Processing coordinated structures: Incrementality and connectedness

Cognitive Science 29 (2005) 291–305

Processing Coordinated Structures: Incrementality and Connectedness

Patrick Sturt[a], Vincenzo Lombardo[b]

[a]*Department of Psychology, University of Glasgow, Scotland*
[b]*Dipartimento di Informatica, Università di Torino, Italy*

Received 1 June 2004; received in revised form 11 October 2004; accepted 15 October 2004

Abstract

We recorded participants' eye movements while they read sentences containing verb-phrase coordination. Results showed evidence of immediate processing disruption when a reflexive pronoun embedded in the conjoined verb phrase mismatched the sentence subject. We argue that this result is incompatible with models of human parsing that employ only bottom-up parsing procedures, even when flexible constituency is employed. Models need to incorporate a mechanism similar to the adjoining operation in Tree-Adjoining Grammar, in which one structure is inserted into another.

Keywords: Human sentence processing; Eye-tracking; Reading; Parsing strategies

1. Introduction

It is widely assumed that human language processing is incremental: In other words, the processor parses words in a left-to-right order while developing a representation of the sentence meaning (Marslen-Wilson, 1973). At the syntactic level, for example, this assumption is supported by garden path effects, which are usually interpreted as evidence that people do not wait for potentially useful disambiguating information before committing to a syntactic analysis (see Altmann & Steedman, 1988; Cuetos & Mitchell, 1988; Frazier & Rayner, 1982; Kamide & Mitchell, 1999; Trueswell, Tanenhaus, & Garnsey, 1994, among many others). Models of incremental human sentence processing differ in the degree of eagerness with which the syntactic structure is built. Some theories assume that each new word in the input is immediately incorporated into the partial syntactic representation, with no possibility for delaying structure building (e.g., Crocker, 1996; Konieczny, 1996; Schneider, 1999). Such approaches have been called *strongly incremental* (Lombardo & Sturt, 2002a). Other theories allow for in-

Requests for reprints should be sent to Patrick Sturt, Department of Psychology, University of Glasgow, 58 Hillhead Street, Glasgow G12 8QB. E-mail: patrick@psy.gla.ac.uk

cremental structure building in many but not all cases, predicting delays in some circumstances. Examples of such theories are head-driven approaches (Pritchett, 1991) and related activation-based models (Stevenson, 1994; Tabor & Hutchins, 2004; Vosse & Kempen, 2000) in which a structure is built (or activated) in response to recognition of its lexical head.

The amount of eagerness achieved by any given model depends on at least two factors; the parsing strategy and the grammar formalism. The parsing strategies differ in the degree of eagerness because they define different points at which structures are recognized in response to lexical input. For example, bottom-up parsers build a node only when all of its daughter nodes have been built (and this predicts delays to processes that depend on structure building), whereas top-down parsers build nodes before their daughter nodes have been built (and this predicts no delays; see Abney & Johnson, 1991, for a thorough review). The head-driven strategy mentioned previously and the left-corner strategy represent an intermediate position between these two extremes (they build some structure as soon as some relevant daughter [the head daughter or the leftmost daughter, respectively] has been recognized).

However, the parsing strategy is not the only factor to affect eagerness, because delays in structure building also depend on the nature of the constituents that are licensed by the grammar formalism. In fact, the various formalisms allow different portions of the input to be packaged as constituents (Shieber & Johnson, 1993). For example, although a noun-phrase (NP) subject followed by a transitive verb (e.g., *John devoured*) is not a constituent in traditional phrase-structure grammar, it is a constituent in the framework of Combinatory Categorial Grammar (CCG), which we discuss later.

This article considers incremental syntactic processing from the perspective of the parsing of coordinated (conjoined) phrases. Despite the fact that coordinate constructions occur in over 50% of sentences in written text,[1] very little is known about how conjoined phrases are processed, especially, in relation to the time course of structure building.[2] So, the study of coordination, although being motivated *per se* by knowing more about processing of coordination, may tell us about previously unknown properties of the human language processor. We show that the processing of coordination involves a high degree of eagerness, with no structure-building delays. In particular, we consider the consequences for incremental processing of sentences with verb-phrase (VP) coordination, such as

(1) The pilot embarrassed Mary and put himself in a very awkward situation.

Most theories of syntax treat constituent coordination, using a schema in which a conjunction such as *and* takes a constituent of Type *X* on either side, to form a new constituent of Type *X*. This can be represented, for example, by a context-free rule schema such as $X \rightarrow X$ and X; in (1), the case addressed in this article, the schema specializes into $VP \rightarrow VP$ and VP. Computational models of human language processing must account for ambiguity in the selection of the type of conjunct (sentence [S], VP, NP, …) as well as the attachment site of the conjunction (Frazier, 1978). However, in this article we abstract away from these issues because we are concentrating on the time course of structure building, and we investigate how and when the coordination schema is incorporated into the syntactic representation. Our concern here is to determine the degree of eagerness of incremental structure building by experimentally investigating human processing behavior in the case of VP coordination. To provide a global over-

view, we now temporarily draw back from processing at the word-by-word level to consider the overall problem. We discuss three crucial constituents of the sentence (1), namely the subject NP (*the pilot*), the first VP conjunct (*embarrassed Mary*), and the second VP conjunct (*put himself in a very awkward situation*), and we examine the way the parsing strategies assemble them, using the coordination schema previously mentioned. We contrast two models that can be viewed, respectively, as general examples of eager and delayed structure-building approaches. The first approach, which we call the *adjunction model,* first assembles the subject NP and the first VP conjunct into a sentence S (*The pilot embarrassed Mary*); then, the sentence structure must be modified to accommodate the coordination schema and the second VP conjunct (*put himself in a very awkward situation*). The second approach, which we call the *bottom-up CCG model* (Steedman, 2000), assembles a complete sentence structure only after the two VP conjuncts have been assembled into the coordinated VP through the coordination schema (*embarrassed Mary and put himself in a very awkward situation*). The choice of a CCG-based model is not incidental, because CCG has a particularly extensive account of coordination phenomena at the competence level, and because the flexible constituency employed in the formalism allows a high degree of eagerness in structure building in many cases (Steedman, 2000).

Now we briefly describe the two models and then we go to the experiment we designed to arbitrate between them.

2. Adjunction model

The adjunction model is a strongly incremental model that displays a very high degree of eagerness by keeping the syntactic structure fully connected while going from left to right. The adjunction model derives its name from the adjunction operation, which is well known in Tree-Adjoining Grammar (TAG; Joshi, Levy, & Takahashi, 1975). TAG encodes syntactic knowledge in terms of elementary trees instead of rules: The TAG lexicon consists of *initial trees,* which represent predicate–argument structures, and *auxiliary trees,* which represent recursive structures, including modifiers and, in our case, the coordination schema. Initial trees replace argument nodes in the syntactic structure through the *substitution* operation, whereas auxiliary trees are inserted to modify predicate–argument structures through the *adjunction* operation. In particular, the adjunction operation can expand a structure by disrupting existing immediate dominance relations by inserting modifiers in the middle of the predicate–argument structure. Similar adjunction operators have also been used in related formalisms (e.g., Gäartner & Michaelis, 2003).

The approach is illustrated in Fig. 1.[3] The elementary tree for the coordination schema in our example is a VP auxiliary tree, that is, a tree that can be inserted into the left-context structure spanning the left fragment *The pilot embarrassed Mary* at the VP node (Figure 1A). The resulting structure has a Substitution VP node (the one with a down arrow) where the initial tree for *put* can be inserted via a Substitution operation (Figure 1B). Finally, the other arguments will be substituted (Figure 1C).

It has been pointed out that an eager incremental strategy can be implemented with a top-down parser paired with a traditional phrase-structure grammar (Abney & Johnson 1991;

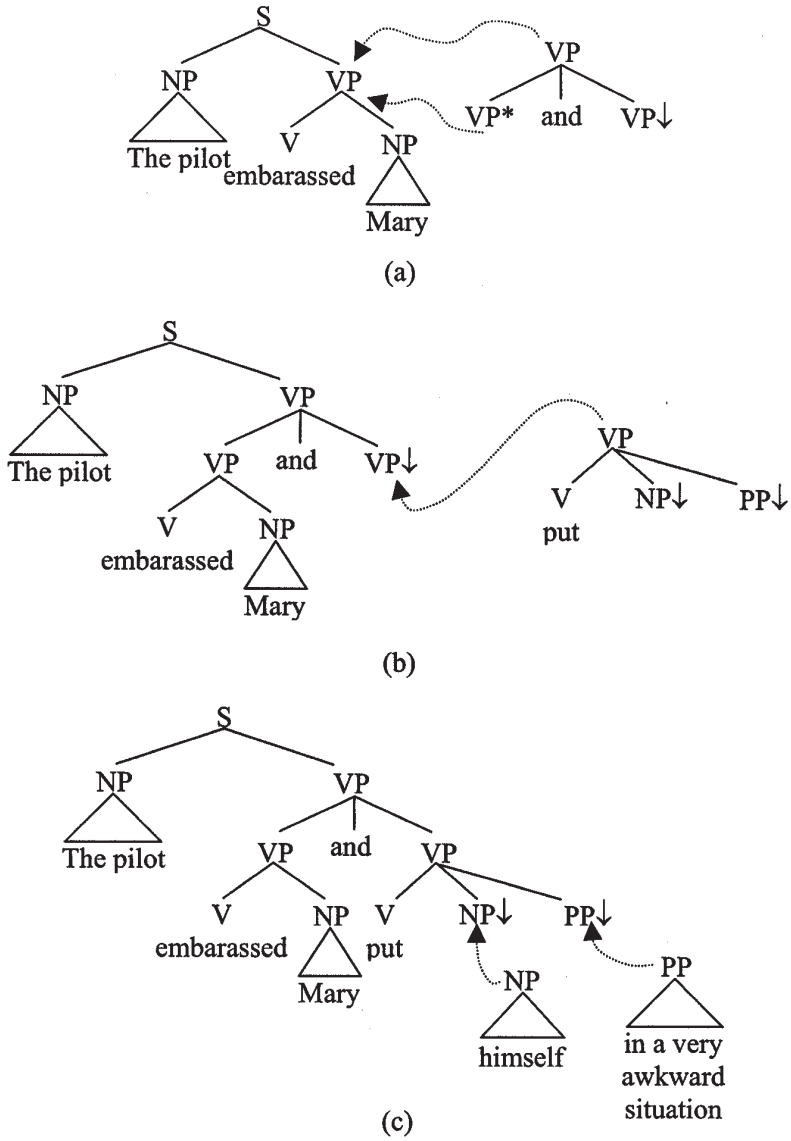

Fig. 1. VP coordination as adjunction

Roark, 2001). However, the top-down approach requires the intuitively implausible step of predicting the coordination schema in advance (before encountering the conjunction *and* in the input). To avoid predicting a coordination without evidence, it seems, then, that a very eager strategy requires a mechanism that is able to insert material directly into a structure that is already completed. One such mechanism is the *adjunction* operation discussed previously. In this article we claim that adjunction, or a similar operation, is necessary for any strongly incremental model to deal with coordination.

3. Bottom-up CCG model

The bottom-up CCG model was developed by Steedman (2000) by pairing a bottom-up parsing algorithm (preferred on a simplicity basis) with CCG, which allows flexible constituency. This approach leads to a high degree of eagerness, despite the use of a bottom-up parsing strategy. To discuss this approach, we first need to describe the grammar formalism.

In CCG, as in other categorial approaches, the categories of lexical items are either atomic category labels (e.g., NP, or S), or otherwise functions that take categories as arguments, and yield categories as the result. These functions are specified for the direction in which the argument is expected to be found: For example, X/Y means that a Y category is required on the right to yield an X, whereas X\Y means that a Y category is required on the left to yield an X. In Fig. 2B we can see a CCG derivation cast in phrase-structure terms. This is contrasted with a bottom-up derivation using traditional phrase structure (Fig. 2A). Numbers on edges indicate the sequence of derivation steps.[4]

A CCG derivation combines functional categories through a sequence of combinatorial operations, the simplest of which is *functional application,* which simply applies the functional category to an argument category to yield the result category. An example is the application of the functional category S/NP (meaning a sentence missing an NP on the right, and spanning "John thinks Peter likes") to an argument NP ("Lucy") on its right to become a sentence S (Operation 11 in Fig. 2B). A high degree of eagerness is made possible by creating more complex categories through the operation of *type raising* and combining two functions through *function composition.* In Fig. 2B, the category of *John,* which is NP, has been type raised to S/VP (in other words, a category that requires a VP on the right, to yield a sentence—Operation 3). Notice that this effectively reverses the head-dependent relation between the subject NP and the VP. This category can then combine with the sentence complement verb *thinks,* through function composition (Operation 4). The derivation proceeds incrementally through function compositions and applications until the whole sentence has been processed.

It is important to notice that in the CCG derivation (Fig. 2B) each new word completes an interpretable constituent, resulting in a left-branching structure. The purely left-branching

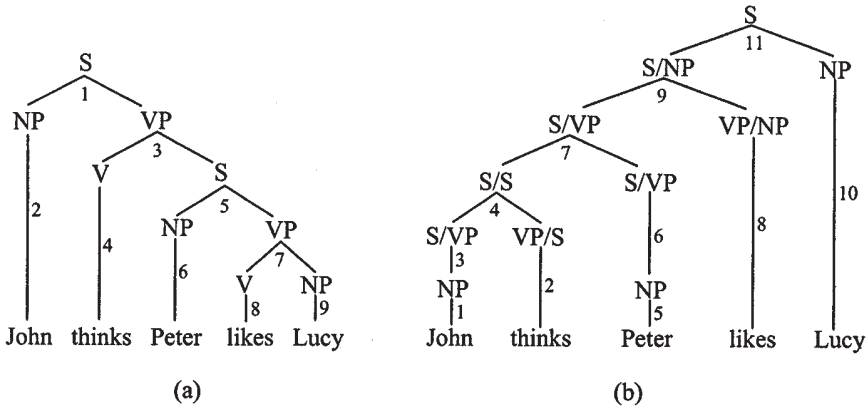

Fig. 2. Bottom-up derivation with traditional phrase structure (left), and flexible constituency in CCG (right).

structure yielded by the CCG derivation achieves a very high degree of eagerness for the purely bottom-up strategy. This left-branching analysis is only possible because of the flexible constituency allowed by CCG, as evidenced by the existence of some constituents (e.g., the S/VP "John thinks Peter") that are not traditionally considered as constituents. This is in contrast with the bottom-up derivation on the same sentence employing right-branching traditional phrase structure (Fig. 2A), where there is a considerable delay in the completion of the constituents. Such delays in building right-branching structure make purely bottom-up approaches paired with traditional phrase-structure grammars implausible for cognitive modeling (Abney & Johnson, 1991), unless extra mechanisms are employed to allow the semantic interpreter to inspect the parser's internal state (Shieber & Johnson, 1993). The discussion in this article is based on the assumption that the parser directly handles grammatical structures.

The combination of CCG and bottom-up parsing can yield a high degree of eagerness in many syntactic constructions, sometimes adopting left-branching structures. However, there exist cases where the CCG–bottom-up approach predicts attachment delays, such as those where only right-branching analyses are available (Phillips, 2003). One of these cases, discussed by Schneider (1999), is coordination, where the CCG bottom-up parser waits for the second conjunct to be complete before combining the two conjuncts into the coordination structure (see Schneider, 1999, pp. 19–21, for a discussion). For this reason, coordination provides a useful test case to evaluate the level of eagerness in human sentence processing and to adjudicate between alternative theories of the time course of structure building.

Fig. 3 schematically illustrates a bottom-up CCG analysis for sentence (1). Again, CCG assumes a coordination schema of the type described previously (*VP → VP CONJ VP*). Therefore, to process a VP coordination structure, a predominantly right-branching analysis of the sentence needs to be made—for example, although *the pilot* can initially combine with *embarrassed* to initiate a left-branching structure, this analysis does not yield the VP node required by the coordination schema and, thus, cannot become part of the overall derivation. Thus, the bottom-up procedure implies a delay in the attachment of the two VP conjuncts, because the input to the schema requires two complete VPs on the left and the right of conjunction, respectively. For example, at the point where the word *himself* is processed, the second conjunct is still missing a predicted prepositional phrase, with the result that the phrase has the type VP/PP rather than the VP type required by the coordination schema. In fact, the verb phrase only be-

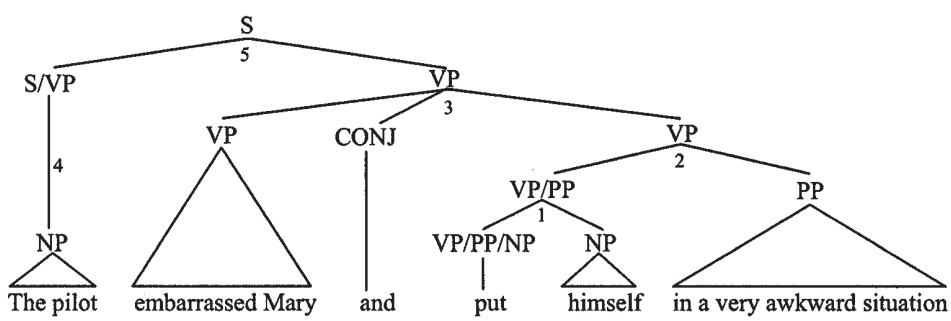

Fig. 3. CCG derivation of VP coordination in phrase structure terms

comes potentially complete at the final word, where Operation 3 (Fig. 3) applies the coordination schema to combine the two VPs and Operation 5 finally combines the subject NP with the conjoined VP.[5]

At first blush, it may appear intuitively obvious that people do not delay the attachment of conjoined phrases until the end of the second conjunct, because a second conjunct may be arbitrarily long, and it is implausible to believe that people wait for an arbitrarily long time before making an attachment. Thus, there is a question about whether it is necessary to test the bottom-up CCG hypothesis experimentally at all. However, this intuition becomes less clear when we consider that the crucial factor for the bottom-up CCG model is whether a phrase can be interpreted as *potentially* complete at a certain point in the string. For example, consider the following:

(2) The pilot embarrassed Mary and criticized himself for being rude.

In (2), the end of the second VP conjunct coincides with the end of the sentence, but the VP can *potentially* end at the word *himself,* as the material following this word is grammatically optional. This means that during processing, the string *criticized himself* can be temporarily assigned the type VP. At this point in processing, therefore, the bottom-up CCG model predicts that the VP coordination schema can in fact be used, allowing a connected representation for the sentence up to the word *himself.* Although this structure will not turn out to coincide with the analysis of the whole sentence (which will have to incorporate the optional phrase *for being rude*), the bottom-up CCG model does indeed predict a high degree of eagerness in structure building for (2). Thus, to provide evidence against the bottom-up CCG model, using VP coordination, we need to demonstrate (a) that structural relations are available at a certain point in the string and (b) that the second VP conjunct cannot be analyzed as potentially complete at that point. Such a demonstration requires experimental methods.

4. Experiment

The two models illustrated previously make different predictions in the time course of processing VP coordination, in particular on the degree of eagerness required. To explore the actual degree of eagerness, we designed an eye-tracking experiment. The experiment exploits the fact that, as pointed out by Schneider (1999), using a similar example, strategies with a low degree of eagerness predict that binding cannot immediately take place between the reflexive *himself* and its antecedent *the pilot* in (1), because the structural relations on which binding relies will not be available until the end of the second VP. We remain agnostic with respect to the exact definition of these structural constraints (e.g., whether based on c-command (Chomsky, 1981), o-command (Pollard & Sag, 1994), co-argumenthood (Reinhart & Reuland, 1993), or scope in logical formulas (Steedman, 2000).

4.1. Design and stimuli

The experiment used eye-movement recording during reading to gain a picture of the time course of processing. The design involved a manipulation of stereotypical gender (Carreiras,

Garnham, Oakhill, & Cain, 1996; Osterhout, Bersick, & McLaughlin, 1997). Consider (1 ′), for example:

The pilot embarrassed Mary and put herself in a very awkward situation.

Here, the morphological gender of *herself* does not match the stereotypical gender of *pilot,* which is usually interpreted as masculine. This means that processing disruption should be found when people read (1′) compared with a condition in which the reflexive matches the stereotypical gender (i.e., *himself;* see also Sturt, 2003). Note that the use of stereotypical gender allows us to avoid the use of ungrammatical sentences.

According to standard assumptions, the detection of such processing difficulty should imply that the relevant structural relation between the reflexive and the antecedent has been established. For example, in the framework of Chomsky (1981), *the pilot* c-commands *himself* (i.e., every branching node dominating *the pilot* also dominates *himself*), and the two NPs are also in the same local domain (i.e., they both appear in a clause containing a governor for *himself* and a subject). These conditions allow *the pilot* and *himself* to be coindexed under Principle A.

However, it is important to control for other explanations of this potential effect. For example, processing difficulty at *herself* in (1′) could occur if the reflexive is simply associated with the first mentioned character (Gernsbacher, Hargreaves, & Beeman, 1989). It is possible that this could occur during the early stages of processing, and at a superficial level of analysis, in which case, processing difficulty associated with the initial reading of *herself* in (1′) could occur even in the absence of a structural relation between the anaphor and *the pilot.* For this reason, the experiment included control conditions using simple pronouns (*him* and *her*), which also either matched or mismatched the stereotypical gender of the subject. If the early influence of gender match occurs regardless of structural relations, then it should occur equally with the simple pronoun and with the reflexive. In contrast, if the difficulty of (1′) is due to structural configurations, then there should be no difficulty when a simple pronoun mismatches in gender with *the pilot,* because this is not a possible antecedent for the pronoun according to binding theory (this is because the same structural relations that hold in the reflexive case—c-command and locality—*prevent* coindexation between *the pilot* and a simple pronoun).

The full design is illustrated as follows:

(3a) Reflexive/Subject match
The pilot embarrassed John and put himself in a very awkward situation.
(3b) Reflexive/Subject mismatch
The pilot embarrassed Mary and put herself in a very awkward situation.
(3c) Pronoun/Subject match
The pilot embarrassed John and put him in a very awkward situation.
(3d) Pronoun/Subject mismatch
The pilot embarrassed Mary and put her in a very awkward situation.

All conditions included a second character (*John* or *Mary*), which always matched with the gender of the anaphor. This was always a grammatical antecedent in the pronoun conditions, both in terms of gender matching and in terms of binding theory (because *John*/*Mary* does not

c-command the pronoun). Thus, all conditions allowed a grammatical antecedent for the reflexive/pronoun. The gender manipulation of the second character (*John* vs. *Mary*) was mirrored in the reflexive conditions to maintain balance in the design (Sturt, 2003, showed that the gender of ungrammatical antecedents has little effect on the early stages of the resolution of reflexive binding in reading).

Twenty-four stimuli were constructed on the model of (3). Half of the stimuli used stereotypically male nouns (such as *pilot*) and half used stereotypically female nouns (such as *nurse*). These nouns were the same as those used by Sturt (2003).

All materials were constructed so that the sentence could not grammatically end at the reflexive/pronoun, as judged by an offline pretest (described later). For example, in (3), the use of the ditransitive verb *put* places a very strong requirement for extra words following the anaphor. As we have mentioned previously, this consideration is crucial for making a proper experimental test of the bottom-up CCG proposal.

4.1.1. Material evaluation

For reasons discussed previously, to evaluate the bottom-up CCG model, we need to demonstrate that the second conjunct cannot be analyzed as potentially complete at the point of the reflexive/pronoun. To test this, 18 members of the University of Glasgow community participated in a continuation task. Each experimental item, from the first word to the reflexive (inclusive), was embedded in a preposed subordinate clause, forming a sentence fragment, as in

(4) After the pilot put himself …

The subordinate clause was used to force participants to continue the sentence in all stimuli. The dependent variable was the proportion of responses in which the subordinate clause was continued before the participant continued the main clause (e.g., "After the pilot put himself *in a difficult situation,* he was admired by the passengers"). To control for any overall bias in the frequency of continuations of the subordinate clause, each experimental fragment was paired with a control sentence fragment, where according to our own judgments, the subordinate clause clearly could grammatically end at the reflexive, as in (5), which uses a monotransitive verb:

(5) After the pilot embarrassed himself …

The experimental and control sentence fragments were printed in booklets, in a pseudorandom order, such that no experimental fragment appeared in the same half of the experiment as its corresponding control fragment. The participants were asked to write the first continuation that occurred to them.

Analysis of responses showed that experimental subordinate clauses were continued in 96% of the trials, whereas the control subordinate clauses were continued in only 13% of the trials. This difference was present for each individual item pair (N =24) and for each individual participant's responses (N = 18), (sign test: ps < .001). This supports the conclusion that the experimental sentences could not terminate at the reflexive pronoun, or were at least very strongly biased against this, and, therefore, that the bottom-up CCG model cannot predict a fully connected representation at this point.

4.2. Eye-movement measures

All psycholinguistic theories predict that processing difficulty should at some point be found in (3b) relative to (3a), after the reflexive is read, and that such a difference should not be found between the two pronoun conditions (3c) and (3d). This should lead to a statistical interaction between the two factors of *subject match* and *anaphor type.* The important consideration is *when* this effect manifests itself in the eye-movement record. A number of different eye-movement measures are available, which allow different inferences to be made about the timing of cognitive processes. Fig. 4 illustrates the five different eye-movement measures that we employ in this article, in terms of the analysis region corresponding to the word *himself.*

The *first-fixation* measure is the time of the first fixation on the region of interest. Because mean fixation duration in reading is generally only around 250 msec, a difference between conditions in first-fixation duration indicates a very early effect. The *gaze-duration* measure is simply the sum of the duration of all initial fixations, before the gaze moves on to another word, either to left or right. *Right-bounded reading times* are the sum of all fixations on the word before the eye gaze first moves to the right of the word (including those made on the word after a regressive eye-movement to previous words). *Regression path times* are the sum of all fixations that are made (including to the left of the word) before the eye gaze first moves to the right of the word. All of these measures are crucial for this experiment because they are informative about the processing that occurs before the gaze moves beyond the anaphor (i.e., at a point where bottom-up approaches predict a disconnected structure). In contrast, *total time* is the sum of all fixations on a word, including those made after subsequent words have been fixated.

The analyses were conducted on the pronoun/reflexive. Short high-frequency words such as pronouns receive very few initial fixations, so we extended the region of analysis to include not only the space before the pronoun/reflexive, but also the space after it. In cases where no initial fixation was made on the anaphor plus preceding and following space, we iteratively extended the word's left boundary by one character space at a time, until a fixation was found.[6] If no fixation was still found after the left boundary had been shifted by four character spaces, the relevant data point was excluded from analysis. This procedure resulted in a first-pass fixation rate

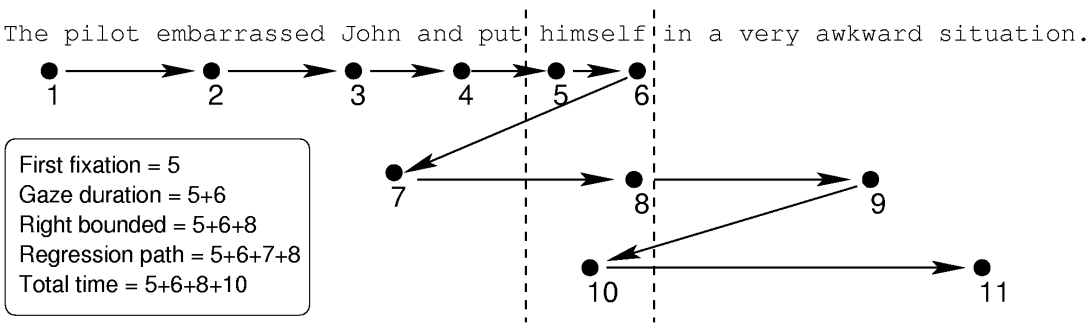

Fig. 4. Example sequence of fixations illustrating eye-movement measures for the critical region (vertical axis is added only for clarity)

of 98% for the reflexives and 87% for the pronouns. The analysis did not include short fixations (less than 100 msec).

4.3. Participants

Twenty-eight members of the University of Glasgow community participated in the experiment. All were native speakers of English and had normal uncorrected vision.

4.4. Procedure

Four stimulus lists were constructed, rotating the four experimental conditions in a Latin-square design. Each stimulus list consisted of the 24 experimental stimuli combined with 48 stimuli from an unrelated experiment on word recognition. Presentation order was random.

The participants were tested individually in a darkened room, using a Fourward Technologies (Buena Vista, Virginia) Generation 5 Dual-Purkinkje-Image eye tracker. The participant's head was immobilized using a bite bar with dental impression compound, and a brief calibration procedure was conducted at the start of the experiment. The whole experiment was usually completed in around 40 min.

4.5. Results

Analyses of variance were computed on the data for the critical anaphor region, on both participant (F_1) and item (F_2) means. *Anaphor type* and *subject match* were treated as within-subjects and within-items factors.

Table 1 shows the means and standard errors for the participant analysis, along with the F ratio for the critical interaction of Anaphor type × Match, for both participant and item analyses. The basic pattern of results is illustrated in Fig. 5, which shows the results for the right-bounded reading time measure.

It can be seen that the critical interaction was fully significant with $\alpha = .05$, for all measures in the participants' analysis, and was significant in the items analysis for all measures except

Table 1
Means (and standard errors) for the four first-pass measures and total time (based on the analysis by participants) and F-ratios for the anaphor × match interaction, for analysis by participants (F_1) and items (F_2)

	First fixation	Gaze duration	Right bounded	Regression path	Total time
Reflexive/Subj-match	241 (7)	260 (8)	272 (8)	305 (15)	342 (14)
Reflexive/Subj-mismatch	260 (10)	288 (16)	313 (16)	358 (22)	438 (31)
Pronoun/Subj-match	260 (9)	272 (9)	276 (10)	310 (23)	295 (12)
Pronoun/Subj-mismatch	261 (10)	270 (10)	270 (10)	284 (10)	306 (14)
Ana × match: F_1 (1,27)	4.52**	4.72**	8.54***	6.36**	6.33**
Ana × match: F_2 (1,23)	1.29^{ns}	3.07*	5.48**	6.89**	13.64***

Note. ns = nonsignificant.
*$p < 0.1$. *$p < 0.05$. *$p < 0.01$.

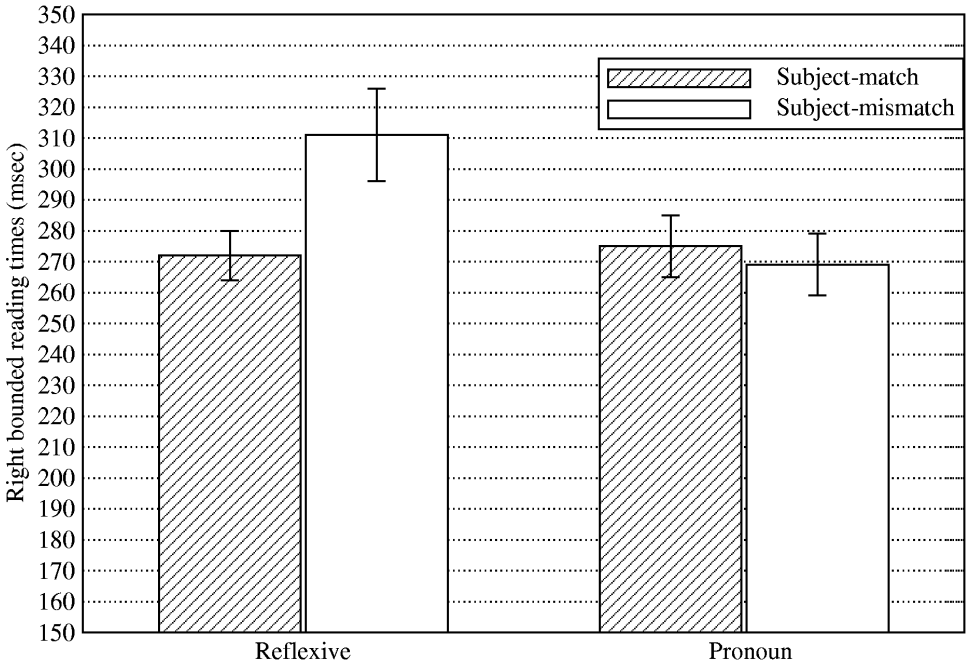

Fig. 5. Right-bounded reading times for the critical region, including standard errors (participants' analysis)

first fixation and gaze duration (although the latter approached significance). The pattern was as predicted: Reading times were longer in the reflexive/subject-mismatch condition than the reflexive/subject-match condition, but there was no reliable difference between the two pronoun conditions. This claim is backed up by *t* tests, which were computed separately on the data for the reflexives and pronouns, respectively. For the reflexives, subject mismatch was slower than subject match, in all measures (first fixation: $p_1 < .05$, $p_2 < .05$, gaze duration: $p_1 < .06$, $p_2 < .05$; right bounded: $p_1 < .01$, $p_2 < .05$; regression path: $p_1 < .05$, $p_2 < .07$; total time: $p_1 < .01$, $p_2 < .001$), whereas the respective pronoun conditions never differed significantly in any measure (all *p*s > .1).

5. Discussion

The experiment shows that the computation of structural relations required for anaphor binding involves full connectivity, even in coordinate structures. In fact, the difference between the two reflexive conditions was found in measures informative of the earliest stages of processing, such as the duration of the first fixation. This demonstrates that the relevant structural relations are available as soon as the reflexive is lexically accessed. A similar first-fixation effect is reported for reflexive binding in Sturt (2003), and the numerical difference between the match and mismatch conditions was of a similar magnitude to that reported for the reflexive conditions here. As that article did not use coordinated structures, the presence of a similar first-fixation effect in the two experiments implies that the relevant structural relations are available equally early, whether or not coordination is used. This result is compatible with

an eager approach to incremental processing and rules out incremental models where the mismatch effect in sentences involving VP coordination is predicted to be delayed because of bottom-up structure building. Therefore, to model the time course of the processing of binding relations, we must turn our attention to those processing models that include a parsing component that keeps a fully connected structure in cases such as the coordination examples used in the experiment. In the introduction, we outlined an approach to the processing of coordination based on the adjunction operation in TAG. We believe that any model of sentence processing will need to employ either adjunction or a mechanism that performs a similar function, such as the approach to coordination based on transitions between processing states (Milward, 1994). However, a TAG approach to coordination needs to supplement the adjunction operation with a tree-unification mechanism to implement the semantics correctly (this is described in Sarkar & Joshi, 1996). Moreover, at the competence level, CCG achieves the best overall coverage of coordination phenomena among computationally well-understood formalisms. One interesting line of research might therefore be to introduce an operation analogous to adjunction into CCG, or a notion analogous to categorial types into TAG.

Alternatively, another line of research would be to encode strong incrementality in a TAG-related formalism that already provides adjunction. Lombardo and Sturt (2002b) proposed a dynamic version of TAG (DV–TAG), which incorporates an eagerly incremental derivation process in the formalism definition (it is a so-called *dynamic grammar*). The DV–TAG derivation process forces input to be eagerly combined word-by-word in a left-to-right order (see also Milward, 1994; Phillips, 2003).

Notes

1. This figure is based on a search of the British National Corpus. Similar figures can be found in other corpora, such as the Turin Treebank (Italian), the Penn Treebank (English), and the Negra corpus (German).
2. Although see Frazier, Munn, and Clifton (2000) for experiments on a related issue.
3. Although the figure shows a flat structure for the coordination, the account is equally compatible with a right-branching structure (Munn, 1999).
4. Notice that this derivation is depicted differently from the usual proof style adopted in the categorical framework. This phrase-structure representation helps in the comparison among the various approaches to incrementality. Note also that the *VP* label is a shorthand for the complex category $S\backslash NP$.
5. To facilitate comparison with Fig. 2, we have included a type-raising operation on the subject NP before combining it with the VP. However, identical conclusions apply whether or not this step is included.
6. Similar procedures were used by Sturt (2003) for first fixation and gaze duration.

Acknowledgments

This research was supported by a Royal Society joint project grant awarded to both the authors and by ESRC Grant R000230572 awarded to the first author. We gratefully acknowledge

the assistance of Jacqueline Craig and Jeremy Pacht for running experimental participants. We thank Simon Garrod, Leonardo Lesmo, Alessandro Mazzei, Paola Merlo, Mark Steedman, Whitney Tabor and an anonymous reviewer for comments on previous versions of the manuscript.

References

Abney, S. P., & Johnson, M. (1991). Memory requirements and local ambiguities of parsing strategies. *Journal of Psycholinguistic Research, 20,* 233–250.

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition, 30,* 191–238.

Carreiras, M., Garnham, A., Oakhill, J., & Cain, K. (1996). The use of stereotypical gender information in constructing a mental model: Evidence from English and Spanish. *Quarterly Journal of Experimental Psychology, 49,* 639–663.

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

Crocker, M. W. (1996). *Computational psycholinguistics: An interdisciplinary approach to the study of language.* Dordrecht, The Netherlands: Kluwer Academic.

Cuetos, F., & Mitchell, D. C. (1988). Cross-linguistic differences in parsing: Restrictions on the use of the late closure strategy in Spanish. *Cognition, 30,* 72–105.

Frazier, L. (1978). *On comprehending sentences: Syntactic parsing strategies.* Unpublished doctoral dissertation, University of Connecticut, Storrs.

Frazier, L., Munn, A., & Clifton, C. (2000). Processing coordinate structure. *Journal of Psycholinguistic Research, 29,* 343–368.

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

Gärtner, H.-M., & Michaelis, J. (2003, August). A note on countercyclicity and minimalist grammars. In P. Monachesi, G. Penn, & S. Wintner (Eds.), *Proceedings of the 8th Conference on Formal Grammar: FGVienna* (pp. 103–114), Vienna, Austria.

Gernsbacher, M. A., Hargreaves, D., & Beeman, M. (1989). Building and accessing clausal representations: The advantage of first mention versus the advantage of clause recency. *Journal of Memory and Language, 28,* 735–755.

Joshi, A. K., Levy, L. S., & Takahashi, M. (1975). Tree adjunct grammars. *Journal of Computer and System Sciences, 10,* 136–163.

Kamide, Y., & Mitchell, D. C. (1999). Incremental pre-head attachment in Japanese parsing. *Language and Cognitive Processes, 14,* 631–632.

Konieczny, L. (1996). *Human sentence processing: A semantics-oriented approach.* Unpublished doctoral dissertation, University of Freiburg, Germany.

Lombardo, V., & Sturt, P. (2002a). Incrementality and lexicalism: A Treebank study. In S. Stevenson & P. Merlo (Eds.), *The lexical basis of sentence processing: Formal, computational and experimental issues* (pp. 137–155). Amsterdam: John Benjamins.

Lombardo, V., & Sturt, P. (2002b, May). Towards a dynamic version of TAG. In *Proceedings of the Sixth International Workshop on Tree Adjoining Grammar and Related Formalism*s (pp. 101–110). Università di Venezia, Italy.

Marslen-Wilson, W. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature, 244,* 522–533.

Milward, D. (1994). Dynamic dependency grammar. *Linguistics and Philosophy, 17,* 561–605.

Munn, A. (1999). First conjunct agreement: Against a clausal analysis. *Linguistic Inquiry, 30,* 643–668.

Osterhout, L., Bersick, M., & McLaughlin, J. (1997). Brain potentials reflect violations of gender stereotypes. *Memory and Cognition, 25,* 273–285.

Phillips, C. (2003). Linear order and constituency. *Linguistic Inquiry, 34,* 37–90.

Pollard, C., & Sag, I. (1994). *Head-driven phrase structure grammar.* Chicago: University of Chicago Press.

Pritchett, B. L. (1991). Head position and parsing ambiguity. *Journal of Psycholinguistic Research, 20,* 251–270.

Reinhart, T., & Reuland, E. (1993). Reflexivity. *Linguistic Inquiry, 24,* 657–720.

Roark, B. (2001). Probabilistic top-down parsing and language modeling. *Computational Linguistics, 27,* 249–276.

Sarkar, A., & Joshi, A. (1996, August). Coordination in tree adjoining grammars: Formalization and implementation. In *Proceedings of 16th International Conference on Computational Linguistics* (*COLING–96;* pp. 610–615), Copenhagen, Denmark.

Schneider, D. (1999). *Parsing and incrementality.* Unpublished doctoral dissertation, University of Delaware, Newark.

Shieber, S., & Johnson, M. (1993). Variations on incremental interpretation. *Journal of Psycholinguistic Research, 22,* 287–318.

Steedman, M. (2000). The syntactic process. Cambridge, MA: MIT Press.

Stevenson, S. (1994). Competition and recency in a hybrid network model of syntactic disambiguation. *Journal of Psycholinguistic Research, 23,* 295–321.

Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language, 48,* 542–562.

Tabor, W., & Hutchins, S. (2004). Evidence for self-organized sentence processing: Digging in effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 431–450.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language, 33,* 285–318.

Vosse, T., & Kempen, G. (2000). Syntactic structure assembly in human parsing: A computational model based on competitive inhibition and a lexicalist grammar. *Cognition, 75,* 105–143.