WEAK CROSSOVER AND THE ABSENCE OF TRACES

Mary Dalrymple, Ronald M. Kaplan, and Tracy Holloway King
Xerox PARC

Proceedings of the LFG01 Conference
University of Hong Kong, Hong Kong
Miriam Butt and Tracy Holloway King (Editors)
2001
CSLI Publications
http://csli-publications.stanford.edu/

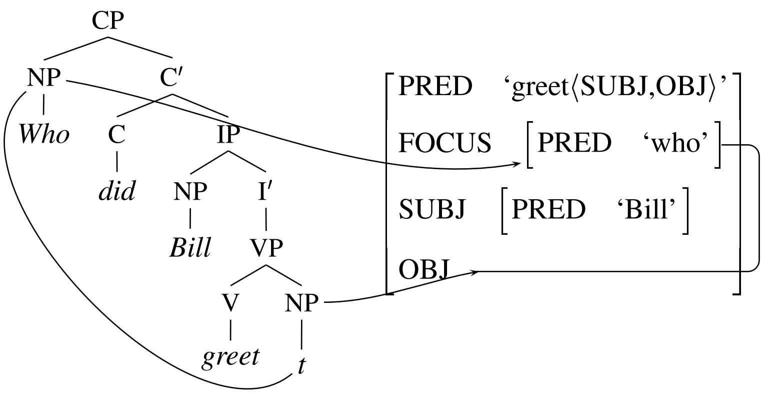
Abstract

We provide a new definition of the linear prominence constraints between pronouns and operators (wh-words and quantifiers) which correctly rules out examples that violate weak crossover. Previous analyses of weak crossover relied on the presence of a trace in the extraction site of a wh-question; in contrast, our analysis enables a traceless account of examples previously cited in support of traces. Since no other incontrovertible evidence for traces has been put forward, our account allows a return to the traceless and therefore more constrained account of long-distance binding in LFG proposed by Kaplan and Zaenen (1989).

1 Long-distance dependencies in LFG

Traces were introduced in transformational grammar (Chomsky, 1973) in order to give a phrasal account for a wide range of long-distance dependency phenomena: a long-distance dependency between a displaced element in a wh-question or relative clause was assumed to arise by movement of the displaced element, which leaves behind a 'trace' in its original position. The original LFG treatment of long distance dependencies (Kaplan and Bresnan, 1982) was based on an adaptation of the transformational/phrase-structure scheme. The relation between the displaced wh-phrase and its within-clause function was defined in terms of a relation between the wh-word and an empty c-structure constituent, a trace, within the clause. In example (1), the relation between the NP node dominating *who* and the NP node dominating its trace, represented by t, ensures that the wh-phrase *who* is both the FOCUS and the OBJ of the sentence:

(1)



Subsequently, however, Kaplan and Zaenen (1989) proposed that constraints on long distance dependencies are best stated in functional and not phrasal terms. As such, *functional uncertainty* offers a more accurate and direct characterization of the long distance dependency. A rule like the one in example (2) establishes two roles for the NP daughter of CP: it is the FOCUS, and it plays a grammatical role defined by the functional uncertainty path COMP* OBJ:

(2) $$\text{CP} \longrightarrow \begin{array}{cc} \text{NP} & \text{C}' \\ (\uparrow \text{ FOCUS}) = \downarrow & \uparrow = \downarrow \\ (\uparrow \text{ COMP* OBJ}) = \downarrow & \end{array}$$

In example (3), the path consisting simply of OBJ is chosen, and the FOCUS f-structure also appears as the OBJ of *greet*:

(3) Who did Bill greet?

$$\begin{bmatrix} \text{PRED} & \text{‘greet}\langle\text{SUBJ,OBJ}\rangle\text{’} \\ \text{FOCUS} & \begin{bmatrix} \text{PRED} & \text{‘who’} \end{bmatrix} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{‘Bill’} \end{bmatrix} \\ \text{OBJ} & \end{bmatrix}$$

With this account, long distance dependencies no longer provide arguments that trace-like devices are *necessary* in the theory of grammar. And in fact, a traceless theory is preferred for two reasons: it is more adequate descriptively, and it offers theoretical advantages. We outline these reasons below before discussing examples involving weak crossover and how a traceless theory can account for them.

A traceless theory is descriptively more adequate. That is, it can describe data that the traceful theory cannot, as argued in detail by Kaplan and Zaenen (1989). Kaplan and Zaenen (1989) cite evidence from wandering adverbs in Icelandic, cross-categorial dependencies in English, and cross-conjunct relativization constraints in Japanese that can be accounted for naturally in a traceless, but not a traceful, analysis. More recently, Sag (1998) has provided additional evidence against traces in his analysis of the conjunct constraint and of floating quantifiers and adverbs.

For several reasons, a traceless theory is also theoretically preferred. That is, even if the traceless and traceful theories accounted for exactly the same data, the traceless theory is preferable. First, the traceless theory is less redundant in that there is only a single way of dealing with long distance dependencies; it is a general scientific principle that less redundant theories are preferred to more redundant ones. Second, it is more restrictive in that it has fewer powerful mechanisms and is more tractable mathematically. Third, it avoids spuriously ambiguous analyses for fillers without ‘canonical’ phrase-structure positions. For example, in extraction of obliques and adverbs there is more than one possible extraction site, even in a relatively fixed word order language like English, as shown in (4).

(4) a. Oblique extr.: About what did John talk (t) to Mary (t)?

b. Adverb extr.: How often did John talk (t) to Mary (t) about apples (t)?

Finally, with the traceless theory, the listener's perceptual problem is simpler: the listener must guess only the function of the displaced element, not both a function and a position. Thus, there are both empirical and theoretical reasons to prefer the traceless account.

2 Weak crossover: A counterexample?

Despite the appeal of a traceless analysis of long distance dependencies, there is evidence that seems to suggest that traces cannot be eliminated: *weak crossover* (Postal, 1971; Wasow, 1979) seems to indicate that the linear position encoded by a trace does affect grammaticality. The sentence in (5) exemplifies a weak crossover violation:

(5) *Who$_i$ did his$_i$ mother greet t?
(cannot mean: Whose$_i$ mother greeted him$_i$?)

The name 'crossover' comes from the transformational analysis of wh-question formation: in a crossover violation, the wh-phrase 'crosses over' a coreferential pronoun when it is moved to the beginning of a sentence. In later analyses, the violation has been reformulated in terms of the position of the trace: a crossover violation ensues when, as in (5), a coreferential pronoun precedes the trace. We concentrate here on *weak crossover* violations, those in which the pronoun precedes but does not c-command the trace.

Crossover effects are found in *operator binding*, in particular, binding by WH-operators and quantifiers (Reinhart, 1983):

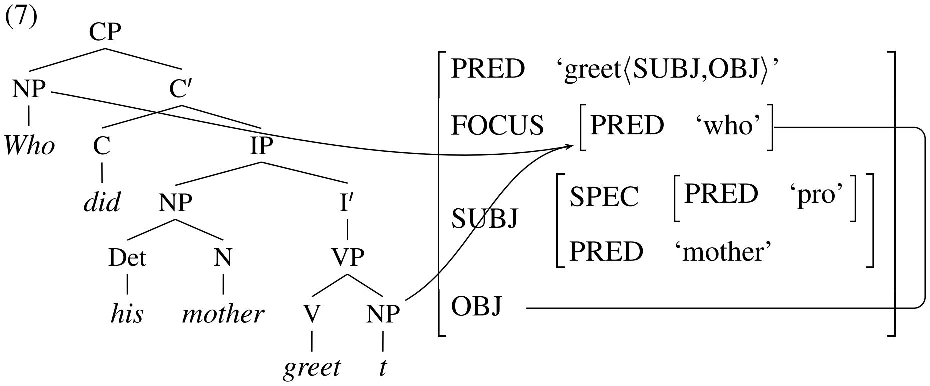
(6) a. *Who$_i$ did his$_i$ mother greet?

b. *His$_i$ mother greeted everyone$_i$.

Our main focus here is on the question of the necessity of traces, and our discussion will center around operator binding in wh-questions. We make only some brief remarks below about quantifier binding in German.

Some recent LFG accounts of operator binding and weak crossover propose a return to a theory of long distance dependency with traces (Bresnan, 1994, 2001; Berman, 2000); these theories adopt the view that the relation between the pronoun and a trace of the displaced wh-phrase is the important one for characterizing weak crossover violations. These theories propose a representation like the one in example (7) for a question like *Who did his mother greet?*, and rule out coreference between the pronoun *his* and the operator *who* by reference to the relation between the position of the pronoun and the position of the trace of *who*:

(7)



In the following, we provide a reformulation of the constraints on the relation between a pronoun and an operator that allow an account of these examples *without* traces.

3 Prominence

Bresnan (1994, 1995, 2001) proposes that operator binding relations are constrained by (at least) two dimensions of *prominence*: syntactic prominence depends on the functional rank of the operator and the pronoun it binds, and linear prominence depends on the linear order between an operator and a pronoun that it binds. We first discuss Bresnan's definitions and then provide a revised set of definitions of these two prominence dimensions.

3.1 Syntactic and linear prominence

Bresnan (2001) defines operator binding requirements in the following way:

(8) **Syntactic Prominence** (A unit containing) the pronoun may not be higher than (a unit containing) the operator on the grammatical function hierarchy:
SUBJ $>$ OBJ $>$ COMP $> \ldots$

Linear Prominence The pronoun must not f-precede the operator.

Bresnan further proposes that languages can vary in which of these constraints apply: some languages impose both kinds of prominence constraints, some require only one or the other, some require a disjunction of the two.

We believe that Bresnan's basic insight is correct: prominence requirements between an operator and a pronoun are multifaceted, and different languages can impose different kinds of prominence requirements between an operator and the pronoun it binds. It turns out, however, that a simple redefinition of the linear prominence condition makes many of the same predictions as the definitions above, including for the crucial cases of weak crossover, without assuming traces.

3.2 Rethinking linear prominence

The intuition behind our reformulation is that linear precedence requirements between an operator and a pronoun are determined by the overt material which indicates the *syntactic role* of the displaced phrase. This intuition follows the proposal of Sag (1998), though it differs in detail.

A theory with traces accounts for the unacceptability of an example like **Who did Sue talk about his mother to?* by reference to the relation between the pronoun *his* and the trace in the object position of the pronoun *to*, as in (9a). In contrast, our theory (like Sag's) refers to the relation between the position of the pronoun and the position of the stranded preposition *to*. The same results obtain, but without positing a trace, as in (9b).

(9) a. Traces: *Who$_i$ did Sue talk about his$_i$ mother to t_i?

b. No traces: *Who$_i$ did Sue talk about his$_i$ mother [to]?

We first define the notion of *coargument* to encompass the arguments as well as the adjuncts of a single predicate:

(10) **Coarguments**:

the arguments and adjuncts of a single predicate.

We also assume the definition of f-precedence in (11), though our proposed redefinition of linear precedence does not depend on adopting this particular formulation of f-precedence as opposed to other definitions of f-precedence that have been adopted in the literature (e.g. Bresnan, 1995).

(11) **F-precedence**:

f_1 f-precedes f_2 if and only if all c-structure nodes corresponding to f_1 precede all nodes corresponding to f_2. (e.g. Kaplan, 1989)

We now present our revised prominence requirements:

(12) Let CoargOp and CoargPro be coargument f-structures such that CoargOp contains O and CoargPro contains P. Then:

Syntactic Prominence An operator O is more prominent than a pronoun P if and only if CoargOp is at least as high as CoargPro on the functional hierarchy.

Linear Prominence An operator O is more prominent than a pronoun P if and only if CoargOp f-precedes P.

The most important difference between our revised definitions and those in (8) is that linear prominence depends on the f-precedence properties of *CoargOp*, an f-structure containing the operator, *not* on the operator.

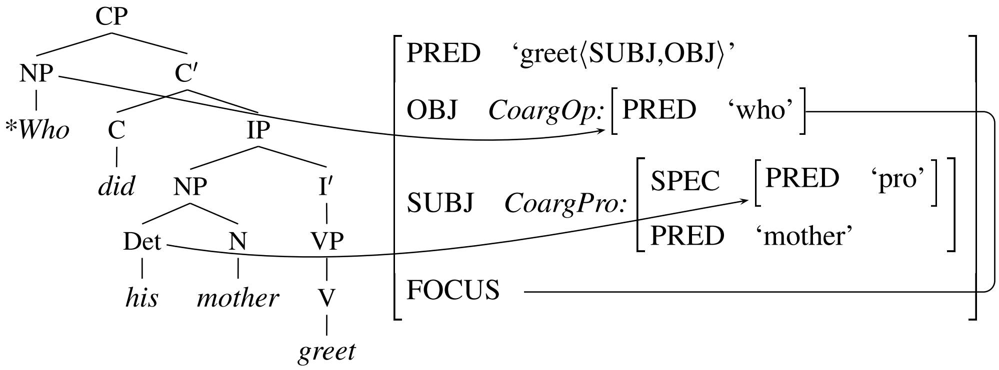
4 Evaluating the proposal

In this section, we show how our revised definitions can account for weak crossover phenomena in English, German, and Malayalam, without positing traces.

4.1 English

Following Bresnan (1995), we assume that in English, the operator must outrank any pronoun it binds in both Linear Prominence and Syntactic Prominence. With this, we correctly predict the unacceptability of example (13):

(13) *Who$_i$ did his$_i$ mother greet?



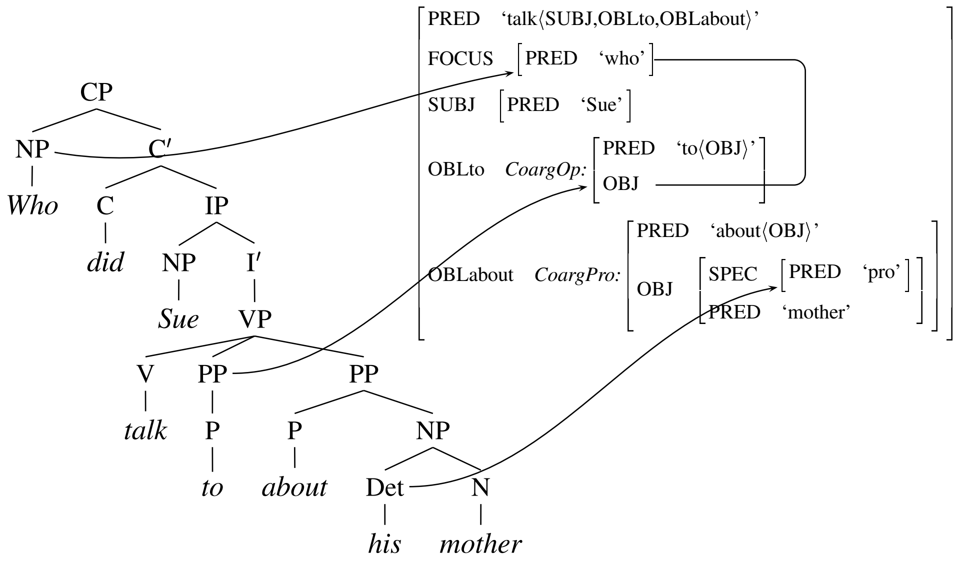
To apply our definitions of prominence, we must first locate the f-structure coarguments that contain the operator and the pronoun. We observe the following:

- SUBJ and OBJ of *greet* are coarguments
- pronoun is *his*
- SUBJ of *greet* contains the pronoun, so it is CoargPro
- OBJ of *greet* contains the operator, so it is CoargOp

On this basis, we check whether both Linear Prominence and Syntactic Prominence are satisfied. In (13), the OBJ *who* f-precedes the pronoun, since every c-structure node corresponding to the OBJ f-structure precedes the pronoun; thus, the Linear Prominence requirement is satisfied. However, the Syntactic Prominence requirement is violated, since CoargPro (SUBJ) outranks CoargOp (OBJ) on the functional hierarchy. This account for the unacceptability of the example.

Next, we examine the grammatical sentence *Who$_i$ did Sue talk to about his$_i$ mother?*:

(14) Who$_i$ did Sue talk to about his$_i$ mother?



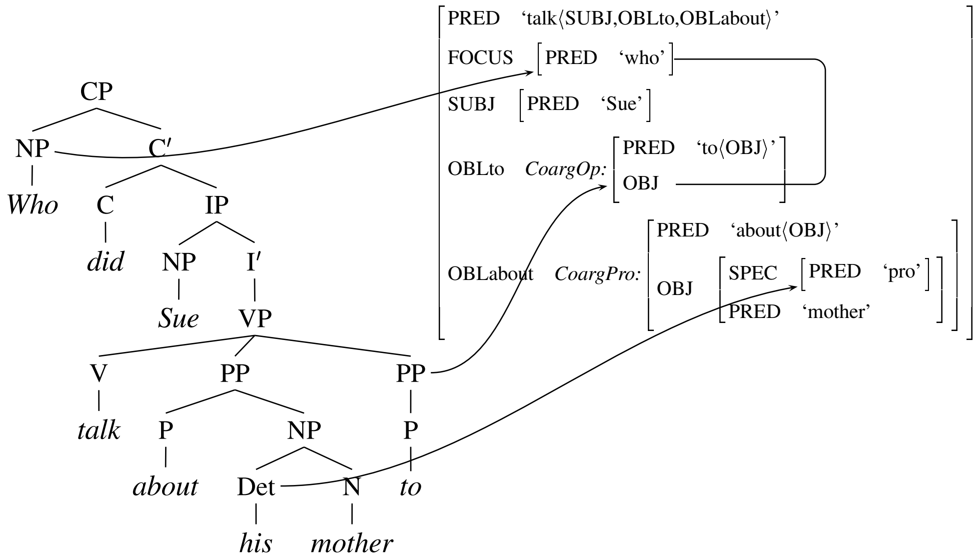
In this example:

- OBLto and OBLabout of *talk* are coarguments
- pronoun is *his*
- OBLabout of *talk* contains the pronoun, so it is CoargPro
- OBLto of *talk* contains the operator, so it is CoargOp

Again, we check to make sure that both Linear Prominence and Syntactic Prominence are satisfied. The syntactic prominence requirement is satisfied: we assume that all obliques occupy the same rank on the functional hierarchy, so that CoargOp (OBLto) is at least as high as CoargPro (OBLabout) on the functional hierarchy. The Linear Prominence requirement is also met. The CoargOp f-structure corresponds to the c-structure nodes PP and P dominating *to*. CoargOp f-precedes the pronoun f-structure, and the Linear Precedence condition is met. Since both requirements are satisfied, the example is correctly predicted to be grammatical.

Finally, we examine the ungrammatical example **Who$_i$ did Sue talk about his$_i$ mother to?*:

(15) *Who$_i$ did Sue talk about his$_i$ mother to?



(15) has the same coarguments as (14). Again, the Syntactic Prominence requirement is met, since the two obliques occupy the same rank on the functional hierarchy. The c-structure nodes corresponding to CoargOp are the PP and P dominating *to*; these nodes do not f-precede the pronoun, the Linear Precedence requirement is not met, and the example is correctly classified as ungrammatical.

Thus, we have seen that traces are not necessary to account for the English weak crossover data once coarguments are taken into account.

4.2 German

Following Berman (2000) and Bresnan (2001), we assume that operator binding in German requires that *either* the Linear Prominence *or* the Syntactic Prominence requirement must be met; unlike English, meeting a single requirement suffices for grammaticality.

First, we consider some patterns of quantifier binding. Both of the examples in (16) are grammatical:

(16) . . . dass [jeder] [seine Mutter] mag

(16)	. . . dass	[jeder]	[seine Mutter]	mag
	that	everyone-NOM	his mother	likes
		(CoargOp, SUBJ)	(CoargPro, OBJ)	

. . . dass	[seine Mutter]	[jeder]	mag
that	his mother	everyone-NOM	likes
	(CoargPro, OBJ)	(CoargOp, SUBJ)	

'. . . that everyone$_i$ likes his$_i$ mother'

In example (17), the quantifier *jeder* 'everyone' is the CoargOp SUBJ, and CoargPro is the OBJ:

- SUBJ and OBJ of *mag* 'like' are coarguments
- pronoun is *seine*
- OBJ of *mag* 'like' is CoargPro
- SUBJ of *mag* 'like' is CoargOp

The Syntactic Prominence requirement is met in both examples, since SUBJ outranks OBJ. Since only a single requirement must be satisfied, the Linear Prominence requirement need not hold, and the SUBJ and OBJ may appear in either order.

When the Linear Prominence requirement is met, the Syntactic Prominence requirement need not be met. In example (17), CoargOp is the OBJ *jeden* 'everyone', and CoargPro is the SUBJ *seine Mutter* 'his mother'. Though the Syntactic Prominence requirement is not met, the CoargOp f-precedes the pronoun, and the example is grammatical:

(17)	. . . dass	[jeden]	seine Mutter	mag
	that	everyone-ACC	his mother	likes
		(CoargOp)		

'. . . that his$_i$ mother likes everyone$_i$'

- SUBJ and OBJ of *mag* 'like' are coarguments
- pronoun is *seine*
- SUBJ of *mag* 'like' is CoargPro
- OBJ of *mag* 'like' is CoargOp

In example (18), which has the same coargument structure of (17), neither requirement is met, and the sentence is ungrammatical:

(18) *...dass [seine Mutter] [jeden] mag
that his mother everyone-ACC likes
(CoargPro, SUBJ) (CoargOp, OBJ)
'... that his$_i$ mother likes everyone$_i$'

We now turn to examples involving long distance dependencies. In example (19), CoargPro is the SUBJ of 'say', *seine Mutter* 'his mother'. CoargOp is the COMP of 'say'. Since SUBJ outranks COMP, CoargOp is not more syntactically prominent than CoargPro. And since CoargOp does not f-precede the pronoun, the linear prominence requirement does not hold either. Since neither requirement holds, the sentence is classified as ungrammatical. Note that the reason for the ungrammaticality of example (19) is not that there is a trace in the subordinate clause; instead, the subordinate clause itself is treated as the important constituent in determining precedence requirements between the operator and the pronoun.

(19) *jeden/wen meinte [seine Mutter], [habe sie getröstet]
everyone/who said his mother has she consoled
(CoargPro, SUBJ) (CoargOp, COMP)
'Everyone$_i$, his$_i$ mother said that she consoled./
Who$_i$ did his$_i$ mother say that she consoled?'

- SUBJ and COMP of *meinte* 'say' are coarguments
- pronoun is *seine*
- SUBJ of *meinte* 'say' is CoargPro
- COMP of *meinte* 'say' is CoargOp

In example (20), we evaluate the prominence conditions with respect to the coargument SUBJ and OBJ of *getröstet* 'consoled': the SUBJ *seine Mutter* 'his mother' is CoargPro, and the OBJ is CoargOp. The Syntactic Prominence condition does not hold, since the CoargPro SUBJ outranks the CoargOp OBJ. But the Linear Prominence condition holds, since the CoargOp f-precedes the pronoun. Thus, the operator binding conditions for German are met, and the sentence is predicted to be grammatical:

(20) [jeden/wen] sagte sie, habe seine Mutter getröstet
everyone/who said she has his mother consoled
(CoargOp)
'Everyone$_i$, she said that his$_i$ mother consoled./
Who$_i$ did she say that his$_i$ mother consoled?'

- SUBJ and OBJ of *getröstet* 'consoled' are coarguments
- pronoun is *seine*
- SUBJ of *getröstet* 'consoled' is CoargPro
- OBJ of *getröstet* 'consoled' is CoargOp

Thus, we see that in German simplex and complex clauses, there is no need to posit traces to account for the weak crossover data.

4.3 Malayalam

Finally, we briefly examine the behavior of 'null pronouns', phonologically unrealized pronominal elements. As shown by Mohanan (1982) and Bresnan (2001), only the Linear Prominence condition is relevant in Malayalam: the CoargOp must f-precede the pronoun, but no syntactic prominence condition is relevant. Example (21) is ungrammaticalwith an overt pronoun:

(21) *in̠n̠ale awar̄e s̄akaar̄icca st̠riikaḷe in̠n̠ə [oor̄oo kuṭṭiyum] sahaayiccu
yesterday they scolded woman today each child helped
(CoargOp)
'Today each child$_i$ helped the woman who scolded them$_i$ yesterday.'

- SUBJ and OBJ of *sahaayiccu* 'helped' are coarguments
- pronoun is *awar̄e*
- OBJ of *sahaayiccu* 'helped' is CoargPro
- SUBJ of *sahaayiccu* 'helped' is CoargOp

In this example, CoargOp is the quantifier phrase *oor̄oo kuṭṭiyum*. CoargOp does not f-precede the overt pronoun *awar̄e*, and so the Linear Prominence requirement is not met.

However, with a null pronoun, the sentence is grammatical:

(22)	iṉṉale	∅	s̄akaar̄icca	sṯriikaḷe	iṉṉə	[oor̄oo	kuṭṭiyum]	sahaayiccu
	yesterday	(they)	scolded	woman	today	each	child	helped
						(CoargOp)		

‘Today each child$_i$ helped the woman who scolded them$_i$ yesterday.’

The same coargument information holds for (22) as for (21). However, the null pronoun does not correspond to any c-structure nodes, and so according to the definition of f-precedence in (11), it vacuously f-precedes and is f-preceded by every other f-structure. In particular, it vacuously satisfies the Linear Prominence condition. This explains the difference between the overt and null pronouns in examples (21) and (22).

5 Empirical differences between trace and traceless accounts

Though the data from English, German, and Malayalam discussed above are classified correctly by our theory of prominence as well as by theories that assume traces, there are empirically testable differences between the two approaches; in particular, the availability of certain sets of data would allow us to distinguish between our traceless account and the accounts of Bresnan (1994), Bresnan (2001), and Berman (2000), which assume traces. Though we have not yet been able to find the crucial data that would allow us to decide between the two theories, we provide characterizations of some of these differences in the following, in hopes that some of these crucial data can be found and examined.

5.1 Language 1

We first examine what we will call Language 1, a language in which the Syntactic Prominence condition does not apply. More specifically, Language 1 has the following characteristics:

- Fixed word order
- Wh-phrases are displaced to the beginning of the sentence
- Only Linear Prominence is relevant

In Language 1, our theory predicts that an example like (23) would be grammatical, since CoargOp precedes the pronoun, and by hypothesis the Syntactic Prominence condition does not apply:

(23)	[who$_i$]	did [his$_i$ mother]	see (t_i)
	(CoargOp, OBJ)	(CoargPro, SUBJ)	

In contrast, the traceful theory predicts ungrammaticality, since the pronoun outranks the (trace of the) operator in Linear Prominence.

5.2 Language 2

We next examine Language 2, a language in which the object precedes the subject. Language 2 has the following characteristics:

- Fixed word order; Object precedes subject
- Wh-phrases are displaced to the beginning of the sentence
- Both Linear Prominence and Syntactic Prominence must hold

In Language 2, our theory predicts grammaticality for example (24):

(24)	[who$_i$]	saw his$_i$ book	(t_i)
	(CoargOp, SUBJ)	(CoargPro, OBJ)	

Here the Syntactic Prominence requirement holds, since CoargOp is SUBJ and CoargPro is OBJ. The Linear Prominence requirement is also met, since CoargOp precedes the pronoun. In contrast, the traceful theory predicts ungrammaticality, since the pronoun precedes the trace.

5.3 Language 3

Finally, we examine Language 3, a language in which either prominence requirement must hold. Language 3 has the following characteristics:

- Fixed word order; Subject precedes object
- Wh-phrases are displaced to the beginning of the sentence
- Either Linear Prominence or Syntactic Prominence must hold

Our theory predicts grammaticality for example (25):

(25)	[what$_i$]	saw its$_i$ owner	(t_i)
	(CoargOp, OBJ)	(CoargPro, SUBJ)	

The Syntactic Prominence requirement is not met in this example, since CoargOp is OBJ and CoargPro is SUBJ. However, the Linear Prominence requirement is met, since CoargOp precedes the pronoun. The traceful theory predicts that this example is ungrammatical, however, since once again the pronoun precedes the trace.

6 Summary

Our traceless theory of weak crossover accounts for the data on weak crossover that motivated previous researchers to posit traces in long distance dependencies. We accomplished this by providing a new definition of Linear Prominence which applies to f-structure units that contain the operator. Since we are aware of no other evidence supporting the existence of traces in long distance dependencies, we believe that traces remain unmotivated in the theory of grammar.

Acknowledgments

We are grateful to Yehuda Falk and the audience at LFG2001 for helpful discussion and comments.

References

Berman, Judith. 2000. *Topics in the Clausal Syntax of German*. Ph.D. thesis, University of Stuttgart, Stuttgart, Germany.

Bresnan, Joan. 1994. Linear order vs. syntactic rank: evidence from weak crossover. In Katie Beals, Jeannette Denton, Bob Knippen, Lynette Melnar, Hisami Suzuki, and Erika Zeinfeld (editors), *Papers from the Thirtieth Regional Meeting of the Chicago Linguistic Society*. University of Chicago: Chicago Linguistic Society.

Bresnan, Joan. 1995. Linear order, syntactic rank, and empty categories: On weak crossover. In Mary Dalrymple, Ronald M. Kaplan, John T. Maxwell, and Annie Zaenen (editors), *Formal Issues in Lexical-Functional Grammar*, pp. 241–274. Stanford University: CSLI Publications.

Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Oxford: Blackwell Publishers.

Chomsky, Noam. 1973. Conditions on transformations. In Stephen Anderson and Paul Kiparsky (editors), *A Festschrift for Morris Halle*. New York: Holt, Reinhart, and Winston.

Dalrymple, Mary, Ronald M. Kaplan, John T. Maxwell, III, and Annie Zaenen (editors). 1995. *Formal Issues in Lexical-Functional Grammar*. Stanford University: CSLI Publications.

Kaplan, Ronald M. 1989. The formal architecture of Lexical-Functional Grammar. In Chu-Ren Huang and Keh-Jiann Chen (editors), *Proceedings of ROCLING II*, pp. 3–18. Reprinted in Dalrymple et al. (1995, pp. 7–27).

Kaplan, Ronald M. and Joan Bresnan. 1982. Lexical-Functional Grammar: A formal system for grammatical representation. In Joan Bresnan (editor), *The Mental Representation of Grammatical Relations*, pp. 173–281. Cambridge, MA: The MIT Press. Reprinted in Dalrymple et al. (1995, pp. 29–130).

Kaplan, Ronald M. and Annie Zaenen. 1989. Long-distance dependencies, constituent structure, and functional uncertainty. In Mark Baltin and Anthony Kroch (editors), *Alternative Conceptions of Phrase Structure*, pp. 17–42. Chicago University Press. Reprinted in Dalrymple et al. (1995, pp. 137–165).

Mohanan, K. P. 1982. Grammatical relations and clause structure in Malayalam. In Joan Bresnan (editor), *The Mental Representation of Grammatical Relations*, pp. 504–589. Cambridge, MA: The MIT Press.

Postal, Paul M. 1971. *Cross-Over Phenomena*. New York: Holt, Rinehart, and Winston.

Reinhart, Tanya. 1983. *Anaphora and Semantic Interpretation*. Chicago: The University of Chicago Press.

Sag, Ivan A. 1998. Without a trace. MS, Stanford University.

Wasow, Thomas. 1979. *Anaphora in Generative Grammar*. Ghent: E. Story.

Addresses:
{dalrymple,kaplan,thking}@parc.xerox.com
Xerox PARC, Palo Alto, CA 94304 USA