 95% confidence intervals around the mean differences between conditions in comparisons of interest. In order to analyze regression patterns, we performed repeated measures ANOVAs on first-pass regressions-out and regression path duration in all regions after the antecedent. We also performed these analyses on the regressions-in measure in all regions. Tables 4 and 5 also report the F-values of interest for these analyses. Table 6 reports the mean for each measure in each condition by region of interest.

Sluicing antecedent region—For the Recent and Distant conditions, as well as the Single VP conditions, the material in the antecedent differed in order (e.g., …*drank coffee and typed something* versus …*typed something and drank coffee; In the morning… typed something* versus …*typed something in the morning*). Thus, only measures on the sluicing verb can be compared between conditions. In the Conjunction Available conditions, the *wh*-item is not the same as in the Conjunction Unavailable conditions, and so cannot be directly compared at the SES.

There were no differences in first-pass times in the Recency x Number of VPs analysis. In second-pass times, there was a marginal main effect of Number of VPs by participants, which was significant by Items (see Table 4). There was no main effect of Recency by participants, but a marginal main effect of Recency by Items (see Table 4).

In this same analysis of total reading times on this region, there were main effects of both Number of VPs and Recency, but no interaction (see Table 4). To test whether total reading times on this region varied as a function of either factor, we compared Recent, Distant, Double VP, and Single VP conditions to each other as pairwise comparisons. The Recent, Double VP, and Single VP conditions differed from each other, respectively. The Recent Double VP condition was read on average 48 ms longer than the Recent Single VP condition (367 vs. 319, 95% CI = 9 to 88 ms), showing the effect of Number of VPs. The Distant Double VP condition was read on average 44 ms longer than the Recent Double VP condition (411 vs. 367, 95% CI = 1 to 87 ms), and the Distant Single VP condition was read on average 74 ms longer than the Recent Single VP condition (393 vs. 319, 95% CI = 34 to 113 ms), showing again that more distant antecedents were read longer even in a Single VP context.

To test for effects on regressions, which may reflect insufficient information for processing at the SES, we compared percentage of regressions into this region by condition. In the regressions-in measure on the sluicing antecedent region, there was a main effect of Recency (see Table 4). Pairwise comparisons showed that both the Double VP and Single VP conditions differed from each other, showing the impact of decreasing recency on the percentage of regressions-in. The Distant Double VP condition engendered on average 16 percent more regressions into this region than the Recent Double VP condition (31 vs. 15, 95% CI = 9 to 24). The Distant Single VP condition engendered on average 18 percent more regressions into this region from further along in the sentence than the Recent Single VP condition did (33 vs. 15, 95% CI = 10 to 27). This pattern suggests that distant antecedents required readers to regress back to them more often than recent antecedents did.

In the Recency x Conjunction Availability analysis on the sluicing antecedent region, there were no differences in the sluicing verb region in first- or second-pass times. However, in total reading times on this region, there was main effect of Recency and a marginal main effect of Conjunction Availability (see Table 5). Pairwise comparisons showed that both the Available and Unavailable conditions differed reliably from each other as a function of Recency. The difference between the Unavailable conditions was reported above in the Recency x Number of VPs analysis as the Double VP conditions. The Distant Available condition was read on average 55 ms longer than the Recent Available condition (385 vs. 330, 95% CI = 10 to 99 ms). It is important to note that the notion of Recency in the Conjunction Available conditions is obscure; we cannot in principle know which antecedent a reader has selected for interpretation, though it appears that the Distant condition is read for a longer period than the Recent one.

To test this region for effects on regressions, we conducted the Recency x Conjunction Availability analysis. The regressions-in measure showed a main effect of Recency in this region (see Table 5). Pairwise comparisons showed that both the Available and Unavailable conditions differed from each other, respectively. The difference between the Unavailable conditions is reported above in Recency x Number of VPs analysis. The Distant Available condition engendered on average 13 percent more regressions into this region from further along in the sentence than the Recent Available condition (31 vs. 18, 95% CI = 1 to 25). There were no other effects on any measure for this analysis in the sluicing antecedent region.

Sluicing ellipsis site (SES) region—Under the hypothesis that either factor may affect regressions, we compared first-pass regressions out of the SES. There was a main effect of Number of VPs in the Recency x Number of VPs (see Table 4). Pairwise comparisons also showed that Distant Double VP condition was regressed out of on average 5 percent more often than in the Distant Single VP condition (9 vs. 4, 95% CI = 1 to 9). There was also a marginal interaction in the regressions-in measure (see Table 4). There were no other effects in this region for any other measures in either the Recency x Number of VPs analysis or the Recency x Conjunction Availability analysis.

First spillover region—In the Recency x Number of VPs analysis, there was a main effect of Number of VPs on second-pass reading times (see Table 4), such that the Double VP conditions were read longer than the Single VP conditions. Pairwise comparisons revealed that the Recent conditions differed reliably from each other. The Recent Double VP condition was read on average 26 ms longer than the Recent Single VP condition (140 vs. 114, 95% CI = 3 to 48 ms).

To test for regression effects, we performed the Recency x Number of VPs analysis of first-pass regressions-out of the first spillover region, and there was an interaction (see Table 4).

Pairwise comparisons revealed that the Distant conditions were significantly different from each other; the Distant Double VP condition was regressed out of on average 7 percent more often than the Distant Single VP condition (10 vs. 3, 95% CI = 2 to 12). There were no other effects on any measure for this analysis. There were no effects under the Recency x Conjunction Availability analysis in any measure in the first spillover region.

Second spillover region—In the Recency x Number of VPs analysis of first-pass times, there was a main effect of Recency and a marginal main effect of Number of VPs that was significant by Items (see Table 4). Pairwise comparisons showed that none of the conditions differed reliably from one another.

Under this same analysis of second-pass times in this region, there was a main effect of Number of VPs and a main effect of Recency that was significant only by Items (see Table 3). Pairwise comparisons showed that the Distant conditions differed from each other, such that the Distant Single VP condition was read on average 39 ms longer than the Distant Double VP condition (160 vs. 121, 95% CI = 5 to 71 ms).

Also in the Recency x Number of VPs analysis, regression path durations showed a main effect of Recency and marginal main effect of Number of VPs that was significant by Items (see Table 4). Pairwise comparisons revealed that none of the conditions differed reliably from one another; there was a marginal difference between the Distant Double VP condition and the Recent Double VP condition, such that readers spent on average 24 ms more in this region before moving past it to the final region under the Distant Double VP condition than the Recent Double VP condition (254 vs. 230, 95% CI = −4 to 51 ms), although this difference did not reach significance. There were no other effects on any measure for the Recency x Number of VPs analysis in this region. There were no effects on any measures in the Recency x Conjunction Availability analysis in this region.

The pattern of main effects on the antecedent and interaction effects in later regions suggests that readers are launching regressions to the antecedent verb after first processing the SES, and doing so more often and for longer periods when distance and interfering information increases. The addition of conjoined VP competitor in an intervening serial position, the Double VP condition, engendered longer total reading times on the antecedent. This aligns with the negative impact on accuracy by serial position, as found in Experiment 1.

The Conjunction Availability conditions did not show the pattern of effects that would be predicted if the difficulty were to stem from accessing the conjunction - that is, the availability of the conjunction as antecedent, did not make processing more difficult. There was a non-significant pattern where Conjunction Available conditions had lower total reading times on the antecedent. It appears that the Conjunction Available conditions were neither reliably more difficult nor reliably easier than the Double VP conditions where one positions is selected. The so-called "Recent" Available condition was read faster and resulted in readers spending less time on the antecedent verb *typed* than they did in the "Distant" Available condition. Unfortunately, we are only able to compare reading times on *typed* within the antecedent, so it is not clear what is causing this difference. We speculate that perhaps there is some plausibility difference in the order of events between "drinking coffee and typing something" and "typing something and drinking coffee," although we note that this difference did not present itself or was masked by other effects in Experiment 1. There is no consistent pattern suggesting that the availability of the conjunction as an antecedent makes processing difficult. Another possibility is that *where* may be slightly biased towards which VP it seems to be moved out of, although it seems unlikely that there would be systematic differences in *where he drank coffee* versus *where he typed something*.

In summary, we found that readers spent more time rereading a distant antecedent verb than a recent one. This finding suggests that the quality of the retrieved representation of the antecedent is reduced when comprehenders process more material before the SES, including when there is another VP present. When the quality of the retrieved information is not sufficient to support interpretation, readers may have to regress back to the antecedent region to construct an acceptable interpretation, which was confirmed by the effects found in measures of regression at the spillover regions. This finding aligns with the demonstrably lower asymptotic performance for Distant as compared to Recent conditions in Experiment 1. Notably, the interaction between Recency and Number of antecedent positions that was expressed in asymptotic accuracy in Experiment 1 seems to be reflected in measures of regression on the first spillover region in Experiment 2. Comparing the conjunction available cases to the conjunction unavailable ones showed that there does not seem to be a cost of taking the whole conjunction as an antecedent, which arguably implies that dechunking the conjoined constituent in memory is not costly (nor accessing its structural domain), and thus is probably not what is costly in the Double VP cases. What we have observed a cost for is selecting one position as its item's recency decreases, especially in the presence of a competing position for the antecedent whose occupant's features overlap with those of the target.

General discussion

The results of both experiments suggest that distance between antecedent and retrieval site and the number of antecedent positions present affected the likelihood of successful comprehension, but did not affect the time it takes to retrieve and interpret the antecedent. Thus, it is likely that the quality of the antecedent representation and the diagnosticity of retrieval cues to the antecedent in memory are negatively impacted by these factors. Crucially, these factors do not affect the time needed to access the representation of the antecedent, which is inconsistent with all classes of search mechanisms during retrieval. These results exclude, for the first time, search mechanisms that use syntactic information as a search criterion or restrict the search space as a function of syntactic information.

These results are fully consistent with the hypothesis that sluicing is resolved with the same type of content-addressable process that has been argued to mediate the resolution of other types of nonadjacent dependencies (Martin & McElree 2008, 2009; McElree, 2000; McElree et al., 2003). What remains to be explored is *what* information or combination of cues is used during retrieval to elicit the antecedent directly. This question proves challenging in a morphologically impoverished, fixed word order language like English, where manipulating cues at retrieval or features of the antecedent often leads to ambiguity and hampers designs with true competitors. Nonetheless, there are two contributions of these results to cue-dependent direct-access accounts of dependency resolution in sentence processing. First, another type of dependency, importantly one at the intersection of ellipsis and *wh*-dependencies, has been shown to fit into the framework of content-addressable representations and direct-access retrieval. A key point in this approach to which our data speak regards *how* the antecedent is elicited from memory; sluicing antecedents, though prone to factors that generally encumber retrieval from memory, are elicited directly, by virtue of their content, not by a search guided by structural position information. Second, we have made steps towards uncovering the source of the processing complexity during sluicing found in the literature, essentially by appealing to the constructs of retrieval interference and serial position effects, both of which are supported by convergent evidence from other areas of cognitive psychology.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Anderson, MC.; Neely, JH. Interference and inhibition in memory retrieval. In: Bjork, EL.; Bjork, RA., editors. Handbook of perception and memory. Vol. 10. San Diego: Academic Press; 1996. p. 237-313.Memory

Bornkessel I, McElree B, Schlesewsky M, Friederici AD. Multidimensional contributions to garden path strength: Dissociating phrase structure from relational structure. Journal of Memory and Language. 2004; 51:495–522.

Chung S, Ladusaw W, McCloskey J. Sluicing and logical form. Natural Language Semantics. 1995; 3:239–282.

Clark SE, Gronlund SD. Global matching models of recognition memory: How the models match the data. Psychonomic Bulletin & Review. 1996; 3:37–60.

Dosher BA. Empirical approaches to information processing: Speed-accuracy tradeoff functions or reaction time. Acta Psychologica. 1979; 43:347–359.

Foraker S, McElree B. The role of prominence in pronoun resolution: Active versus passive representation. Journal of Memory and Language. 2007; 56:357–383.

Frazier, L. PhD dissertation. University of Connecticut; Storrs: 1978. On comprehending sentences: Syntactic parsing strategies.

Frazier L, Clifton CE Jr. The syntax-discourse divide: Processing ellipsis. Syntax. 2005; 8:121–174. [PubMed: 17356681]

Gibson E, Pearlmutter N, Canseco-Gonzalez E, Hickok G. Recency preferences in the human sentence processing mechanism. Cognition. 1996; 59:23–59. [PubMed: 8857470]

Gordon PC, Hendrick R, Johnson M. Memory interference during language processing. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2001; 27:1411–1423.

Gordon PC, Hendrick R, Johnson M. Effects of noun phrase type on sentence complexity. Journal of Memory and Language. 2004; 51:97–114.

Gordon PC, Hendrick R, Levine WH. Memory-load interference in syntactic processing. Psychological Science. 2002; 13:425–430. [PubMed: 12219808]

Gronlund SD, Edwards MB, Ohrt DD. Comparison of the retrieval of item versus spatial position information. Journal of Experimental Psychology: Learning, Memory, & Cognition. 1997; 23:1261–1274.

Johnson, NF. The role of chunking and organization in process recall. In: Bower, GH., editor. Psychology of language and motivation. Vol. 4. New York: Academic Press; 1970.

Judd, CM.; McClelland, GH. Data analysis: A model comparison approach. San Diego: Harcourt Brace Jovanovich; 1989.

Keppel G, Underwood BJ. Proactive inhibition in short-term retention of single items. Journal of Verbal Learning & Verbal Behavior. 1962; 1:153–161.

Kimball J. Seven principles of surface structure parsing in natural language. Cognition. 1973; 2:15–47.

Lewis RL. Interference in short-term memory: The magical number two (or three) in sentence processing. Journal of Psycholinguistic Research. 1996; 25:93–115. [PubMed: 8789368]

Lewis RL, Vasishth S. An activation-based model of sentence processing as skilled memory retrieval. Cognitive Science. 2005; 2:375–419.

Lewis RL, Vasishth S, Van Dyke J. Computational principles of working memory in sentence comprehension. Trends in Cognitive Science. 2006; 10:447–454.

Martin AE, McElree B. A content-addressable pointer underlies comprehension of verb-phrase ellipsis. Journal of Memory and Language. 2008; 58:879–906.

Martin AE, McElree B. Memory operations that support language comprehension: Evidence from verb-phrase ellipsis. Journal of Experimental Psychology: Learning, Memory, & Cognition. 2009; 35:1231–1239.

McElree B. The locus of lexical preference effects in sentence comprehension: A time-course analysis. Journal of Memory and Language. 1993; 32:536–571.

McElree B. Sentence comprehension is mediated by content-addressable memory structures. Journal of Psycholinguistic Research. 2000; 29:111–123. [PubMed: 10709178]

McElree B. Working memory and focal attention. Journal of Experimental Psychology: Learning, Memory & Cognition. 2001; 27:817–83.

McElree, B. Accessing recent events. In: Ross, BH., editor. The psychology of learning and motivation. Vol. 46. San Diego: Academic Press; 2006.

McElree B, Dosher BA. Serial position and set size in short-term memory: Time course of recognition. Journal of Experimental Psychology: General. 1989; 18:346–373.

McElree B, Dosher BA. Serial retrieval processes in the recovery of order information. Journal of Experimental Psychology: General. 1993; 122:291–315.

McElree B, Foraker S, Dyer L. Memory structures that subserve sentence comprehension. Journal of Memory and Language. 2003; 48:67–91.

McElree B, Nordlie J. Literal and figurative interpretations are computed in equal time. Psychonomic Bulletin & Review. 1999; 6:486–494. [PubMed: 12198788]

McElree B, Pylkkänen L, Pickering MJ, Traxler M. The time course of enriched composition. Psychonomic Bulletin & Review. 2006; 13:53–59. [PubMed: 16724768]

McNamara TP, Hardy JK, Hirtle SC. Subjective hierarchies in spatial memory. Journal of Experimental Psychology: Learning, Memory, & Cognition. 1989; 15:211–227.

Merchant, J. The Syntax of Silence: Sluicing, Islands and the Theory of Ellipsis. Oxford University Press; 2001.

Miller GA. The magical number seven plus or minus two: Some limits on our capacity for processing information. Psychological Review. 1956; 63:81–97. [PubMed: 13310704]

Nairne JS. The myth of the encoding-retrieval match. Memory. 2002; 10:389–395. [PubMed: 12396651]

Neath I. Distinctiveness and serial position effects in recognition. Memory & Cognition. 1993; 21:689–698.

Neath I, Knoedler A. Distinctiveness and serial position effects in recognition and sentence processing. Journal of Memory and Language. 1994; 33:776–795.

Öztekin I, McElree B. Retrieval dynamics of proactive interference: PI slows retrieval by eliminating fast assessments of familiarity. Journal of Memory and Language. 2007; 57:126–149.

Öztekin I, McElree B, Staresina BP, Davachi L. Working memory retrieval: Contributions of left prefrontal cortex, left posterior parietal cortex and hippocampus. Journal of Cognitive Neuroscience. 2008; 21:581–593.

Pearlmutter NJ, Gibson E. Recency in verb phrase attachment. Journal of Experimental Psychology: Learning, Memory, & Cognition. 2001; 27:574–590.

Rayner, K.; Pollatsek, A. The psychology of reading. New York: Prentice-Hall; 1989.

Reed AV. Speed-accuracy trade-off in recognition memory. Science. 1973; 181:574–576. [PubMed: 17777808]

Reed AV. The time course of recognition in human memory. Memory & Cognition. 1976; 4:16–30.

Ross, JR. 'Guess Who?'. In: Binnick, R.; Davison, A.; Green, G.; Morgan, J., editors. Papers from the 5th Regional Meeting of the Chicago Linguistics Society. 1969.

Sturt P, Scheepers C, Pickering M. Syntactic ambiguity resolution after initial misanalysis: The role of recency. Journal of Memory and Language. 2002; 46:371–390.

Van Dyke JA. Interference effects from grammatically unavailable constituents during sentence processing. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2007; 33:407–430.

Van Dyke JA, Lewis RL. Distinguishing effects of structure and decay on attachment and repair: A retrieval interference theory of recovery from misanalyzed ambiguities. Journal of Memory and Language. 2003; 49:285–316. [PubMed: 18079972]

Van Dyke JA, McElree B. Retrieval Interference in Sentence Comprehension. Journal of Memory and Language. 2006; 55:157–166. [PubMed: 18209744]

Watkins OC, Watkins MJ. Build-up of proactive inhibition as a cue overload effect. Journal of Experimental Psychology: Human Learning and Memory. 1975; 104:442–4152.

Waugh NC, Norman DA. Primary Memory. Psychological Review. 1965; 72:89–104. [PubMed: 14282677]

Wickelgren WA. Speed-accuracy tradeoff and information processing dynamics. Acta Psychologica. 1977; 41:67–85.
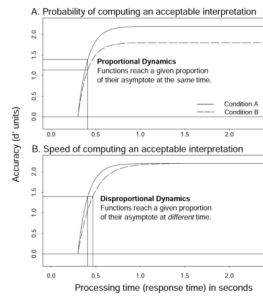
Figure 1.
Hypothetical SAT functions illustrating two conditions that differ by asymptote only (Panel A) or rate (Panel B). The intersection of the horizontal and vertical lines shows the point in time (abscissa) when the functions reach two-thirds of their respective asymptote (ordinate). When dynamics are proportional (Panel A), the functions reach the two-thirds point at the same time.
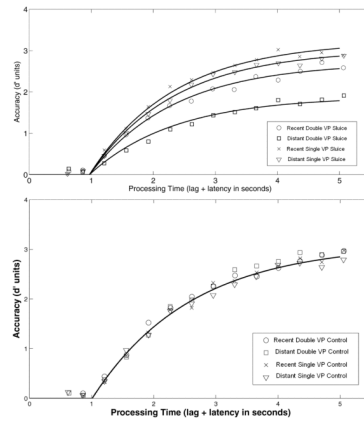
Figure 2.
Average *d′* accuracy (symbols) as a function of processing time (lag of the interruption cue plus latency to response) for Recent and Distant, Double VP and Single VP Sluiced conditions (Top Panel) and Recent and Distant Double VP and Single VP Control conditions (Bottom Panel) from Experiment 1. Smooth curves show the best-fitting exponential fit (see text).

Table 1

Example of materials used in Experiment 1

Recent Double VP, Sluicing, Acceptable

1a. Michael slept and studied, but he didn't tell me what.

Distant Double VP, Sluicing, Acceptable

2a. Michael studied and slept, but he didn't tell me what.

Recent Single VP, Sluicing, Acceptable

3a. In the morning, Michael studied but he didn't tell me what.

Distant Single VP, Sluicing, Acceptable

4a. Michael studied in the morning, but he didn't tell me what.

Recent Double VP, Control, Acceptable

5a. Michael slept and studied, but he didn't tell me much.

Distant Double VP, Control, Acceptable

6a. Michael studied and slept, but he didn't tell me much.

Recent Single VP, Control, Acceptable

7a. In the morning, Michael studied, but he didn't tell me much.

Distant Single VP, Control, Acceptable

8a. Michael studied in the morning, but he didn't tell me much.

Recent Double VP, Sluicing, Unacceptable

1b.* Michael slept and studied, but he didn't tell me which.

Distant Double VP, Sluicing, Unacceptable

2b.* Michael studied and slept, but he didn't tell me which.

Recent Single VP, Sluicing, Unacceptable

3b.* In the morning, Michael studied, but he didn't tell me which.

Distant Single VP, Sluicing, Unacceptable

4b.* Michael studied in the morning, but he didn't tell me which.

Recent Double VP, Control, Unacceptable

5b.* Michael slept and studied, but he didn't tell me must.

Distant Double VP, Control, Unacceptable

6b.* Michael studied and slept, but he didn't tell me must.

Recent Single VP, Control, Unacceptable

7b.* In the morning, Michael studied, but he didn't tell me must.

Distant Single VP, Control, Unacceptable

8b.* Michael studied in the morning, but he didn't tell me must.

Table 2

Example of Materials used in Experiment 2 with analysis regions

Region 3 – Sluicing antecedent verb region

Region 6 – Sluicing site region

Region 7 – Spillover region 1

Region 8 – Spillover region 2

Recent Double VP *Regions of interest are underlined*

1. Michael/ slept and/ studied/,/ but he didn't tell me/ what,/ because/ he was/ still tired./

Distant Double VP

2. Michael// studied/ and slept,/ but he didn't tell me/ what,/ because/ he was/ still tired./

Recent Single VP

3. In the morning,/ Michael/ studied/,/ but he didn't tell me/ what,/ because/ he was/ still tired. /

Distant Single VP

4. Michael// studied/ in the morning,/ but he didn't tell me/ what,/ because/ he was/ still tired. /

Recent Conjunction Available

5. Michael/ slept and/ studied/,/ but he didn't tell me/ where,/ because/ he was/ still tired./

Distant Conjunction Available

6. Michael// studied/ and slept,/ but he didn't tell me/ where,/ because/ he was/ still tired./

Table 3

Empirical *d′* and parameter estimates for Experiment 1

Condition	Average Empirical *d′* of Lags 14 & 15 across participants	4λ-1β-1δ parameter estimates from individual fits of participants' data	4λ-1β-1δ parameter estimates of the average empirical *d′*	1λ-1β-1δ parameter estimates of the average empirical *d′*
Recent Double VP Sluice	2.65	2.86	2.71	Asymptote λ: 2.86
Distant Double VP Sluice	1.86	2.02	1.89	Rate β: .59 β⁻¹: 1.69s
Recent Single VP Sluice	2.92	3.4	3.23	Intercept δ: .853
Distant Single VP Sluice	2.84	3.21	3.03	
		Rate β: .912 β⁻¹: 1.1s	Rate β: .727 β⁻¹: 1.37s	
		Intercept δ: 1.13	Intercept δ: .97	
Recent Double VP Control	2.85	x	x	Asymptote λ: 3.15
Distant Double VP Control	2.67	x	x	Rate β: .63 β⁻¹: 1.58s
Recent Single VP Control	3.04	x	x	Intercept δ: 1.02
Distant Single VP Control	3.02	x	x	

Table 4

Statistical analysis of eye-movement patterns in Experiment 2: Recency x Number of VPs

Measure/Region	Source	df	F_1	p<	df	F_2	p<
Sluicing Verb							
2nd pass (ms)	Recency	1, 16	.75	.40	1, 14	3.72	.07
	Number of VPs	1, 16	4.21	.06	1, 14	8.63	.01
Total time (ms)	Recency	1, 39	14.64	.001	1, 35	13.61	.001
	Number of VPs	1, 39	5.23	.05	1, 35	5.51	.05
Regressions-in (%)	Recency	1, 39	36.81	.001	1, 35	41.50	.001
SES Wh-item							
Total time (ms)	Recency	1, 39	10.45	.01	1, 35	5.45	.05
1st pass regressions-out (%)	Number of VPs	1, 39	5.20	.05	1, 35	1.99	.17
Regressions-in (%)	Recency x Num. VPs	1, 39	3.70	.06	1, 35	2.29	.14
Spillover 1							
2nd pass(ms)	Number of VPs	1, 19	6.36	.05	1, 23	5.21	.05
1st pass regressions-out (%)	Recency x Num. VPs	1, 39	5.71	.05	1, 35	2.16	.15
Spillover 2							
1st pass (ms)	Recency	1, 33	6.44	.05	1, 32	5.05	.05
	Number of VPs	1, 33	2.86	.1	1, 32	7.06	.05
2nd pass (ms)	Recency	1, 15	.04	.84	1, 13	7.30	.05
	Number of VPs	1, 15	4.40	.05	1, 13	.93	.35
Regression path duration (ms)	Recency	1, 33	5.21	.05	1, 32	3.63	.07

Table 5

Statistical analysis of eye-movement patterns in Experiment 2: Recency x Conjunction Availability

Measure	Source	df	F_1	$p <$	df	F_2	$p <$
Total time (ms)	Recency	1, 39	8.48	.01	1, 35	6.50	.05
	Conjunction Availability	1, 39	3.07	.09	1, 35	4.48	.05
Regressions-in (%)	Recency	1, 39	13.97	.001	1, 35	19.42	.001

Table 6

Means for Experiment 2 with mean standard error in parentheses

Measure/Condition	Sluicing Verb	SES	Spillover 1	Spillover 2
1st pass (ms)				
Recent Double VP	230(7)	201(9)	218(8)	229(8)
Distant Double VP	225(8)	211(7)	216(6)	248(10)
Recent Single VP	200(7)	206(6)	221(8)	246(11)
Distant Single VP	224(8)	227(13)	206(6)	268(11)
Recent Available	218(6)	223(9)	223(7)	242(8)
Distant Available	221(8)	215(8)	215(6)	239(8)
2nd pass (ms)				
Recent Double VP	137(9)	116(15)	140(8)	133(11)
Distant Double VP	131(9)	91(10)	133(8)	121(9)
Recent Single VP	119(10)	105(6)	114(10)	153(16)
Distant Single VP	112(6)	129(22)	131(6)	160(14)
Recent Available	112(6)	124(11)	118(5)	126(9)
Distant Available	114(9)	122(18)	131(7)	124(7)
Total time (ms)				
Recent Double VP	367(18)	320(10)	358(16)	389(24)
Distant Double VP	411(21)	321(10)	356(19)	420(24)
Recent Single VP	319(18)	313(9)	335(19)	399(19)
Distant Single VP	393(18)	347(13)	337(17)	428(22)
Recent Available	330(17)	344(11)	346(20)	408(28)
Distant Available	385(19)	327(11)	385(30)	442(29)
1st pass regressions-out (proportion)				
Recent Double VP	x	7(2)	4(1)	16(3)
Distant Double VP	x	9(2)	10(2)	17(3)
Recent Single VP	x	5(1)	5(2)	16(4)
Distant Single VP	x	4(2)	3(1)	13(3)
Recent Available	x	5(1)	8(2)	16(3)
Distant Available	x	8(1)	10(3)	21(4)
Regression path duration (ms)				
Recent Double VP	x	204(9)	215(7)	230(8)
Distant Double VP	x	212(7)	218(7)	254(12)
Recent Single VP	x	207(6)	218(8)	249(11)
Distant Single VP	x	225(12)	204(6)	266(10)
Recent Available	x	222(10)	223(7)	243(8)
Distant Available	x	214(8)	214(6	246(9)
Regressions-in (%)				
Recent Double VP	15(3)	2(1)	19(4)	14(3)
Distant Double VP	31(3)	7(2)	19(4)	11(3)
Recent Single VP	15(3)	6(2)	17(3)	9(2)

Measure/Condition	Sluicing Verb	SES	Spillover 1	Spillover 2
Distant Single VP	33(4)	4(1)	18(3)	12(3)
Recent Available	18(4)	9(2)	21(4)	15(3)
Distant Available	31(4)	6(3)	21(4)	15(3)

[*] Double VP conditions also serve as the Conjunction Unavailable conditions.