The Quarterly Journal of Experimental Psychology

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pqje20

Processing ambiguous Spanish *se* in a minimal chain

Enrique Meseguer [a], Carlos Acuña-Fariña [b] & Manuel Carreiras [c]

[a] Universidad de La Laguna, Tenerife, Spain

[b] Universidad de Santiago de Compostela, Santiago de Compostela, Spain

[c] Universidad de La Laguna, Tenerife, Spain

Available online: 27 Mar 2009

To cite this article: Enrique Meseguer, Carlos Acuña-Fariña & Manuel Carreiras (2009): Processing ambiguous Spanish *se* in a minimal chain, The Quarterly Journal of Experimental Psychology, 62:4, 766-788

To link to this article: http://dx.doi.org/10.1080/17470210802215228

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2009, 62 (4), 766–788

Psychology Press
Taylor & Francis Group

Processing ambiguous Spanish *se* in a minimal chain

Enrique Meseguer
Universidad de La Laguna, Tenerife, Spain

Carlos Acuña-Fariña
Universidad de Santiago de Compostela, Santiago de Compostela, Spain

Manuel Carreiras
Universidad de La Laguna, Tenerife, Spain

The recovery of pieces of information that are not linguistically expressed is a constant feature of the process of language comprehension. In the processing literature, such missing information is generally referred to as "gaps". Usually, one resolves gaps by finding "fillers" in either the sentence or the context. For instance, in *Peter seemed to be upset*, *Peter* is really the subject of *being upset* but appears as surface subject of *seems*. Sometimes constituents move, leaving gaps behind. Various Romance languages such as Spanish or Italian have a grammatical particle *se*/*si*, which, as it is extremely ambiguous, licenses different sorts of gaps. In Spanish, *se* can encode at least reflexive, impersonal, and passive meanings. In an eye-tracking experiment we contrast reflexive structures containing postverbal subjects with impersonal structures with no subjects (GAP *se vendó apresuradamente el corredor*/"the runner bandaged himself hurriedly" vs. GAP *se vendó apresuradamente al corridor*/"*(someone)* bandaged the runner hurriedly"). In a second manipulation we contrast the presence of an extra argument with *se*-passives (GAP *se vendó el tobillo el corredor*/"the runner bandaged his ankle" vs. GAP *se vendó el tobillo al corridor*/"the runner's ankle was bandaged"). Our comparisons involve contrasting standard transitive structures with nonstandard word order (postverbal subject and a preverbal subject gap) against inherently complex and less habitual structures such as impersonals (with no subject) or *se*-passives (with subjects in canonical object position). We evaluate the *minimal chain principle* (de Vincenzi, 1991), according to which displacement is costly because it entails complex (derivational) "chains" that must be undone before phrasal packaging can commence. We show the minimal chain principle to be essentially correct when contrasting more complex but more frequent structures with less complex but less frequent structures. A noteworthy feature of this research is that the gaps appear before the fillers in the structures that we analyse.

Keywords: Eye tracking; Minimal chain principle; Ambiguous Spanish *se*; Movement; Impersonals; Reflex-passives; Reflexives.

Understanding language involves the constant recovery of missing material—that is, of constituents of the coded predications that are not actually pronounced or written but are nevertheless understood to be latent. For instance, in *John hoped at the time to be able to go to London with all his*

Correspondence should be addressed to Enrique Meseguer, Departamento de Psicología Cognitiva, Campus de Guajara, Universidad de La Laguna, Tenerife 38205, Spain. E-mail: emesegue@ull.es

Preparation of this article was supported by Grants SEJ2004-07680-C02-02/PSIC, SEJ2006-09238/PSIC, and SEJ2005-08911 from the Spanish Ministry of Science and Education and Grant PGIDT01PXI20401PR from the Autonomous Government of Galicia.

http://www.psypress.com/qjep
DOI:10.1080/17470210802215228

family, the infinitive "to go" has no expressed subject, but language users still know *who* is able to go to London. Usually, missing elements are recovered by either looking back into the context or looking at other parts of the sentence. In the previous example, the missing subject of the infinitival clause is recovered by looking back to the subject of the main clause: *John*. So, *John* is the agent of both the "hoping" predication and the "being able" predication, but since the two predications share the same agent, it is only coded once (in the main clause). By contrast, in *eating foreign food may be dangerous*, the missing subject of "eating" must be recovered pragmatically (using knowledge of the world, etc.), so this "gap" can only be resolved by looking outside the sentence. Since these "gaps" are the rule, rather than the exception, languages typically invest heavily in their regulation, especially if they are to be resolved inside the sentence. The movement or displacement of sentence constituents usually brings about momentary gaps. For instance, in *that girl I wouldn't be caught dead going out with*, we understand "that girl" to be the complement of the preposition "with" (i.e., *going out with that girl*), but that complement has been displaced from its canonical position and has therefore left a gap behind (right after "with"). In Romance languages, such movement operations are extremely common, and they typically involve the most salient argument in predications: the subject. Thus, for instance, in the Spanish sentence *creo que mañana por la mañana va a venir Juan* ("I think Juan is coming tomorrow morning"), the subject of the subordinate clause, "Juan", is moved to the very end of the sentence, which means that the subordinate clause starts with an empty subject position. Usually such movements are driven by information structure necessities. In the case at hand, Spanish language users would interpret "Juan" to be new information (Lambrecht, 1994). Here we focus on these sorts of gaps.

One obvious way of approaching a psycholinguistic study of structures involving subject gaps that are not to be resolved contextually consists in measuring the processing cost of the displacement of the missing subject inside the sentence. Independently of how one captures movement (or its negation) in a linguistically formal way (e.g., Chomsky, 1981, 1986; Gazdar, Klein, Pullum, & Sag, 1985) one must assume that, in reference to a hypothesized subject–verb–object (SVO) word order, when something moves it leaves some sort of gap behind. From a psycholinguistic perspective, therefore, movement implies at least one form of filler-gap dependency. Formal theories of parsing presuppose that there are mechanisms for identifying, representing, and resolving such dependencies. Two well-known gap-filling theories are the *active filler hypothesis* (Clifton & Frazier, 1989; Frazier, 1987; Frazier & Flores d'Arcais, 1989) and the *minimal chain principle* (de Vincenzi, 1991, 1998). Here we focus on the latter.

Using a *government & binding* approach to grammar (Chomsky, 1981, 1986), De Vincenzi (1991) postulates her minimal chain principle as follows:

> Minimal Chain Principle: Avoid postulating unnecessary chain members at S-structure, but do not delay required chain members.

Based on Rizzi (1990), she defines a chain as a "set of elements nondistinct in indices, bearing one thematic role and one grammatical case… ". Informally, a chain is an anaphoric connection between two or more positions in syntactic trees, a sort of discontinuous constituent with a unitary thematic role (agent, patient, etc.) and a unitary function (subject, object, etc.). One particular prediction of the minimal chain principle is that the parser prefers to postulate a singleton chain to postulating multimember chains. For instance, it predicts that in the ambiguous Italian construction with *pro* (the abstract, missing subject of tensed clauses) in (1):

1. Ha chiamato Gianni
   a. pro has called Gianni "he/she/it has called Gianni"
   b. pro_i has called $Gianni_i$ "Gianni has called"

the parser will opt for (a) because in (b), assuming that Italian is a subject–verb–object language, the movement of Gianni from the canonical preverbal subject position to the displaced position after the verb must be undone. More specifically, notice

that (1a)–(1b) involve choosing between two different empty *pro* subjects: a context-linked full *pro* with specific case and thematic role, like any other explicit pronoun (*he*, *she*, etc), which therefore forms a singleton chain by virtue of the fact that it is in a position where it can directly receive such case and thematic role (say, *Mary*$_{i}$ *got the phone; pro*$_{i\ (she)}$ *has called Gianni*); and an expletive (dummy) *pro*, which forms a chain with the inverted subject to which it transmits case and thematic role (hence the shared index and the complexity of the chain). The first kind of *pro* gap amounts to an instruction saying, "subject missing because you know who I am talking about". The second, expletive *pro* entails a different instruction, "subject missing but it is a new entity: Wait for it until after the verb". In the government & binding terms that de Vincenzi uses, the first *pro* has content (like other pronouns: *he*, *she*, etc.); the second is just a place-holder for the moved subject (like the first "there" in *there is a book there*). De Vincenzi's theory may give way to research agendas that are too dependent on the specifics of specific linguistic theories. Here we are concerned with the simple notion that displacement of an entity (such as *Gianni*) leaves a gap in the vacated position and thus creates the need for the gap and the displaced entity to be co-indexed somehow (Clifton & Frazier, 1989). At this level of description, when an element moves it forms a complex chain with its gap (as in 1b). If it does not move overtly, the deep-to-surface mapping is direct and therefore not complex. An anonymous reviewer points out that it is not clear whether de Vincenzi intended the minimal chain principle to apply to anaphoric chains without movement as well as to obvious cases of movement and that she probably intended it to apply to both. Here we take the notion of a complex chain to apply when there is obvious overt movement of the subject in the surface structure of the sentence, as in (1b), where *Gianni* appears postverbally. This notion of movement is therefore tangible and less theory driven.

As de Vicenzi has commented, the one-member chain preference actually "amounts to saying that the parser prefers to analyze an element as being in its deep-structure position, that is, in the position where it directly receives a thematic role" assigned by the verb (de Vincenzi, 1991, p. 14), as *Gianni* is in (1a) above. In other words, the parser has "a general preference to posit unmoved elements over moved ones" (p. 124). This is a suggestion that de Vincenzi herself traces back to Fodor's (1979) "superstrategy". The motivation for such a strategy is computational economy: Only after the parser has been able to sort out the thematic and case features of a constituent (that is, what kind of role it has in the predication) can it structure that constituent in a phrasal package and send it to the semantic processor, instead of holding it unstructured. In multimember chains the parser has to link all the members of the chain (undo movement) before phrasal packaging can start.

De Vincenzi's (1991) comments on the difficulty of processing "complex derivations" are to be seen also in the same context where others such as Hemforth (1993), or Featherston, Gross, Munte, and Clahsen (2000) have referred to the cognitive cost of derivations, and to current linguistic theorizing by Chomsky (1995) that sees movement as an extremely costly operation also in terms of grammar. According to Featherston et al. (2000), for instance, their event-related potentials study of noun phrase trace and PRO (the missing subject of infinitivals and gerunds, as in *she*$_{i}$ *would like PRO*$_{i}$ *to come*) provides electrophysiological evidence that trace-filler coindexing is costly for the mind. Using *raising* constructions such as *the sheriff*^{i} *seemed* (t^{i}) *to be able to sentence the offender* (where *the sheriff* has moved to matrix subject position from the subordinate clause subject position, leaving a trace behind) and *control* constructions like *the sheriff*^{i} *hoped PRO*^{i} *to be able to sentence the offender* (where *the sheriff* is also the true subject of *hope* and therefore has not moved there from anywhere), they found that their raising constructions elicited a significantly stronger P600 effect than their control constructions. In line with previous event-related potentials results that show that the amplitude of the P600 reflects the cost of syntactic processing (Osterhout, 1994),

they concluded that the extra processing cost of *noun phrase trace* is due to the fact that it requires "an extra computational operation which is not required in control constructions" (Featherston et al., 2000, p. 153): what we have been defining here as a chain. This line of reasoning sits well with de Vincenzi's (1991) formulation of the minimal chain principle (notice that the German effects apply even in the absence of the overt surface movement that is evident in the Italian construction in (1b), so they involve "reconstruction" of a more fine-grained nature). Using event-related potentials and functional magnetic resonance imaging, Fiebach, Schlesewsky, and Friederici (2001, 2002) also provide evidence that maintaining *wh*-fillers ("unintegrated syntactic information") in working memory delays processing, but that the strain is different in cases of object *wh*-fillers (as opposed to subject fillers) despite unambiguous case marking. Several German studies have also served to make the point that the rationale of the minimal chain principle might not simply be a heuristic that grants rapid access to argument structure. Since an advantage of subject–object (SO) word order over OS order has been documented in verb-last German sentences, the minimal chain principle might also reflect a structural parsing preference that works independently of verb information (Frisch, Schlesewsky, Saddy, & Alpermann, 2002; see also Frazier & Flores d'Arçais, 1989, and Lamers, 2001, on Dutch; and Sekerina, 1997, 2003, on Russian). Most of these studies capitalize on case and/or agreement to resolve ambiguities that exist in strings where, keeping the position of nominal phrases constant, case alone is often not enough to avoid ambiguity because of morphological syncretism. This means that the word order advantage of the subject readings over the object readings is measured against the tendency of the language at large (a NOM-marked + ACC-marked configuration is usually read faster than an ACC-marked + NOM-marked one, where NOM = nominative, and ACC = accusative). In the *Gianni* case mentioned above, as well as in the sentences used in our manipulations (see below), preference for one word order or another is measured by noticing the effects of moving a subject constituent and leaving a gap behind. So it is overt movement, rather than overt case, that we intend to study here.

We proceed now to a description of the structures under analysis here. Various Romance languages such as Spanish or Italian have a grammatical particle *se*/*si*, which can be used to form a wide variety of structures (see Cinque, 1988, 1995, for Italian, and Sánchez López, 2002, for Spanish). A reflexive pronoun, as in (2) and (3), Spanish *se* is also used in impersonals such as (4) and *reflex* passives such as (5). (In the following sentences, DO = direct object; V = verb; IO = indirect object.)

2. Two-argument reflexive:
   Se vendó apresuradamente *el* corredor
   DO V ADJUNCT SUBJECT
   "Himself bandaged hurriedly the runner".
   The runner bandaged himself hurriedly.
3. Three-argument reflexive:
   Se vendó el tobillo *el* corredor.
   IO V DO SUBJECT
   "To himself bandaged the ankle the runner".
   The runner bandaged his ankle.
4. Impersonal:
   Se vendó apresuradamente *al* corredor.
   Clitic V ADJUNCT DO
   "Bandaged hurriedly to-the runner".
   The runner was bandaged hurriedly.
5. Passive:
   Se vendó el tobillo *al* corredor.
   Clitic V SUBJECT IO
   "Was bandaged the ankle to-the runner".
   The runner's ankle was bandaged.

When the reflexive structures have displaced postverbal subjects as in (2) and (3), they differ from impersonals and passives only minimally. (Compare 2–4 and 3–5.) The difference between the two reflexive structures is that (1) is made up of a direct object reflexive pronoun (*se*), the verb (*vendó*), an adjunct (*apresuradamente*), and a displaced subject (*el corredor*), while (3), which has no adjunct after the verb, has an extra argument (*el tobillo*) that takes on the direct object function and forces the reflexive pronoun

to be reanalysed as an indirect object. As the English translation shows, English does not have the corresponding three-argument reflexive. As for the difference between the impersonal and the passive, the impersonal, which has an optional adjunct (*apresuradamente*), has no subject, while in the passive the postverbal noun phrase *el tobillo* is the subject that agrees with the verb (*se vendó el tobillo al corredor* vs. *se vendaron los tobillos al corredor*). As the examples show, the only difference between the two-argument reflexive and the impersonal is the segment *el* (nonaccusative) versus *al* (accusative). Likewise, the only difference between the three-argument reflexive and the passive is also *el* versus *al*. This is the disambiguating region for both pairs of structures. Note that *al* is a contraction of preposition *a* ("to") and the definite article *el* ("the"). A phrase headed by *al* can therefore never be a subject phrase: It can only be a direct or an indirect object phrase. Importantly also, a phrase headed by *el* cannot be an object in the passive sentences. If we discard nonprimary constituents first, after "se vendó el tobillo" a phrase headed by "el" can only be a subject, the reason being that the previous "el" phrase (*el tobillo*) makes it impossible both to have two subject noun phrases (NPs) or a passive subject (*el tobillo*) followed by an object phrase headed by "el" (one cannot have "se vendó el tobillo el corredor" with *el corredor* being the recipient of the action: That would require "al"). This means that *el corredor* is as unambiguous as *al corredor* once it occurs after *el tobillo*. This is not the case when there is only one "el" phrase, since this is, in principle, ambiguous between nominative and accusative forms.

All the four structures under analysis (the two-argument reflexive and the impersonal on the one hand, and the three-argument reflexive and the passive on the other) are declaratives that start out without any subject noun phrase before the verb (as *se* is a direct or indirect object in the reflexive structures and a verbal clitic in the passive and the impersonal). Assuming a basic subject–verb–object order, the fact that a subject must be recovered implies that the preverbal position has some sort of gap that must be filled. Notice that the verb is tensed, so the latent presence of a subject is justified.

A crucial feature of our *se* examples is that their subject gaps are found before any phrase can be evaluated as a possible filler. This means that the form of gap filling that we are examining is not filler driven. Indeed, most gap-filling experiments have been carried out with English materials and contain structures involving *wh*-traces, where the *wh*-word appears first (this applies to the Fiebach et al., 2001, 2002, data as well). This makes these experiments more amenable to an approach in line with the active filler hypothesis (Frazier, 1987; Frazier & Flores d'Arcais, 1989). For instance, in *which book*$_i$ *did John read*—$_i$ *that summer at junior camp*? the filler *which book* announces the imminent presence of a gap. This is indeed a stable property of *wh*-words: to announce that something is not in place (movement). The active filler hypothesis is thus designed to account for structures whose fillers are clearly recognizable as such before their gaps show up, which means that gaps are easy to *predict* (Osterhout & Swinney, 1993). This is not the case in our structures. A major concern of the present study, therefore, is to study how the parser copes with the filling of a gap that cannot be so predicted.

In sum, the basic aim of this research is to study the minimal chain principle in Spanish, since there is no comparable study in this language. However, even similar languages like Italian and Spanish differ in complex and subtle ways. This means that the application of general strategies like the minimal chain principle may interact with the specifics of each new structure in unexpected ways. For instance, the Italian *Gianni* structure where a postverbal NP can be a direct object or a displaced subject is not possible in Spanish because animate direct objects in Spanish require a preposition, while subjects, displaced or not, obviously do not (*pro ha llamado **a** Juan*/*pro has called (**to**) Juan*). In the same way, the very complex reflex passive of Spanish in (5) contains a postverbal inanimate NP that may momentarily be a passive subject or a typical active direct object in situ, a fact with potentially interesting consequences for a fuller

evaluation of the minimal chain principle. In short, in the light of the peculiarities of the structures in the present analysis, the following specific objectives are pursued here:

The first objective is to measure the processing cost of displaced material in very simple structures such as the two-argument reflexive in (2) above, which is compared with the impersonal in (4), from which it differs minimally (*el* vs. *al*). Just like Italian, Spanish is a pro-drop language. This means that the simplest reading is in principle one in which the subject has been dropped (and is therefore discourse linked), thereby leading to a minimal chain in de Vincenzi's (1991) formulation. Note that there is evidence from a recent study on Turkish to suggest that pro-drop structures are highly acceptable even in the absence of a specific, supporting context (see especially Demiral, Schlesewsky, & Bornkessel-Schlesewsky, 2008, but also Carminati, 2005). Like the *Gianni* example, the reflexive construction has a displaced subject, whereas the impersonal sentence has no subject but nothing is displaced. The reflexive structure therefore involves the same kind of expletive *pro* as that found in the *Gianni* example (1b). Since, at least in the intuitive conception of chains invoked here, it involves a complex chain, the prediction is that it should be slower to process, according to the minimal chain principle. Alternatively, the linguistic complexity of a type of structure (the impersonal) where the absence of a subject (explicit or implicit) has been strangely grammaticalized might make this structure harder to deal with. It is interesting that English has no corresponding construction, as in this non-pro-drop language every thematically empty subject position must be announced by the insertion of expletives like *there* or *it* (as in *there is a man at the door*, where *there* does not mean anything, or in *it is essential that you be here in time*, where *it* has the sole function of anticipating the extraposed subject clause *that you be here in time*).

Our second objective involves a comparison between the three-argument reflexive in (3) and the passive in (5). It consists in measuring the cost of complex chain coindexing in structures where there is more primary material (Frazier & Clifton, 1996) intervening between the gap and the filler, thus making the movement operation apply over a segment with more stringent integration/storage needs. Integration costs in connection with storage needs have often been emphasized. For instance, Gibson's *syntactic prediction locality theory* (Gibson, 1998) uses both an integration cost component and a memory cost component. Integration cost is hypothesized to increase with the distance between the constituents to be integrated, with distance being measured in terms of intervening discourse referents. The passive structure in (5) has an extra discourse referent as compared to the impersonal in (4). Likewise, the three-argument reflexive in (3) has an extra discourse referent as compared to the two-argument reflexive in (2). Given the minimal chain principle ("avoid movement") and integration/storage cost logic ("avoid intervening referents"), the prediction is that (a) the passive will be easier than the three-argument reflexive because the latter involves movement, and (b) the effects of undoing movement in the three-argument reflexive should be greater than those in the two-argument reflexive because of the increased integration cost due to the extra referent. Given the privileged role of subjects in predications, it is important not to lose sight of the fact that integration and memory costs are at work while looking for a subject constituent in particular (i.e., the unique external argument), not for any kind of constituent.

Notice that this series of predictions rests on two peculiarities of our structures. The first is that since we are dealing with gap-before-filler parsing, integration cost involves storing a gap in memory (instead of a filler) while the left-to-right gap-filling process unfolds. The second has to do with movement itself. Recall that, at least under the interpretation of the minimal chain principle adopted here, both reflexives involve movement because, as in the *Gianni* example, the gap in the subject position is in a complex chain with material yet to come (*el corredor*/"the runner"). By contrast, impersonals like (4) are simply subjectless by definition, which means that there is no overt movement in them. As for passives, *government & binding* theory assumes that an empty slot (e)

occupies the subject position in these. In English passives this (e) position is filled by the movement of the active object (see below). Irrespective of the linguistic technicalities, however, since in Spanish *se*-passives the active object remains postverbal, there is no overt movement in these, either. In sum, the two reflexives should be more problematic than both passives and impersonals due to movement in the former. Additionally, the three-argument reflexive should prove the most difficult due to movement along a more demanding integration path. These predictions hold if the minimal chain principle and integration cost hold: If there is movement there is extra difficulty, and if there are more obstacles along the movement path, then there is even more difficulty.

Our third objective involves subjecting the minimal chain principle to a tighter test than the *Gianni* structure tested by de Vincenzi (1991). Notice that assuming a discourse linked *pro* as a default strategy in pro-drop languages in general entails reprocessing in both the three-argument reflexive and the reflex passive. Thus, in the former, as soon as the subject phrase *el corredor* appears, the pro-drop reading must of course be abandoned as sentences cannot have two subjects. In the latter, as soon as the indirect object phrase appears (*al corredor*), the discourse-linked interpretation whereby somebody bandages himself must be discarded too (?? "somebody bandaged himself (to) the runner"). Now, since reprocessing is mandatory, what are the predictions we can make? Notice that *el tobillo* is a temporarily ambiguous phrase: In the reflexive structure it occupies a direct object position after the verb in reference to a hypothetical underlying subject–verb–object template. This means that it is in a simple object chain and should therefore be invisible to the search for a subject filler. In the passive, however, *el tobillo* is a subject, not an object, but there is an interesting twist—that subject function is not signalled by an overt change of position since in *se*-passives subjects are normally postverbal (80% of the time, according to Barrenechea & Manacorda de Rossetti, 1977; see also Sepúlveda Barrios, 1988; intuition clearly confirms this). That postverbal noun phrase is now the segment of interest, and the question arises as to whether the parser is sensitive to its functional shift. In English-style periphrastic passives the functional shift from object to subject is accompanied by a clear configurational shift (i.e., movement), as the postverbal noun phrase object in the active becomes the preverbal noun phrase subject in the passive:

6. I killed **the tiger** ACTIVE TRANSITIVE
7. **The tiger** was killed (by me) PASSIVE

Notice that the noun phrase cannot remain postverbally: **was killed the tiger by me*. A configurationally defined principle like the minimal chain principle is able to handle that movement directly through the notion of a complex chain: Movement creates a complex chain, and chain complexity delays processing (e.g., noun phrase trace; Featherston et al., 2000). But, crucially, there is no such configurational shift in Spanish *se*-passives:

8. Juan se **vendó los tobillos** (*John bandaged his ankles*) ACTIVE TRANSITIVE
9. Se (le) **vendaron los tobillos** a Juan (*John's ankles were bandaged*) PASSIVE

Thus, the contrast between the three-argument reflexive and the reflex passive allows us to see whether the parser finds it more difficult to deal with a reflexive structure that simply has a displaced subject, or with a passive structure where nothing is displaced in reference to the deep order, but where the surface subject is in a deep object position. If the passive were to prove easier to process, that would mean that the parser is able to "see" its perfect deep SVO order and ignore the surface function of the first postverbal phrase. It would also mean that the apparently simple overt movement operation in the reflexive would be problematic simply because it violates the deep word order preference. Although cross-linguistic comparisons are difficult, it might be relevant to remember that in a series of experiments looking at phrase structure revision and revision of case making in German, Bornkessel, McElree, Schlesewsky, and Friederici (2004) have recently shown that these two processes are associated with two different event-related potential (ERP) components: Structure revision was associated with a P600 component while

revision of case marking was associated with an N400 component. Their ERP and behavioural data might be taken to indicate that structural reanalysis, and especially reconstruction of expected underlying word order, takes a greater toll than functional reanalysis (as signalled by case markers), even in a morphologically rich language like German.

In sum, our predictions are in line with the philosophy of the minimal chain principle in the sense that we expect the two reflexive structures to incur a processing cost due specifically to the movement of the displaced subject in both. They are also in line with the logic of integration cost in that we also expect movement to be made more difficult if it takes place along a path filled with referential competition. However, our structures also allow us to examine filler-gap coindexing in a language where functional roles and configurationality tend to diverge much more than in English, and where overt configurationality cannot therefore be relied upon as a safe clue to the establishment of the correct thematic relations among the various constituents of the sentence. For these structures (*se*-passives), it will be interesting to see whether easy-to-recognize overt movement causes more problems than hard-to-recognize deep-to-surface functional mismatches.

On the other hand, there is a particular type of model that may predict that the opposite should happen. Recently, some theories have proposed that parsing is highly influenced by the frequency with which a given structure occurs in the language (e.g., constraint-based models: MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell & Tanenhaus,1994; the tuning hypothesis: Mitchell, Cuetos, Corley, & Brysbaert, 1995). In general, these models assert, although with important differences between them, that resolution of structural ambiguities will be determined by the frequencies with which alternative disambiguations occur in the language. Thus, if it turns out to be true that reflexive structures are more frequent than passive structures, these models would make the opposite prediction to that of the Minimal Chain Principle hypothesis. Therefore, it is very important not only to investigate the online parsing preferences of readers, but to evaluate the frequency with which these structures occur in Spanish.

Corpus study

Since *se* is such a frequent and polysemic word in Spanish, the possibility that strong co-occurrence propensities might explain a particular processing bias is a serious one, so we conducted a corpus study of the structures under consideration. Additionally, we wanted to verify whether the general pro-drop reflexes of Spanish were evident in the particular structures under investigation, as our predictions rest on the assumption that the preferred reading in the structures tested is one that is not satisfied in any: that of a discourse-linked full *pro* due to pro-drop. Our corpus contains about 450,000 words extracted from different books, newspapers, and magazines in Spanish. We found 2,395 occurrences of the different types of the word *se.* The results are shown in Table 1.

As the table shows, the vast majority of the *se* occurrences in our corpus appear in the category "others". This category is comprised almost entirely of *se* that co-occurs with Spanish verbs like "ir" (*go*), "beber" (*drink*), or "marchar" (*leave*)—that is, aspectual *se*, which disappears in the English translation ("Juan se fue"/*John left*). It also contains *middle* uses of verbs such as "secar" (*dry*) as in "la ropa se secó rápidamente", which also disappear in the translation (*the clothes dried quickly*). The fact that they disappear when translated means that the reflexive, passive, or

Table 1. *Frequency and percentage of the occurrence of the different types of Spanish* se

Type of se	*Frequency*	*Percentage*	*Cumulate percentage*
Reflexive	335	14.0	14.0
Impersonal	178	7.4	21.4
Passive	191	8.0	29.4
Indirect	65	2.7	32.1
Reciprocal	75	3.1	35.2
Other	1,551	64.8	100
Total	2,395	100	

impersonal meanings that we are interested in are not involved in such uses of *se*.

As can be seen from the table, reflexive *se* appears as often as impersonal and passive *se* combined. However, only 11 out of 335 types of *se* in our corpus have displaced subjects like those we have used with our reflexive structures. Due to the multilayered nature of language, it is hard to imagine the grain that ought to be privileged for statistical analysis, if any. Despite the displaced character of the subjects of the reflexive structures used in our experiments, reflexives are an instance of a general syntactic mould (the transitive) that is much more frequent than the passive or any other major structural type. Additionally, the reflexives in our corpus appear much more often without a subject than with a subject (postverbal or not); that is, they typically co-occur with a silent *pro*: 238 without subject, 86 with a preverbal subject, and 11 with a postverbal one. This confirms the general pro-drop tendency of Spanish and adds to the body of evidence uncovered by Demiral et al. (2008) and others to sustain our view that the natural reading of all our structures involves a missing pronominal subject with full content (that is, one in a singleton chain, in de Vincenzi's, 1991, terminology). In any case, it is nevertheless also true that reflexive sentences do not usually appear with explicit displaced subject, so we offer these data but remain agnostic on the grain problem. The fact is that if lexical frequency is taken to exert its influence even before syntactic templates are recognized, then it is clear that only aspectual *se* of the sort not examined here (with a 65% overall frequency) would have a right to claim the benefit from frequency, followed by the reflexive *se*, and then all the other types of *se*, including passive and impersonal.

EXPERIMENT

Method

Participants

A total of 44 undergraduate students of the University of La Laguna received course credit for their participation in this experiment. All the participants were native Spanish speakers with normal, uncorrected vision, and they were naive with respect to the purposes of the experiment.

Design and materials

A total of 40 groups of sentences similar to those in Table 2 were constructed. The design of the experiment was a 2 × 2 with repeated measures on the factors: (a) type of sentence: *reflexive sentences, with presence of* **el** *versus reflex passive or impersonal sentences, with presence of* **al**, and (b) type of postverbal phrase: *postverbal noun phrase versus postverbal adjunct*, which could be an adverb or a prepositional phrase. Thus, each sentence had four conditions (see Table 2): postverbal adjunct reflexive (*el* sentence); postverbal noun phrase reflexive (*el* sentence); postverbal adjunct impersonal (*al* sentence); and postverbal noun phrase reflex passive (*al* sentence).

The participants read 10 experimental sentences in each condition. The four experimental conditions were counterbalanced so that each participant read each experimental sentence only in one condition. All experimental sentences were presented on one line, as is shown in Table 2. The slashes in the sentences in Table 2 represent the regions into which they were split for data analysis purposes. The slashes were not visible during the experiment. The 40 experimental sentences were presented to the participants randomly intermixed with 100 unrelated sentences.

Apparatus

The sentences were presented in lower-case letters (with upper-case letters where appropriate) on a video screen interfaced with a compatible personal computer. The participants were seated 70 cm away from the monitor, and three characters equalled 1º of visual angle. A Fourward Technologies Dual Purkinje Image Eyetracker interfaced with a computer was used to record the participants' eye movements. The eye tracker had a resolution of 10 arcmin (half a character). Viewing was binocular, but eye position was recorded only from the right eye. The signal

Table 2. *Sample sentences used in the experiment*

Condition	*Sample sentence*
1. Postverbal adjunct reflexive (*el*)	Se vendó/ apresuradame/nte *el*/ corredor/ de fondo/ antes de empezar la carrera./ The long distance runner bandaged himself hurriedly before the race.
2. Postverbal noun phrase reflexive (*el*)	Se vendó/ el tobi/llo *el*/ corredor/ de fondo/ antes de empezar la carrera./ The long distance runner bandaged his ankle before the race.
3. Postverbal adjunct impersonal (*al*)	Se vendó/ apresuradame/nte *al*/ corredor/ de fondo/ antes de empezar la carrera./ The long distance runner was bandaged hurriedly before the race.
4. Postverbal noun phrase reflex passive (*al*)	Se vendó/ el tobi/llo *al*/ corredor/ de fondo/ antes de empezar la carrera./ The long distance runner's ankle was bandaged before the race.

Note: Slashes indicate the limits of the regions used for the data analyses. They were not present during the reading task of the experiment.

from the eye tracker was sampled every millisecond by the computer.

Procedure

When a participant arrived for the experiment, he or she was seated in front of the monitor. A chinrest and a headrest with a securing strap were used to prevent head movements. The participants were asked to silently read the sentences displayed on the monitor. They were told that they would be asked questions about the sentences and that they should read for comprehension. Once the instructions were given, the eye-tracking system was calibrated, which took about 5 min. Each participant then read 10 practice sentences in order to become familiar with the procedure. When the participant pressed a button to begin a trial, a sentence immediately appeared on the screen. After reading the sentence, the participant pressed a response button, which resulted either in the presentation of a question in the centre of the screen, with the two alternative answers below it (one at the right and the other at the left of the screen), or in the row of calibration check boxes for the next trial. Questions were asked in one third of the experimental trials and in half of the filler trials. Questions in the experimental items were always related to the last part of the sentence, so they had no relation to the manipulations. For the sentence in our example the question was: "¿Cuándo se vendó al/el corredor?" (*When was the runner bandaged?/When did the runner bandage himself?*). The two alternative answers were: "Antes de la carrera/Después de la carrera" (*Before the race/After the race*). When a question appeared on the screen, the participant had to press one of two buttons to state which answer was the correct one (left button for the left answer and right button for the right answer). Half of the questions had the correct answer on the right, and half had the correct answer on the left. No feedback was given about question–answer accuracy.

After the participants completed each third of the experiment, they were given a break. After each break, the equipment was recalibrated, and the experiment was resumed at the point at which it had been interrupted. Apart from the two routine recalibrations, if a failure in calibration was detected before any sentence, the experimenter recalibrated the apparatus.

Results

The mean question–answer accuracy rate was 95%. The participants had a maximum of 12 errors out of the total 70 comprehension questions in the experiment, which represents a minimum accuracy rate of 83%. In the experimental sentences, the participants had a maximum of 5 errors out of the 20 comprehension questions, which represents a minimum accuracy rate of 79%.

For analysis purposes, the experimental sentences were divided into six regions, as can be seen in Table 2. Data are reported only for the initial four regions, as we did not find results of

interest in the two last regions. The first region always contained the word *Se* plus the main verb of the sentence, which was an optional transitive verb in singular, third person, simple past form. The second region contained the postverbal noun phrase—for Conditions 2 and 4—or the postverbal adjunct of the matrix verb (an adverb or a prepositional phrase)—for Conditions 1 and 3—except for its three last characters. The third region contained the three last characters of the preceding word plus the word *el* (Conditions 1 and 2) or *al* (Conditions 3 and 4) and included the space between this word and the previous one. The fourth region included the recipient of the action expressed by the matrix verb—for Conditions 3 and 4—or the subject noun phrase—for Conditions 1 and 2. In the four conditions this word was always the same. The reason for extending the third region three characters to the left (four including the space) was to incorporate in the measures for this region the duration of fixations that can be parafoveally sensitive to the *al*/*el* manipulation. It is well known that short words are skipped frequently because they can be easily processed during a fixation close to their left (Ehrlich & Rayner, 1983; Garrod, Freudenthal, & Boyle, 1994; Van Gompel & Majid, 2003). So, we could keep a more accurate track of the effect of our manipulations by enlarging the *al*/*el* region a few characters to the left.

Measures

In this experiment we report the data of five dependent variables related to the eye movements during the reading of the experimental sentences:

First-pass reading time. This is the sum of the durations of the fixations made in a region from the moment participants first fixate the region until they leave the region in a backward direction or look at a later region. *Total time* is the sum of the durations of all the fixations made in a region without restrictions. *Regression path time* is the summed fixation duration from the moment the region is first fixated until the eyes first move past the region; this includes first-pass time, time spent in previous parts of the sentence following any regressive eye movements, and time due to refixations coming from the left before the eyes move past the region. A very important contribution to this measure, and (as we will see) a very significant one, is the time spent at the left of the region considered. We call this measure *regression path time to the left*, and it includes only time spent in the previous parts of the sentence between the moment the region is first entered and the moment the eyes go past it. As a last measure we consider *second-pass time* as the summed duration of the fixations in which the region was refixated once it was left rightwards. Trials with no fixations in the region being analysed were eliminated in calculating all measures, except the second-pass time measure in which no fixated regions contributed in durations of zero.

The first-pass time is classically considered as an "early" measure that can provide information about the very first processes, both lexical and syntactic, that occur when a word is fixated. Although the other measures are all normally considered as "late", we can distinguish among them according to their different temporal and spatial resolutions. From this point of view, regression path measures can inform us about the point of the sentence at which readers had problems in left to right reading, and second-pass measures can inform us about where readers went to solve those problems (Meseguer, Carreiras, & Clifton, 2002). These two types of measure provide a sharper temporal definition than the classic total time measure, which does not distinguish between first- and second-pass times. This is why we take these measures into account as well as the more standard measures.

Analyses

Table 3 shows the values of the measures that we used in the four regions of interest. The statistical analyses of the five measures in the four regions of interest are included in Tables 4–7. (For the sake of simplicity only significant results are included).

The most important effects found in our experiment are highlighted in Tables 4–7 in bold type.

Table 3. *Means of eye movement measures by region and condition*

		Region			
Measure	*Condition*	*R1*	*R2*	*R3*	*R4*
First-pass time (ms)	PVAd el	690	477	250	397
	PVNP el	712	371	251	381
	PVAd al	722	487	265	407
	PVNP al	707	376	281	385
Total time (ms)	PVAd el	927	715	370	545
	PVNP el	968	563	379	539
	PVAd al	949	718	365	543
	PVNP al	885	506	399	523
Regression path time (ms)	PVAd el	682	687	377	556
	PVNP el	709	550	428	588
	PVAd al	711	687	408	546
	PVNP al	696	524	392	536
Regression path time to the left (ms)	PVAd el	—	106	92	116
	PVNP el	—	103	150	145
	PVAd al	—	102	111	92
	PVNP al	—	88	80	104
Second-pass time (ms)	PVAd el	238	153	125	87
	PVNP el	267	145	154	89
	PVAd al	232	142	115	83
	PVNP al	189	107	129	71

Note: PVAd: postverbal adjunct conditions. PVNP: postverbal noun phrase conditions. el: reflexive conditions. al: impersonal/passive conditions. R1–R4 are the four regions of interest: R1 = Se + verb; R2 = PVAd/PVNP – 3 characters; R3 = 3 characters + el/al; R4 = noun.

As can be seen in Table 7 (Region 4), a preference for the impersonal/passive versions (Conditions 3 & 4: sentences with *al*) versus the reflexive versions (Conditions 1 & 2: sentences with *el*) appeared in Region 4 (noun) in regression path time (541 ms vs. 572 ms), although this preference was only marginally significant in the analysis by participants. As we pointed out previously, an important contribution to this measure comes from the time that readers spend in regressions to previous parts of the sentence, which is captured by *regression path time to the left*. A clearly reliable preference in Region 4 (noun) for the impersonal/passive versions in this measure (98 ms for impersonal/passive vs. 131 ms for reflexive) demonstrates that readers spent more time in regressions from that region in reflexive versions than in the impersonal/passive versions.

The preference for the impersonal/passive (*al*) versions is confirmed by the second-pass time effects in the same direction as that found in Regions 1, 2, and 3 (Tables 4–6). In first region (*Se* + main verb) readers spent more time rereading reflexive versions (Conditions 1 & 2, *el*: 253 ms) than passive/impersonal versions (Conditions 3 & 4, *al*: 211 ms). In the second-pass time of this first region we also found an interaction of the two variables, which we comment on later. The same main effect was found in the rereading of the second and third regions. The second-pass reading time was longer in the second region (postverbal noun phrase/postverbal adjunct minus 3 characters; see Table 5) when it was followed by *el* (reflexive conditions: 149 ms) than when it was followed by *al* (impersonal/passive conditions: 125 ms). The third region (three characters plus *al* or *el*) presented a tendency, only marginally significant in the by-participants analysis, to be refixated more

Table 4. *Statistically reliable effects of the two factors in the first region of interest*

		R1 (Se + main verb)														
		F_1			F_2		F'			F_1			F_2		F'	
	Effects	F_1	*df*	*MSE*	F_2	*df*	*minF′*	*df*	*Simple effects*	F_1	*df*	*MSE*	F_2	*df*	*minF′*	*df*
First pass																
Total time	**interaction**	7.28**	1, 43	16,284	6.03*	1, 39	3.29(*)	1, 80	PVNPpass < PVNPrefl	7.88**	1, 43	19,036	4.76*	1, 39	2.96(*)	1, 75
									PVNPpass < PVAdpass	5.19*	1, 43	16,934	4.10*	1, 39	2.29	1, 80
Regr. path time																
Regr. path time to the left																
Second pass	**imp/pass < refl**	8.34**	1, 43	9,475	8.50**	1, 39	4.21*	1, 82	PVNPpass < PVNPrefl	10.2**	1, 43	13,245	17.2**	1, 39	6.38*	1, 79
	interaction	5.23*	1, 43	10,779	4.56*	1, 39	2.62	1, 82	PVNPpass < PVAdpass	4.44*	1, 43	9,026	4.39*	1, 39	2.2	1, 82

Note: PVAd: postverbal adjunct conditions; PVNP: postverbal noun phrase conditions. Refl: reflexive conditions (*el*); imp/pass: impersonal/passive conditions (*al*).
(*)$p < .07$; *$p < .05$; **$p < .01$.

Table 5. *Statistically reliable effects of the two factors in the second region of interest*

		R2 (PVNP/PVAd − 3 characters)						
		F_1			F_2			
	Effects	F_1	*df*	*MSE*	F_2	*df*	*F′ minF′*	*df*
First pass	PVNP < PVAd	104**	1, 43	4981	15.9**	1, 39	13.7**	1, 51
Total time	PVNP < PVAd	132**	1, 43	11037	25**	1, 39	21**	1, 53
Regr. path time	PVNP < PVAd	68.03**	1, 43	14581	16.58**	1, 39	13.3**	1, 57
Regr. path time to the left								
Second pass	**imp/pass < refl**	4.46*	1, 43	5983	4.23*	1, 39	2.17	1, 82

Note: PVAd: postverbal adjunct conditions; PVNP: postverbal noun phrase conditions. Refl: reflexive conditions (*el*); imp/pass: impersonal/passive conditions (*al*).
*$p < .05$; **$p < .01$.

often in the reflexive *el* conditions than in the impersonal/passive *al* conditions (140 ms vs 122 ms; see Table 6).

Up to this point we have presented the results of five measures informing about a preference for the impersonal/passive (*al*) versions of the sentences versus the reflexive (*el*) versions. It seems that all these effects have their origin in the noun after *al* or *el*. But before these effects take place, we can observe an interaction of the two variables in three related measures: regression path time to the left of the third region, and total time and second-pass time of the first region.

Chronologically speaking, the first place where we can find this interaction is in the regression path time to the left of the third region (3 characters plus *al*/*el*; see Table 6). Paired comparisons showed that readers spent less time to the left of the third region in sentences with a postverbal noun phrase when it was followed by *al* (Condition 4: postverbal noun phrase passive; 80 ms) than when it was followed by *el* (Condition 2: postverbal noun phrase reflexive; 150 ms). None of the other paired comparisons showed reliable effects.

Total time measure showed an effect in Region 1 (*se* plus main verb) similar to that found in Region 3 for regression path time to the left: a reliable interaction between the two factors, (see Table 4). As we found in the previous measure, paired comparisons showed that in sentences with a postverbal noun phrase the first region of the *al* sentences (Condition 4, postverbal noun phrase passive: 885 ms) took less total time than the first region of the *el* sentences (Condition 2, postverbal noun phrase reflexive: 968 ms). Post hoc analyses also showed that *al* sentences took less total time when they had a postverbal noun phrase (Condition 4, postverbal noun phrase passive: 885 ms) than when they had an adjunct (Condition 3, postverbal adjunct passive: 949 ms). This pattern of results, and its resemblance to that of regression path time to the left, seems to indicate that when readers had trouble in Region 3 they returned to the beginning of the sentence to reread it.

This interpretation is supported by the effects found in the second-pass time measure of the first region; actually, a similar interaction (already mentioned) appeared here (see Table 4). As in the other two measures, paired comparisons showed that participants spent less time rereading the first region of the postverbal noun phrase sentences when they were passive (Condition 4, postverbal noun phrase passive—*al*—189 ms) than when they were reflexive (Condition 2, postverbal noun phrase reflexive —*el*—267 ms). As happened in total time, in second-pass time we also found that the first region of the *al* sentences was reread faster when it contained a postverbal noun phrase (Condition 4: 189 ms) than when it contained a postverbal adjunct (Condition 3: 232 ms).

Apart from all the above results (those highlighted in Tables 4–7) we obtained another set

Table 6. *Statistically reliable effects of the two factors in the third region of interest*

		R3 (3 characters + al/el)														
		F_1			F_2		F'			F_1			F_2		F'	
	Effects	F_1	*df*	*MSE*	F_2	*df*	*minF′*	*df*	*Simple effects*	F_1	*df*	*MSE*	F_2	*df*	*minF′*	*df*
First pass	imp/pass > refl	9.73**	1,43	2,330	5.51*	1,39	3.51(*)	1,74								
Total time																
Regr. path time																
Regr. path time to the left	**interaction**	5.0*	1, 43	17,732	4.65*	1, 39	2.4	1, 81	PVNPpass < PVNPrefl	4.11*	1, 4	26,548	5.49*	1, 39	2.35	1, 81
Second pass	**imp/pass < refl**	4.0(*)	1,43	3,738	2.46	1,39	1.52	1,76								

Note: PVAd: postverbal adjunct conditions; PVNP: postverbal noun phrase conditions. Refl: reflexive conditions (*el*); imp/pass: impersonal/passive conditions (*al*). (*)$p < .07$; *$p < .05$; **$p < .01$.

Table 7. *Statistically reliable effects of the two factors in the fourth region of interest*

		R4 (noun)						
		F_1			F_2		F'	
	Effects	F_1	*df*	*MSE*	F_2	*df*	*minF'*	*df*
First pass								
Total time								
Regr. path time	**imp/pass < refl**	3.71(*)	1,43	11440	1.93	1,39	1.26	1,72
Regr. path time to the left	**imp/pass < refl**	6.85*	1,43	8970	7.8**	1,39	3.64(*)	1,82
Second pass								

Note: PVAd: postverbal adjunct conditions; PVNP: postverbal noun phrase conditions. Refl: reflexive conditions (*el*); imp/pass: impersonal/passive conditions (*al*).
(*)$p < .07$; *$p < .05$; **$p < .01$.

of results not related to our manipulations but with uncontrolled factors such as word length and frequency. We present these results below, but as they are of little or no interest for our research, we do not comment on them in the Discussion section.

In the first-pass time measure we found that reading time for the second region (postverbal noun phrase/postverbal adjunct minus 3 characters) was longer for the adjunct versions (Conditions 1 and 3: 482 ms) than for the noun phrase versions (Conditions 2 and 4: 374 ms; see Table 5). It was also found that in the third region (3 characters plus *al*/*el*) the first-pass reading time for the *al* versions (Conditions 3 and 4: 273 ms) was longer than that for the *el* versions (Conditions 1 and 2: 251 ms; see Table 6). No other effects were found for first-pass time in any of the other regions in which the sentences were divided.

In regression path time we found the same effect as in first-pass time in Region 2 (postverbal noun phrase/postverbal adjunct minus 3 characters). Again postverbal noun phrase conditions (2 and 4) took less time (537 ms) than postverbal adjunct conditions (1 and 3; 687 ms; Table 6). And again the same effect was found in the total time of the second region. Postverbal noun phrase conditions (2 and 4) took less time (535 ms) than postverbal adjunct conditions (1 and 3; 717 ms).

As we previously pointed out, lexical frequency and length factors can account for all these last four results. On the one hand, the three effects of preference of postverbal noun phrase versus postverbal adjunct in the second region can be easily explained by the significantly higher average frequency of the words in the postverbal noun phrase (13,370 occurrences per million) than the frequency of those in the postverbal adjunct (5.37 occurrences per million), $t(39) = 6.46$, $p < .0001$; and by the significantly lower average length of the postverbal noun phrase (7.15 characters) than the average length of the postverbal adjunct (10.5 characters), $t(39) = -5.75$, $p < .0001$. On the other hand, the first-pass time effect of preference for the reflexive (*el*) versions found in the third region (3 characters plus *al*/*el*) can be explained by the higher lexical frequency of the article *el* (24,696 occurrences per million) than of the contraction *al* (4,436 occurrences per million).

Discussion

An interaction between the two factors was found in several measures. Planned comparisons showed a preference for postverbal noun phrase *al* sentences versus postverbal noun phrase *el* sentences, and no differences in the postverbal adjunct conditions between *al* and *el* sentences was found, as well as a preference for both *al* (impersonal/passive) conditions versus the *el* (reflexive) conditions. We found these effects distributed along the four regions of the sentence in four out of

the five measures used: total time, regression path time, regression path time to the left, and second-pass time.

The interaction is apparent in three related measures: the regression path time to the left of the third region (3 characters plus *al* or *el*) and second-pass and total times of the first region (*Se* plus main verb). What can be seen in those interactions is an effect of the *al* versus *el* manipulation only in the sentences with a postverbal noun phrase. Apparently, when readers reach the disambiguation they have problems in the reading of the reflexive (*el*) version that has a postverbal noun phrase, and, consequently, they return to the beginning of the sentence to reread it. On the contrary, in the adjunct versions readers do not show any difficulties in the disambiguating region (third region: 3 characters plus *al* or *el*), but then, as they keep reading past it, a second type of effect appears. Now what we have is a main effect of the *al* versus *el*, with a preference for the *al* version, without an interaction with the postverbal noun phrase versus postverbal adjunct factor. This effect appears twice: in a measure that implies going out to the left of a region, regression path time to the left of the fourth region (disambiguation plus 1); and in a measure that implies receiving refixations mainly from the right: second-pass time of Regions 2 and 3. The same main effect was found for the second-pass time of the first region, although in this case, as has already been pointed out, there was also an interaction between both factors. It is possible that the first region received refixations from the disambiguation and from subsequent regions, producing both the interaction and the main effect mentioned.

These effects fit quite well with our predictions: that the reflexive *el* conditions ought to present more reading problems than the impersonal/passive *al* conditions, and that these effects should be stronger in the postverbal noun phrase sentences than in the sentences with postverbal adjuncts. What we found was quite an early effect of the *al* versus *el* manipulation only in the postverbal noun phrase condition, while a main effect of this variable was found later, as a spillover, in postdisambiguating regions. This is an interesting effect, which allows us to discard two other possible interpretations. On one hand, it could be claimed that the fact that reflexives were harder than impersonal/passives may reflect the markedness of out-of-the-blue postverbal subject structures, without further discourse support. However, this hypothesis would predict differences only between the reflexives and the impersonals, since both the passive sentences and the reflexives with postverbal phrases hold postverbal subjects. In contrast, what we found was an overall effect of *el* versus *al*, which was earlier in the case of the two sentences with postverbal subjects. Another possible alternative hypothesis could be that the eye-movement patterns may reflect the difficulty of identifying the correct structures based on the available bottom-up information. When readers see "al corredor" the dative preposition "a" provides a direct cue to the syntactic parse of the NP. In contrast, when they see "el corredor" in the critical region more steps are needed in order to determine that this is a subject NP. An NP of the form "el N" is, in principle, ambiguous between nominative and accusative forms. However, this is only the case when contrasting reflexives and impersonals with postverbal adjuncts. As we have argued earlier, *el corredor* is as unambiguous as *al corredor* once it occurs after *el tobillo*. If we discard nonprimary constituents first, after "se vendó el tobillo" a phrase headed by "el" can only be a subject, the reason being that the previous "el" phrase (el tobillo) makes it impossible both to have two subject NPs or a passive subject ("el tobillo") followed by an object phrase headed by "el" (one cannot have "se vendó el tobillo el corredor" with *el corredor* being the recipient of the action). Thus, assuming an incremental parser, the reflexive and the passive sentences become equally unambiguous when reaching *el* or *al*.

GENERAL DISCUSSION

The gist of the effects we found is that the structures involving displaced subjects (the reflexives) are harder to process than the structures without displaced subjects, and that the difficulty increases if there are more referential constituents along the

path of the moved phrase. As noted in the introduction, there is recent evidence that processing is disturbed when the sequential order of the linguistic input does not correspond to the unmarked word order of a language, which seems to indicate that word order per se has an important role in the shaping of processing biases. According to Featherston et al. (2000), for instance, the processing of noun phrase trace is more difficult than the processing of PRO (the hypothesized silent subject of infinitives) because of movement in the former (even if this movement is "derivational"—that is, leading from something analogous to a deep structure to the actual surface form). This is consistent both with de Vincenzi's (1991) minimal chain principle and the active filler strategy (Frazier & Flores D'Arcais, 1989) and backs up de Vincenzi's general idea that movement is costly for processing. Fiebach et al. (2001, 2002) also provide electrophysiological evidence that maintaining object *wh*-fillers in working memory disturbs processing. In the present study, the most obvious finding is that the structures involving displaced subjects, similar to those of de Vincenzi's Experiment 1, also incur a processing cost.

It appears that the essence of the minimal chain principle is therefore well founded. In what follows, however, we aim to go a little further by examining specific aspects of the structures that we have analysed and its consequences for the minimal chain principle. Recall that a critical difference between the Fiebach et al. (2001, 2002) and the Featherston et al. (2000) data on the one hand and ours on the other is that in our structures it is the appearance of the gap that first prompts the search for the filler, instead of the other way around (like a *wh*-phrase, for instance).

Our results suggest that displacement is such a disturbing process that it causes the parser to allocate more resources to dealing with its effects than to dealing with structures that are complex and deviant in themselves. The two reflexive patterns that we have used are very simple transitive structures. Conversely, the special kind of *se* reflex-passive and the very special kind of impersonal *se* that we have contrasted with the reflexives are far from simple. In fact, linguists have debated about them for a long time (see Jaeggli, 1986, and Sánchez López, 2002, for reviews), and only a minimal consensus on their structures and properties has been reached. Passives—of whatever type—always involve a degree of complexity relative to the active subject–verb–object template from which they are usually derived. For instance, the passive "se" is linguistically very complex because among other reasons, (a) it has postverbal subjects that look like active objects; (b) the subject position remains therefore empty in S-structure; (c) subjects of *se* passives display the typical properties of unaccusatives (including subject-postposition and freedom to dispense with determiners), which testify to their "deep" object status (Burzio, 1981; Levin & Rappaport-Hovav, 1995; for a minimalist account, see Raposo & Uriagereka, 1996).

As for impersonals, they embody the strange grammaticalization of an a priori ungrammatical string (Blevins, 2003): one with no subject, full or dummy. Otero's (1986) asterisked *pro**, as well as other similar explanatory manoeuvres, simply testify to the fact that it is odd for the system of grammar to give its seal of approval to a tensed structure with no recoverable subject (something impossible in English, for instance). The asterisked *pro** (like the passive "e" in part as well) means that (lacking content) some "form" (a place holder or a dummy) ought to be there, but there is none. That is, full *pro* is a contentful pronoun like any other pronoun (*he*, *she*, etc). Expletive *pro* inherits content via coindexation with a displaced lexical filler (the Gianni example). But impersonal *pro** is inherently contentless (it means "one" or "someone unspecified"; that is, it does not have specific referentiality), something one might expect to pose problems for a parser.

Given the asymmetric comparisons that we have established in terms of theoretical linguistic complexity, it is quite impressive that plain reflexive structures have come out the losers just because their subjects were not in place. All the more impressive is the fact that such processing cost be demanded in a language like Spanish, which is not characterized by extreme structural rigidity. In Spanish, subjects bearing the focus of

information are routinely displaced. Displacement is in fact the means that most Romance languages have of accommodating information structure in order to express the informational content that English typically expresses using stress focus (Lambrecht, 1994).

Notice that, as our corpus study showed, it is unlikely that the results of our experiments can be explained as a consequence of frequency biases. Although the level at which statistical propensities should count is unknown, the fact remains that both at the lexical and at the syntactic level nothing in the corpus suggests that the difficulties that subjects had with the two reflexive patterns are due to frequency of occurrence. Morphosyntactically, reflexive *se* is more frequent than either impersonal *se* or passive *se*. Syntactically, reflexive *se* is a member of a large family of active transitive structures, whereas the impersonal construction and the *se* passive construction are relatively marginal. Lexically, only aspectual types of *se* of the kind not investigated here (with over 65% of overall occurrence) would seem to be solid candidates for the benefits of entrenchment and accompanying processing speed. All in all, the fact that reflexive *se* does not tend to have displaced subjects in the corpus does not seem to compensate for all the other counts, especially since another count shows that the vast majority of reflexive structures do not even have a phonetically realized subject, displaced or not, because of pro-drop. Thus, although one cannot conclusively rule out a frequency reading of our results, the accumulated evidence gathered by our corpus study makes it less convincing.

A second, more specific conclusion we can reach is that comparing the two-argument reflexive and the impersonal on the one hand and the three-argument reflexive and the passive on the other reflects an effect of integration cost. Gibson's (1998) *syntactic prediction locality theory* differentiates between an integration cost component and a memory cost component (see also Grodner, Gibson, & Tunstall, 2002). The integration cost is increased if more intervening referents exist along the path hit by the parsing scan during the integration process. Relative to the impersonal and the two-argument reflexive, in the three-argument reflexive and the passive an additional third argument of the verb must be integrated into the ongoing structural representation of the sentence while the subject noun phrase gap must at the same time be kept in memory until a filler is found. The full integration of all the tree nodes seems to take place a little after the displacement operation has been sorted out, as the minimal chain principle would predict. We conclude this from the interaction that we have found in the total time and the second-pass time in the first region where differences affect the comparison between the three-argument reflexive and the passive, but not that between the two-argument reflexive and the impersonal. It appears that in both reflexives there is an effect of movement but that only in the second, three-argument reflexive is this effect followed by greater integration cost.

We proceed now to consider how the information from the eye-tracking methodology may be interpreted in connection with our third objective—whether the mind would somehow register or find trouble with the functional shift that the first postverbal noun phrase (*el tobillo*) experiences in the passive, relative to a hypothetically more basic subject–verb–object active transitive structure (where the same noun phrase would be a canonical object). The minimal chain principle, as a configurational principle, is not expected to deal with such cases a priori. Relative to an active subject–verb–object kernel, in the passive nothing moves, at least overtly. Since Spanish is a pro-drop language, the first option that the minimal chain principle would favour for both the passive and the three-argument reflexive would be one in which a full *pro* is recognized as filling the subject gap. That would most likely be somebody latent in a putative context: say *Ana* in 1. *Ana se casó /Ana got married. 2. (pro) Está encantada/(She) is thrilled*. In fact, Demiral et al. (2008) have recently proved that a full *pro* is invoked by the parser even in the absence of a supporting context (see also Carminati, 2005) This is because such a *pro* (licensed by the verb's inflection) is simply like a subject pronoun placed in

the habitual position right before the verb (from which it receives case and thematic role). The discourse-linked full *pro* option is, however, realized in neither of the two structures, a fact that becomes apparent in the disambiguating region: The nonaccusative/nominative phrase *el corredor* in the reflexive means that the true subject is there now (its displaced position thus causing a complex chain); the accusative phrase *al corredor* in the passive means that full *pro* must also be ruled out as it involves analysing *se* as indirect object ("somebody bandaged the ankle 'to' himself"), and now the accusative introduces the true indirect object ("somebody bandaged the ankle to the runner"). One may assume that since reanalysis is necessary in both structures, and since the initial subject gap cannot come from a previous context (and a full *pro*), the subject gap must be reevaluated, and a search for a filler must be initiated inside the sentence.

Faced with this situation, how can we make sense of the fact that it seems easier for the parser to recover from its initial reflex to build full *pro* in the passive sentence than in the reflexive sentence? Notice that in the *pro ha chiamato Gianni* sentences that de Vincenzi (1991) studied, once the parser realizes that *pro* is coindexed with a displaced Gianni, reanalysis becomes mandatory but straightforward, and it affects only the complex chain interpretation. However, in our comparisons reanalysis is necessary in *both*, and still the passive is less problematic than the reflexive. We suggest that what may be happening is that the parser is able to immediately "see" the postverbal phrase in an underlying SVO template, ignoring its surface function. However, what the parser seems to have trouble ignoring is the fact that in the reflexive sentence, a NP subject appears displaced at the end, as this involves a complex chain that must suddenly be undone. Put differently, even though the NP "el tobillo" is functionally ambiguous, the phrase is in its correct "deep" order, which makes it nonambiguous to a parser committed to looking for "deep" patterns. There is nothing functionally ambiguous with the subject phrase "el corredor". However, its position entails the creation of a complex chain that relates it to the preverbal gap, and that chain delays processing, in the classic de Vincenzi formulation. This also lends further credence to the view that pro-drop structures are highly acceptable even in the absence of a specific, supporting context, since postulating a null subject seems to be easier than repositioning a present one (Demiral et al., 2008). This interpretation is also compatible with ERP and behavioural data (a speed–accuracy trade-off task) by Bornkessel et al. (2004), which suggests that structural reanalysis is more costly than functional reanalysis, even in a morphologically rich language like German.

This need for syntactic orderliness and the effect of integration costs in the structures with more arguments is cognitively intriguing. One might contend that theories such as Gibson's syntactic prediction locality theory, for instance, simply redescribe the facts, and that what is really needed is a psychological explanation for why an additional constituent engenders a cost. Gibson (1998) has emphasized *decay* as the source of processing complexity, but there seems to be no solid evidence for that. Alternatively, other researchers like van Dyke and McElree (2006) have studied *interference effects* in sentence comprehension in a cue-based parsing framework (van Dyke & Lewis, 2003) aimed at resolving situations where the available cues cannot distinguish among competitors. In such a framework, grammatical relations are created via cue-based retrieval of necessary constituents. Grammatical heads thus provide retrieval cues that identify necessary properties of the parsed constituent (such as clause function, grammatical case, or meaning), which are then combined in parallel to access a single retrieval probe. Similarity-based interference arises when there are distractor items very similar to the target item vis-à-vis the retrieval probe. Gordon, Hendrick, and Johnson (2004) also study interference effects in a line of research that is usually designed to show that the memory system that subserves language comprehension draws from the same pool of resources as that used for other domains. In fact, in principle, Gordon et al.'s work might seem more appropriate for present purposes because they specifically deal with referential NPs creating competition. We cannot discard effects caused by different cues in

our experiments. For instance, as one reviewer points out, even if our three-argument reflexive sentences are completely unambiguous at the final relevant segment ("el corredor"), the bottom-up cues from a nominal phrase like "el corredor" might still be less informative than those provided by "al corredor". However, one positive aspect of the tested manipulations is that in our structures there is really no referential competition because the referential properties of the distractor items are identical (*el tobillo* in all cases), and the gap itself is the same in all conditions. In fact, maybe the most favourable cue for a passive reading of the ambiguous segment is absent: The clitic "le", which would normally be used in the passive, was not used in our sentences for reasons of experimental design. Finally, in decontextualized transitive, reflexive structures, as the sentence unfolds, and the unique external argument of the predication is not in sight yet, the mere appearance of a nominative-compatible "el" phrase might be taken to be the best possible cue for the successful completion of the predication. In that sense, one could argue that the nominative-compatible cue is primed by the incrementality of the reading process.

While cue-based research is nevertheless broadly compatible with our findings, here we would like to take a more linguistic stance to stress the fact that the gap that we are dealing with is a subject gap, not just any kind of ellipsis. This is the same in all conditions. Subjects are privileged constituents in grammar, and languages normally go to great lengths to grammaticalize them. In English, for instance, dummy (i.e., semantically empty) *there* or *it* (***there*** *is a book on the table there* / ***it*** *is important that she stays in charge the whole period*) must be summoned to fill the subject position when this is not thematically easy to fill (because it is a whole predication, for instance: *that she stays in charge the whole period is important = it is important that she stays in charge the whole period*). In generative grammar, it has always been mandatory to generate the so-called "external argument" of the predication in a very precise (and early, cyclic) way in the derivation, and there can be no derivation without a subject. In most Indo-European languages, a NP to the left of the verb constitutes a very strong expectation of parsers, as the literature we have reviewed shows, and this includes German and Spanish, two languages with rather liberal word order. What we mean, therefore, is that being a subject NP is not the same kind of cue as so many other potential cues, although this is of course something to be properly measured.

Part of the theoretical importance of these findings has to do with the seemingly unavoidable conclusion that reconstruction of word order is more imperative than reconstruction of clausal function, or that it is at least more essential to the proper understanding of sentence structure. Since word order and function are two very different things, no doubt linguists will find the present results of interest. Note that in some linguistic models, such as generative grammar, in all its historical and current forms, function is a derivative notion: It comes about only as a result of phrases previously occupying the right place in an underlying and "tidy" phrase marker. So the object function, for instance, is derived from a verb phrase (VP) internal configuration when a NP is to the right of the V. Relational grammarians and lexical-functional grammarians, as well as several types of European functionalism, hold a different view. As noted, it is interesting that this word-order toll be demanded in a language with relatively liberal word order such as Spanish. Since subjects bearing the focus of information are usually postponed in this language (as in most Romance languages), the question arises as to whether a contextual manipulation of information flow could be enough to make the displaced phrase more predictable. That is a question that we are working on, but which for the moment remains open.

In sum, we believe that the general validity of the minimal chain principle has now been strengthened by evidence from a different language than Italian, using a methodology (eye tracking) that is well suited to gauge the temporal dynamics of syntactic processing.

Original manuscript received 20 June 2007
Accepted revision received 22 April 2008
First published online 21 August 2008

REFERENCES

Barrenechea, A. M., & Manacorda de Rossetti, M. (1977). La voz pasiva en el español hablado de Buenos Aires. In J. M. Lope Blanch (Ed.), *Estudios sobre el español hablado en las principales ciudades de America* (pp. 395–405). Comisión de lingüística hispanoamericana: Programa interamericano de lingüística y enseñanza de idiomas.

Blevins, J. P. (2003). Passives and impersonals. *Journal of Linguistics*, *39*, 473–520.

Bornkessel, I., McElree, B., Schlesewsky, M., & Friederici, A. (2004). Multi-dimensional contributions to garden path strength: Dissociating phrase structure form case marking. *Journal of Memory and Language*, *51*, 495–522.

Burzio, L. (1981). *Intransitive verbs and Italian auxiliaries*. Unpublished doctoral dissertation. Cambridge, MA: MIT.

Carminati, M. N. (2005). Processing reflexes of the feature hierarchy and implications for linguistic theory. *Lingua*, *115*, 259–285.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, The Netherlands: Foris.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and us*. New York: Praeger.

Chomsky, N. (1995). *The Minimalist program*. Cambridge, MA: The MIT Press.

Cinque, G. (1988). On SI constructions and the theory of ARB. *Linguistic Inquiry*, *19*, 521–581.

Cinque, G. (1995). *Italian syntax and universal grammar*. Cambridge, UK: Cambridge University Press.

Clifton, C., Jr., & Frazier, L. (1989). Comprehending sentence with long distance dependencies. In G. Carlson & M. Tanenhaus (Eds.), *Linguistics structure in language processing* (pp. 273–317). Dordrecht, The Netherlands: Kluwer.

Demiral, S. B., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2008). On the universality of language comprehension strategies: Evidence from Turkish. *Cognition, 106*, 484–500.

De Vincenzi, M. (1991). *Syntactic parsing strategies in Italian*. Dordrecht, The Netherlands: Kluwer.

De Vincenzi, M. (1998). Syntactically based parsing strategies: Evidence from typologically different languages. In D. Hillert (Ed.), *Syntax and semantics* (Vol. 31, pp. 337–344). New York: Academic Press.

Ehrlich, K., & Rayner, K. (1983). Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning and Verbal Behavior*, *22*, 75–87.

Featherston, S., Gross, M., Munte, T. F., & Clahsen, H. (2000). Brain potentials in the processing of complex sentences: An ERP study of control and raising constructions. *Journal of Psycholinguistic Research*, *29*, 141–154.

Fiebach, C. J., Schlesewsky, M., & Friederici, A. (2001). Syntactic working memory and the establishment of filler-gap dependencies: Insights of ERPs and fMRI. *Journal of Psycholinguistics Research*. *30*, 321–338.

Fiebach, C. J., Schlesewsky, M., & Friederici, A. (2002). Separating syntactic memory costs and syntactic integration costs during parsing: The processing of German WH-questions. *Journal of Memory and Language*, *47*, 250–272.

Fodor, J. D. (1979). "Superstrategy". In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Frazier, L. (1987). Syntactic processing: Evidence from Dutch. *Natural Language and Linguistic Theory*, *5*, 519–560.

Frazier, L., & Clifton, C., Jr. (1996). *Construal*. Cambridge, MA: The MIT Press.

Frazier, L., & Flores D'Arcais, G. (1989). Filler-driven parsing: A study of gap filling in Dutch. *Journal of Memory and Language*, *28*, 331–344.

Frisch, S., Schlesewsky, M., Saddy, D., & Alpermann, A. (2002). The P 600 as an indicator of syntactic ambiguity. *Cognition*, *85*, B83–B92.

Garrod, S., Freudenthal, D., & Boyle, E. (1994). The role of different types of anaphor in the on-line resolution of sentences in a discourse. *Journal of Memory and Language*, *33*, 39–68.

Gazdar, G., Klein, E., Pullum, G. K., & Sag, I. A. (1985). *Generalized phrase structure grammar*. Cambridge, MA: Harvard University Press.

Gibson, E. (1998). Syntactic complexity: Locality of syntactic dependencies. *Cognition*, *68*, 1–76.

Gordon, P. C., Hendrick, R., & Johnson, M. (2004). Effect of noun phrase type on sentence complexity. *Journal of Memory and Language*, *51*, 97–114.

Grodner, D., Gibson, E., & Tunstall, S. (2002). Syntactic complexity in ambiguity resolution. *Journal of Memory and Language*, *46*, 267–295.

Hemforth, B. (1993). *Kognitives Parsing: Reprasentation und Verarbeitung sprachlichen Wissens* [Cognitive parsing: Representation and processing of linguistic knowledge]. Sankt Augustin, Germany: Infix.

Jaeggli, O. (1986). Arbitrary plural pronominals. *Natural Language and Linguistic Theory*, *4*, 43–76.

Lambrecht, K. (1994). *Information structure and sentence form*. Cambridge, UK: Cambridge University Press.

Lamers, M. A. L. (2001). *Sentence processing: Using syntactic, semantic and thematic information*. Unpublished PhD dissertation, University of Groningen, Groningen, The Netherlands.

Levin, B., & Rappaport-Hovav, M. (1995). *Unaccusativity at the syntax-lexical semantic interface*. Cambridge, MA: The MIT Press.

MacDonald, M., Pearlmutter, N., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 676–703.

Meseguer, E., Carreiras, M., & Clifton, C., Jr. (2002). Overt reanalysis strategies and eye movements during the reading of mild garden path sentences. *Memory & Cognition*, *30*, 551–561.

Mitchell, D., Cuetos, F., Corley, M., & Brysbaert, M. (1995). Exposure-based models of human parsing: Evidence for the use of coarse-grained (non-lexical) statistical records. *Journal of Psycholinguistic Research*, *24*, 469–488.

Osterhout, L. (1994). Event-related brain potentials as tools for comprehending language comprehension. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Prospectives on sentence processing* (pp. 15–44). Hillsdale, NJ: Lawrence Erlbaum Associates.

Osterhout, L., & Swinney, D. A. (1993). On the temporal course of gap-filling during comprehension of verbal passives. *Journal of Psycholinguistics Research*, *22*, 273–286.

Otero, C. P. (1986). Arbitrary subjects in finite clauses. In I. Bordelois, H. Contreras, & K. Zagona (Eds.), *Generative studies in Spanish syntax* (pp. 82–109). Dordrecht, The Netherlands: Foris.

Raposo, E., & Uriagereka, J. (1996). Indefinite *se*. *Natural Language and Linguistic Theory*, *14*, 749–810.

Rizzi, L. (1990). *Relativised minimality*. Cambridge, MA: MIT Press.

Sánchez López, C. (2002). *Las construcciones con se* [Constructions with se]. Madrid, Spain: Visor Libros.

Sekerina, I. (1997). *The syntax and processing of scrambling constructions in Russian*. Unpublished PhD dissertation, City University of New York, New York.

Sekerina, I. (2003). Scrambling and processing: Complexity, dependencies, and constraints. In S. Karimi (Ed.), *Word order and scrambling* (pp. 301–324). Oxford, UK: Blackwell.

Sepúlveda Barrios, F. (1988). *Voz Pasiva en español del siglo XVII. Contribución a su estudio* [Passive voice in XVII century Spanish: A contribution to its study]. Madrid, Spain: Gredos.

Trueswell, J. C., & Tanenhaus, M. K. (1994). Toward a lexicalist framework of constraint-based syntactic ambiguity resolution. In C. Clifton, Jr., L. Frazier, & K. Rayner (Eds.), *Perspective on sentence processing* (pp. 155–179). Hillsdale, NJ: Lawrence Erlbaum Associates.

Van Dyke, J. A., & Lewis, R. L. (2003). Distinguishing effects of structure and decay of attachment and repair: A retrieval interference theory of recovery from misanalyzed ambiguities. *Journal of Memory and Language*, *49*, 285–316.

Van Dyke, J. A., & McElree, B. (2006). Retrieval interference in sentence comprehension. *Journal of Memory and Language*, *55*, 157–166.

Van Gompel, R. P. G., & Majid, A. (2003). Antecedent frequency effects during the processing of pronouns. *Cognition*, *90*, 255–264.