

FUNCTIONAL CONSTRAINTS ON BACKWARDS PRONOMINAL REFERENCE

Catherine L. Harris
Elizabeth A. Bates

Department of Cognitive Science
University of California, San Diego

ABSTRACT

How does the syntax of a sentence constrain speakers' selection of pronominal referents? Drawing on work by functionalist grammarians, we describe the communicative effect of using a pronoun vs. a definite noun phrase, a matrix vs. a subordinate clause, and the simple past tense vs. anterior/imperfective aspect. Our analysis allowed us to predict differences in coreference judgements for the following three sentence types:

- He worked on a top-secret project when John was ordered to quit.
- He was working on a top-secret project when John was ordered to quit.
- When he worked on a top-secret project, John was ordered to quit.

Coreference judgements from 70 speakers supported our predictions and our research program: An adequate characterization of how syntax constrains sentence comprehension requires reference to the communicative functions performed by syntactic forms.

INTRODUCTION

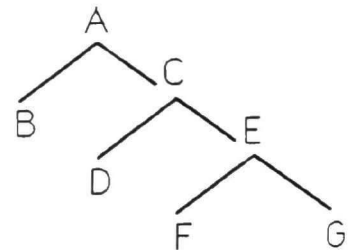
Pronouns are typically used to refer to persons or objects previously identified by name in a sentence or discourse, as illustrated by (1). In some sentences, pronouns can precede their referents, as in (2), while in others, such as (3), native speaker intuitions are that coreference between an initial pronoun and a subsequent lexicalized noun is blocked.

- (1) Harry met Sally when he moved to New York.
- (2) When he moved to New York, Harry met Sally.
- (3) He moved to New York when Harry met Sally.

Because the only change in form across these examples is syntactic, syntax must be responsible for differences in perceived meaning and differences in coreference judgements. How should these changes in syntactic form be characterized, and how do these forms aid listeners in determining the referent intended by speakers?

Linguists have long noted that the relationship between nodes in a phrase-structure tree appears to constrain coreference assignment. Reinhart (1983) has argued that coreference is blocked (prohibited) when an anaphoric element **c-commands** a lexical noun. If the anaphoric element does not c-command the candidate referent, then coreference is free to vary with semantic and pragmatic factors.

A relatively accessible definition of c-command is provided by Radford (1988): A node c-commands its sisters (those nodes on the same level) and their descendants. In the tree at the right, node D c-commands E, F, G, but not A, B, or C. In example (3), *he* is the sentence subject, and thus c-commands *Harry* as well as all other sentence elements. In (2), however, *Harry* c-commands *he*.



Numerous linguistic analyses have questioned whether c-command makes the right predictions in all cases where it is applicable (Bolinger, 1979; Bosch, 1984; Kuno, 1987). For example, because *he* is the sentence subject in (4) and (5), it c-commands *John*. The prediction is thus that coreference should be blocked in both of these sentences. Linguistic intuition is, however, that an interpretation of coreference is odd in (4) but reasonably felicitous in (5). (Examples are from Carden, 1980, cited in Kuno, 1987.)

- (4) Near John, he found a snake.
- (5) Near the girl John was talking with, he found a snake.

From a functionalist perspective, even if c-command (or some other statement of a surface relation between sentence elements) did perfectly divide all utterances into categories such as blocked and free, the job of understanding how syntactic form constrains meaning would be incomplete. What would remain to be developed is an explanation of why this syntactic device (eg., level and direction of embedding) predicts speaker coreference judgements. If the surface devices that comprise c-command prove to be only partially predictive of coreference judgements, we would still want to determine the communicative reasons why this partial correlation between form and meaning is a useful one for selecting pronominal referents.

GENERALIZATIONS ABOUT SOME FORM-FUNCTION MAPPINGS

The sentences in (6)-(9) differ in the grammatical devices they use, but they contain almost identical lexical items and describe roughly the same events. Next to each sentence we have noted our intuitions about whether coreference between the proper name and pronoun is possible.

- (6) **likely:** Bill stayed with some underground writers when he was ordered to leave the country.
- (7) **optional:** When he stayed with some underground writers, Bill was ordered to leave the country.
- (8) **optional:** He was staying with some underground writers when Bill was ordered to leave the country.
- (9) **blocked:** He stayed with some underground writers when Bill was ordered to leave the country.

The obvious changes in form between these sentences are the differences between lexicalized noun (*Bill*) and pronoun, the difference between subordinate and matrix clause, and the difference between imperfective aspect (*was staying*) and the completive aspect which is typically indicated in English with the simple past tense (*stayed*).

In the following sections, we sketch a functionalist explanation of how these differences in syntactic forms lead to differences in coreference judgements. We will then investigate the extent to which speaker judgements of coreference support our hypotheses about form-function mappings.

PRONOUNS AND DEFINITE NOUN PHRASES

What factors influence whether a speaker will use a pronoun or a full nominal to indicate reference to a person or object? Givon (1984) and Prince (1981) have described how pronouns are typically used when the intended referent is the discourse topic, or is otherwise recoverable from context. Use of a pronoun thus tends to be a signal to the listener that the current topic is unchanged: "I'm still talking about the same thing." In contrast, a speaker uses a full lexical noun to indicate a change in topic or when the identity of the pronoun can not be recovered from context. Even if a listener could, with some effort, infer the identity of a referent from the context, a lexical noun may still be used if substantial time has passed since last mention of the entity, or if other discourse characters have been discussed in the interim since last mention of the intended referent.

In the likely assignment example in (6), coding *Bill* as the sentence subject establishes him as the discourse topic. It is thus most natural to interpret the following *he* as referring to *Bill*. In the blocked

case (9), a pronoun (*he*) is the sentence subject and thus established as the current discourse topic.¹ The lexicalization of *Bill* in the following clause signals that *Bill* is new information, and that either a shift in topic has occurred or that the speaker wishes to refer to a discourse entity other than the current topic. The resulting inference is that *he* has an extrasentential reference, and that *Bill* and *he* refer to different people.

This helps explain the difference in referent selection between the **blocked** and **likely** assignment cases, but does not shed light on why backwards pronominalization is allowed when the pronoun appears in a subordinate clause. The next section addresses this point.

MATRIX AND SUBORDINATE CLAUSES

Clauses conjoined with conjunctives/adverbials such as *when, because, before, after, while* are typically viewed by syntacticians to be matrix-subordinate pairs: the clause introduced by the adverbial is the subordinate clause, and the clause beginning with a head noun is the matrix clause. For the purposes of the current paper we will accept this definition of matrix and subordinate.

What is the communicative function of conjoining two events in a matrix-subordinate pair? Matthiessen and Thompson (1988) note that crosslinguistically, matrix-subordinate clause combining involves some type of circumstantial relation: the subordinate clause signals a condition, reason, purpose, cause, setting, manner, or means. The matrix clause is understood to be the main or focal assertion. The subordinate clause thus functions as the *context* in which to interpret or evaluate the speaker's main assertion.

A similar conception of the role of subordinate clauses has been advanced by Chafe (1988), who views matrix-subordinate clause combining as a grammaticization of what he calls *event linking*. He notes that when the subordinate clause is placed before the matrix, the clause linkage is anticipatory: the first clause must be followed by a second clause (barring a change of mind on the part of the speaker). When the matrix clause appears first, however, the listener may not know until the second clause is actually encountered that the utterance is an instance of event linking.

Why might listeners be willing to let a pronoun in a sentence-initial *subordinate* clause -- but not a pronoun in a sentence-initial *matrix* clause -- be coreferential with a lexical noun in a subsequent clause? Because sentence-initial subordinate clauses require a following clause, listeners know that material relevant to the current clause is immediately forthcoming. Because subordinate clauses function as a context in which to evaluate or interpret a main point, listeners will attempt, where possible, to integrate material across the clauses, or to seek reasons why material in the first is relevant to the second.

Why would **speakers** place a pronoun before a referent, violating the convention that a pronoun signals continued reference to an established topic? We doubt that the speaker's *goal* is to violate this convention. Instead, violation should be viewed as a side-effect of other communicative goals. One tendency in English and other languages is that important information, such as establishing the identity of a central discourse participant, is coded with syntactic forms which draw attention to it, such as the sentence subject (Bates and MacWhinney, 1989). How does this tendency to signal importance with sentence-subject coding interact with the subordinate-matrix asymmetry? Speakers place the context or supporting material (the subordinate clause) before the main assertion to produce an effect: it is placed first so that when the information in the main clause is encountered, it is evaluated appropriately. What happens if the supporting material needs to refer to *Bill*? This will be the case if the point of the supporting material is to situate *Bill* in space or time, or to describe some circumstantial or otherwise relevant event. If the topic is identified with a full lexical noun in the supporting material, then the situating force of this material is lessened. Instead of scene setting, the initial clause will be construed as having the

¹Discourse topic is not to be equated with sentence subject. In naturally occurring text, the discourse topic is often so obvious and available that it is either not explicitly mentioned, or it is pronominalized (Givon, 1984). In interpreting sentences which do not have a prior context (such as the first sentence in a text such as a newspaper article) speakers must use what surface cues are available to determine the topic. Sentence subject is the surface device most commonly used to rapidly establish or re-establish a topic.

function of introducing a new discourse character. Example (10), as well as other sentences from Carden's (1982) corpus, illustrate this point.

- (10) While *he* hadn't read the Gifford article, *Associate Dean of Yale College Martin Griffin* said that the 'best administrators are scholars', and that... (*Yale Daily News*, 31, Jan. 78).

Our hypothesis that a subordinate clause licenses backwards pronominalization by virtue of a backgrounding or scene-setting function is complemented by Bolinger's (1979) discussion of loose and tight connectives. He notes that connectives such as *and*, *although*, *before*, and *when* are "loose" connectives: they allow a change in agent or topic, and, when such adverbials are sentence-initial, they are typically accompanied by an intonation break. It is most natural for a speaker to reidentify the referent of a pronoun after a break of some kind.

Thus far, the system of explanation we have developed makes the same blocked vs. free predictions as Reinhart's c-command description: backwards reference is blocked if the pronoun is the sentence subject, but free if it is in an initial subordinate clause. In the next section, we turn to a case where c-command incorrectly predicts blocked coreference. We argue that backwards pronominalization is possible in sentences such as (8) because the imperfective or anterior aspect serves to background the initial clause with respect to a clause encoded with the simple past tense.

THE SIMPLE PAST VERSUS IMPERFECTIVE ASPECT

In this section, we investigate a case of backwards pronominal reference which has not received explicit linguistic analysis (although see Bolinger, 1979; Bosch, 1983). Although the pronoun is the sentence subject in examples (11)-(13), either an intrasentential or an extrasentential reference appears to be possible.

- (11) He had been staring at the control panel for over an hour when Jack received a message from his commander.
- (12) She had just reached the door to her apartment when Lois heard a shriek from the street.
- (13) He was busy looking over the stolen exam when Michael heard someone opening the door.

In English, the simple past tense is typically used to denote discrete, punctate, completed events, as in (14). Events which are encoded with the simple past tense are typically construed as a narrative sequence in which the order of mention reflects the real world order of occurrence. An imperfect marker such as the progressive tense *ing* can be used to indicate that one event overlaps another, as in (15). The past perfect marker *had* signals that an event was completed prior to some reference point, as in (16). We will refer to these types of verbal aspect as **marked** to distinguish them from clauses containing the simple past tense.

- (14) He walked to the store, decided what he wanted, and returned home.
- (15) Walking to the store, he decided what he wanted. Upon returning home...
- (16) He had walked to the store before, but this time...

When narrating a story, speakers need to convey to their listeners which events are part of the story line, and which events are to be interpreted as *comments* or *amplifications* of the story line. Drawing on an analysis of speakers of 20 different languages describing a 6 minute silent film, Hopper (1979) argues that the language of the story line (what he calls the **foreground**) differs from the language of the supporting material (the **background**). One important device for signaling the narrative foreground is the use of the completive aspect (the simple past tense in English). In contrast, when one event is simultaneous with another, or when an event is static or descriptive, the event can be construed as narrative background.

The sentences in (11)-(13) were problematic for c-command because the subject of the sentence is a pronoun, and thus c-commands all other nodes in the tree. But from the perspective of Hopper's analysis of verb aspect in discourse, the initial clause of each of these sentences is a backgrounded event. We

thus hypothesize that coreference is possible in these sentences for the same reasons that coreference is possible with sentence-initial subordinate clauses: because the initial clause functions as the setting or context for evaluating the material in the following clause.

SPEAKER JUDGEMENTS

In the previous section we demonstrated that the canonical communicative functions of clause type (matrix vs. subordinate), NP type (pronoun vs. full noun), and verbal aspect (simple past versus marked aspect) could be related to our linguistic intuitions about coreference assignment. Cases of marked aspect are unusual: backgrounding is achieved in a matrix clause by describing an action with imperfective or anterior aspect. Would a survey of speakers, questioned on a substantial number of such sentences, support our intuitions that verbal aspect functions to background the event, thus allowing backwards pronominalization? Is manipulating aspect as valid or strong a cue to backgrounding as syntactic subordination?

EXPERIMENTAL MATERIALS

In the exploration of constraints on intrasentential coreference, linguists typically make judgements of blocked versus free coreference for single, isolated sentences. As mentioned previously, the most common function of pronouns is to refer back to already established discourse entities (Givon, 1984). A very stringent test of speakers' willingness to allow a pronoun to precede its antecedent would thus be to give speakers a choice between discourse characters occurring before as well as after the pronoun.

We composed 61 three-sentence passages in which the first sentence introduced two characters and the second sentence topicalized one of them. These two sentences functioned as the discourse context for the third sentence, called the **target**. For each of the 61 passages, there were three versions of the target sentence, designed to correspond to the sentence types in (7)-(9). In the **matrix** sentences, the initial clause was not introduced by an adverbial, and occurred in the simple past tense. In the **subordinate** cases, the initial clause was prefaced with an adverbial and contained simple past tense. In the **aspect**, the initial clause was the matrix clause of the sentence, and contained anterior or imperfective aspect.

CONTEXT

Kenneth and Andrew helped prepare each other for the big sumo wrestling match. Andrew weighed in at 193 kilos.

THREE VERSIONS OF TARGET SENTENCE

He put on the wrestler's outfit just as... (**matrix**)
 As he put on the wrestler's outfit... (**subordinate**)
 He had just put on the wrestler's outfit when... (**aspect**)

... Kenneth/Andrew learned the match was canceled.

In each target sentence, either of the two discourse characters could be named. In the **relexicalization condition** the character who had been established as the discourse topic in the second sentence was renamed in the target sentence. In the **new name condition**, the non-topic was lexicalized in the target sentence.

Thirty-four raters were asked to read the 61 passages (along with filler passages) while sitting at a computer display. Each subject saw only one version of a passage. A question about the target sentence and a list of response options appeared simultaneously with each passage. For the passage above, the question and response options were as follows:

Question: Who got into wrestling gear?

Options:	Kenneth Best	Kenneth Better	Either	Andrew Better	Andrew Best	Someone Else	Ungrammatical
----------	-----------------	-------------------	--------	------------------	----------------	-----------------	---------------

An additional 36 raters were asked to make reference judgements to just the target sentence from the 61 passages (the **no context** condition). Instead of being seated at a computer display, raters were

given a paper-and-pencil questionnaire. Because only the target sentences were to be rated, the response options included only one proper name. Response options for the target sentence described above were as follows:

Options:	Kenneth	Kenneth	Either	Someone	Ungrammatical
	Best	Better		Else	

Our two factors were thus sentence type (three levels: **matrix**, **aspect**, **subordinate**) and discourse condition (three levels: **relexicalization**, **new name**, **no context**.)

HYPOTHESES

The **aspect** sentence type resembles the **matrix** in syntactic form, but resembles the **subordinate** in terms of its foreground-background structure. Will coreference judgements to the **aspect** targets be more similar to the pattern for the **subordinate** case or the **matrix** case? Our hypothesis is that the **aspect** sentences will be midway between the two. The reason for this is that these sentence types contain conflicting cues: the sentence-initial pronoun is a cue to extrasentential reference, while the verbal aspect, as a cue to backgrounding, signals the listener to delay pronominal selection until its relevance of the following clause is assessed.

We predict that the **no-context** condition will elicit a higher frequency of intrasentential referents than the **relexicalization** and **new name** conditions. This prediction comes from our characterization of the discourse function of pronouns: pronouns are used to refer to established or known discourse entities. In the absence of syntactic or semantic cues that the pronoun refers to a following NP, raters will try to find an extrasentential referent for the pronouns in our sentences. But without a previous context, raters will have to construct a context and populate it with a candidate referent. The "principle of least effort" suggests that raters will want to *minimize* the number of hypothesized discourse entities (Prince, 1981). Without a context, speakers will, if at all possible, attempt to find a reading of the sentence in which the intrasentential discourse character is the referent for the pronoun.

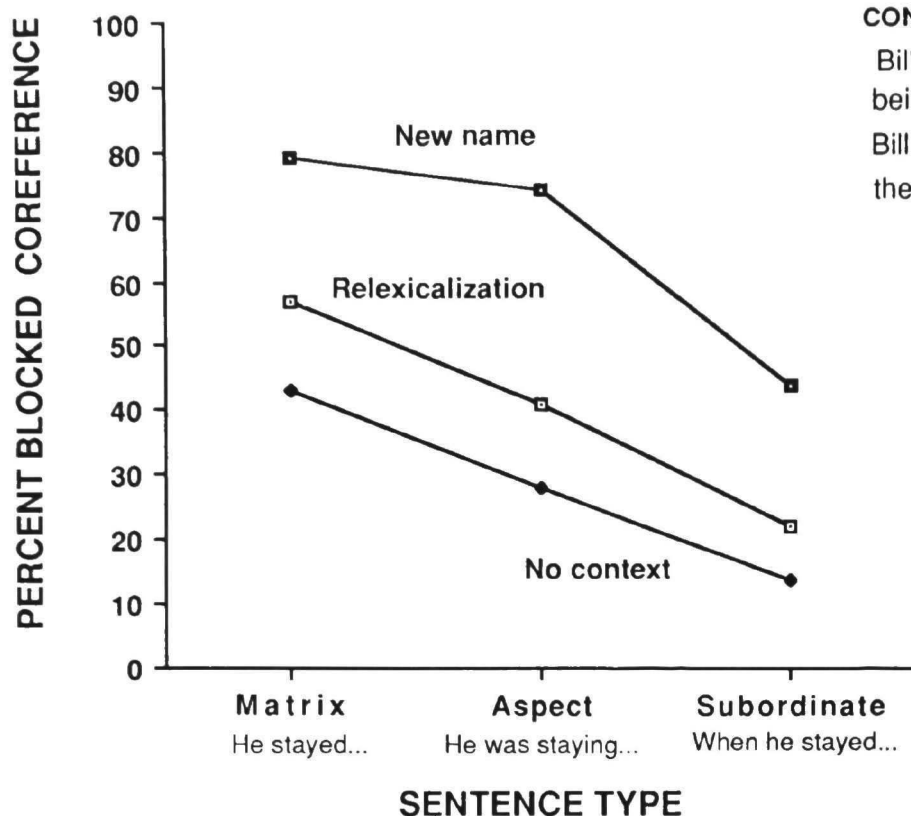
If this prediction is supported, it implies that judgements of blocked vs. free reference for sentences in isolation provide an inflated estimate of the conditions under which backwards pronominal reference is possible. Including a **no context** condition in the current study allows us to compare judgements of intrasentential reference for our three sentence types under two different discourse environments: an environment which encourages extrasentential reference (the **new name** condition), and an environment which encourages intrasentential reference (the **no context** condition).

RESULTS

For each passage, we calculated the percent of judgements in which raters decided that something *other* than the intrasentential proper name was the referent for the pronoun. Selection of the non-intrasentential proper name, or the "someone else" response, or the "ungrammatical" response was coded as a judgement that coreference between the pronoun and intrasentential NP was blocked (or non-preferred). ("Ungrammatical" responses were rare: only four passages received more than six such judgements.)

Figure 1 shows the percent of passages in which raters judged intrasentential reference to be blocked or non-preferred. An item-analysis anova revealed highly reliable main effects for sentence type, $F(2,120) = 72, p < .001$, and discourse condition $F(2,120) = 201, p < .001$, as well as a sentence type X discourse condition interaction, $F(4,420) = 5.2, p < .001$.

It is not surprising that syntactic differences in sentence type and discourse condition result in significantly different reference judgements. Of greater interest is quantifying how the *interaction* of cues affects reference selection. Within the **no context** condition, **matrix** sentence types received more blocked coreference judgements than the **aspect** sentences, $F(1,60) = 18, p < .001$, and **aspect** sentences



CONTEXT:

Bill and John were excited about being in the Soviet Union. Bill had always wanted to meet the dissident community.

New name: He ... John ...

Relexicalization:
He ... Bill ...

FIGURE 1: Percent of target sentences in which raters judged intrasentential reference to be blocked or non-preferred.

received more blocked judgements than **subordinate** sentences, $F(1,60) = 22, p < .001$. Comparable F values were obtained for the **relexicalization** condition. In the **new name** condition, however, the **aspect** sentences did not differ significantly from the **matrix** sentences, $F(1,60) = 2.1, p = .15$, although the **aspect** sentences did differ significantly from the **subordinate** sentences, $F(1,60) = 51, p < .001$.

DISCUSSION

In the first part of this paper we drew on work by functionalist grammarians to describe the communicative function of pronouns compared to lexical nouns, of subordinate clauses compared to matrix clauses, and of completive aspect compared to imperfective and anterior aspect. In the absence of other compelling syntactic or semantic information, listeners will look to the previous discourse context for an antecedent to a pronoun (Givon, 1984; Prince, 1981). What types of information would compel listeners to suspend this default strategy? We hypothesized that both syntactic subordination and durative/anterior aspect cause information to be perceived as background, context or supporting material. Because the listener knows that relevant information immediately follows a backgrounded clause, the default strategy of looking for an antecedent in the previous discourse may be relaxed.

Our survey of raters' judgements strongly supported our hypotheses. When provided with a backgrounding cue (either an aspect cue or syntactic subordination) subjects were more likely to allow a pronoun to precede its antecedent than when no backgrounding cue was present. As we predicted, however, aspect was not as strong a cue as syntactic subordination: in all discourse conditions, more intrasentential judgements were obtained for the **subordinate** sentences than for the **aspect** sentences.

What is the significance of the sentence type by discourse condition interaction? Recall that we analyzed the **aspect** sentences as containing two conflicting cues: the backgrounding cue signals that an

antecedent may be coming up, but coding the pronoun as the sentence subject signals that it refers to established material. In the **no context** and **relexicalization** condition, both cues appear to be contributing to reference judgements, since the percent of blocked/non-preferred intrasentential judgements for the **aspect** case is intermediary between the two other sentence types. Why, in the **new name** condition, was the **aspect** case indistinguishable from the **matrix** case? It appears that when raters are given a legitimate choice between an intrasentential and extrasentential referent, the aspect cue is not strong or salient enough to overwhelm the pronoun-in-subject position cue. However, in a conflict situation -- such as when speakers must decide whether relexicalization of a referent following pronominalization is possible, or must choose a referent in the absence of context -- then the aspect cue does succeed in allowing backwards pronominalization.

Functional approaches to understanding constraints on anaphoric reference have sometimes been represented as attempts to "indicate that coreference is not dependent at all on properties of the syntactic tree" (Reinhart, 1983, p. 94). Our own view -- supported by the data from speakers' judgements presented here -- is that the surface form of a sentence has a strong effect on referent selection. A more accurate characterization of the functionalist position is that speakers have communicative reasons for selecting a particular syntactic encoding (Givon, 1984; Kuno, 1987). An adequate characterization of the constraints on anaphoric reference will thus need to make reference to the communicative functions performed by syntactic forms. Furthermore, the interaction we found between discourse condition and sentence type suggests that syntactic cues do not operate in isolation from each other or from the discourse environment.

In the current paper, we showed that characterizing syntactic forms in terms of their communicative functions allowed us to predict a range of coreference judgements. Future work will be required to better understand the principles constraining the interaction of several syntactic cues and discourse contexts.

REFERENCES

- Bates, E.A. & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney and E. A. Bates (Eds.), *The crosslinguistic study of sentence processing*. New York: Cambridge University Press.
- Bolinger, D. (1979) Pronouns in discourse. In T. Givon (Ed.), *Discourse and Syntax: Syntax and Semantics* 12. New York: Academic Press.
- Bosch, P. (1983). *Agreement and anaphora: A study of the role of pronouns in syntax*. New York: Academic Press.
- Carden, G. (1980). Blocked forward anaphora: C-command the surface-interpretation hypothesis. Presented at the Fifty-fifth Annual Meeting of the Linguistic Society, of America, San Antonio.
- Carden, G. (1982). Backwards anaphora in discourse context. *Journal of Linguistics*, 18, 361-387.
- Chafe, W. (1984). How people use adverbial clauses. *Berkeley Linguistic Society*, 10, 437-449.
- Chafe, W. (1988). Linking intonation rules. In J. Haiman and S. Thompson, (Eds.), *Clause combining in grammar and discourse*. Philadelphia: J. Benjamins.
- Givon, T. (1984). *Syntax: A functional-typological introduction, Vol. 1*. Philadelphia: J. Benjamins.
- Hopper, P.J. (1979). Aspect and foregrounding in discourse. In T. Givon, (Ed.), *Discourse and Syntax, Syntax and Semantics* 12. New York: Academic Press.
- Kuno, S. (1987). *Functional syntax: Anaphora, discourse and empathy*. Chicago: Chicago University Press.
- Matthiessen, C., & Thompson, S. (1988). The structure of discourse and 'subordination.' In J. Haiman and S. Thompson, (Eds.), *Clause combining in grammar and discourse*. Philadelphia: J. Benjamins.
- Prince, E. (1981). Toward a taxonomy of given-new information. In P. Cole, (Ed.), *Radical pragmatics*. New York: Academic Press.
- Radford, A. (1988). *Transformational grammar*. Cambridge: Cambridge University Press.
- Reinhart, T. (1983). *Anaphora and semantic interpretation*. London: Croom Helm.