

**DO WE WANNA (OR HAFTA) HAVE
EMPTY CATEGORIES?**

Yehuda N. Falk

The Hebrew University of Jerusalem

Proceedings of the LFG07 Conference

Miriam Butt and Tracy Holloway King (Editors)

2007

CSLI Publications

<http://csli-publications.stanford.edu/>

Abstract

This paper revisits the relevance of *wanna* contraction for the existence of empty categories. An analysis of *wanna* contraction is proposed under which empty categories have a role in blocking contraction. It is then shown that alternative analyses (subject sharing, local c-command, and morpholexical analyses) are inadequate. It is suggested that empty categories are a last resort, which can only appear in non-subject LDD constructions because these have to be licensed by inside-out functional uncertainty equations, and the empty categories are needed as the c-structure positions on which these equations are annotated.

1. Prologue

The existence of empty categories is very controversial in LFG. So it is important to search for empirical evidence that bears on the existence of empty categories. While in the LFG literature weak crossover has been the focus of the argument over empty categories (as in Bresnan 1995 and Dalrymple, Kaplan, and King 2001), the most enduring such construction in the broader syntactic literature is contraction, particularly the contraction of *want to* to *wanna*. In this paper, we will evaluate the contraction argument for empty categories.¹

Looking ahead, we will show that other attempts to account for the contraction facts are untenable, and therefore that contraction does provide evidence for empty categories. However, the same evidence that shows that empty categories do exist in a limited set of constructions shows that they do not exist in many other places in which they have been hypothesized in the transformational literature: in particular, they do not exist in long-distance dependency constructions involving subjects, nor do they exist in non-long-distance-dependency contexts. Finally, we will explain why it is not so terrible to have empty categories, and how they are constrained.

2. The Claim

The facts about contraction which are alleged to be relevant for the existence of empty categories were brought to light by Lakoff (1970: 632), who credited Larry Horn with the observation. Simplifying the examples somewhat,² Lakoff observed that contraction of *want to* to *wanna* is possible in (1) but (for most speakers) not in (2).

- (1) a. Who do you want to see?
b. Who do you wanna see?

- (2) a. Who do you want to see Pnina?
b. *Who do you wanna see Pnina?

The basic observation, abstracting away from the specific theoretical assumptions made by Lakoff, is that in (2) the preposed *wh* element bears the function of object of *want*. The canonical structural position of the object of *want* intervenes between *want* and *to* and it is this, Lakoff claimed, that causes contraction to be blocked. In (1), on the other hand, *who* is the object of *see*, and thus has a canonical structural position which does not block the contraction. We refer to this as the Lakoff/Horn Generalization. Using the term “locally licensed function” to refer to a non-discourse grammatical function, we can state the generalization as (3).

¹This paper was presented at the LFG07 conference. Thanks to Irit Meir, Yael Ziv, and participants in LFG07 for comments. This paper is dedicated to the memory of my father Paul Falk פול, who passed away 2 Tevet 5767 (22 December 2006). This research was supported in part by the Israel Science Foundation (grant No. 207/07).

²Lakoff presented the examples in terms of ambiguity of sentences, using an optionally transitive verb. Simple grammaticality is more straightforward.

(3) **Lakoff/Horn Generalization**

Want to cannot contract to *wanna* if the canonical position of the locally licensed function of a preposed element intervenes between *want* and *to*.

If the Lakoff/Horn Generalization is correct, linguistic theory needs a way to express it. Ideally, such an expression should be one in which the Lakoff/Horn Generalization results naturally from the system, rather than having to be stipulated. Empty categories provide a way to do this. The central idea is that the canonical structural position of the locally licensed function of a preposed element is occupied by an empty category, conventionally represented *e*.

- (4) a. Who do you want to meet *e*? (= (1))
b. Who do you want *e* to meet Pnina? (= (2))

In (4b), *want* and *to* are not adjacent; the empty category intervenes. On the plausible assumption that contraction requires adjacency (but see footnote 4), the lack of contraction in (2) follows from the postulation of the empty category.

An empty-category-based approach provides an elegant expression of the Lakoff/Horn Generalization. For this reason, much of the literature which is ostensibly about empty categories focuses on the correctness of the generalization. Following our fleshing-out of an empty-category-based analysis in the next section, we will review the alternative descriptions that have been proposed, and discover that they are all flawed.

3. An Analysis

3.1. *To* Attachment

We begin with the infinitival *to*. As discussed by Jacobson (1982), and in more detail by Zwicky (1982), *to* is phonologically subordinate to adjacent material, with this subordinate status manifesting itself either as becoming part of the phonological phrase (what Zwicky refers to as leaning) or part of the phonological word (cliticization). Zwicky, in particular, shows that various puzzles about the distribution of *to* can be accounted for under such an analysis. We will use the term *attachment* as a way of referring to this phonological subordination which is neutral between leaning and cliticization. Zwicky takes the position that while *to* ordinarily functions as a leaner, in *wanna* contraction it cliticizes. The behavior of *to* in *wanna* contraction is, under this analysis, merely an extreme case of the normal behavior of *to*.

Under normal conditions, *to* attaches to the right, but it is also possible for it to attach to the left, primarily when it is stranded. (Parentheses here indicate phonological phrasing.)

- (5) a. (We're nót) (to léave).
b. (We're nót to).

We propose the following constraint on infinitival *to*:

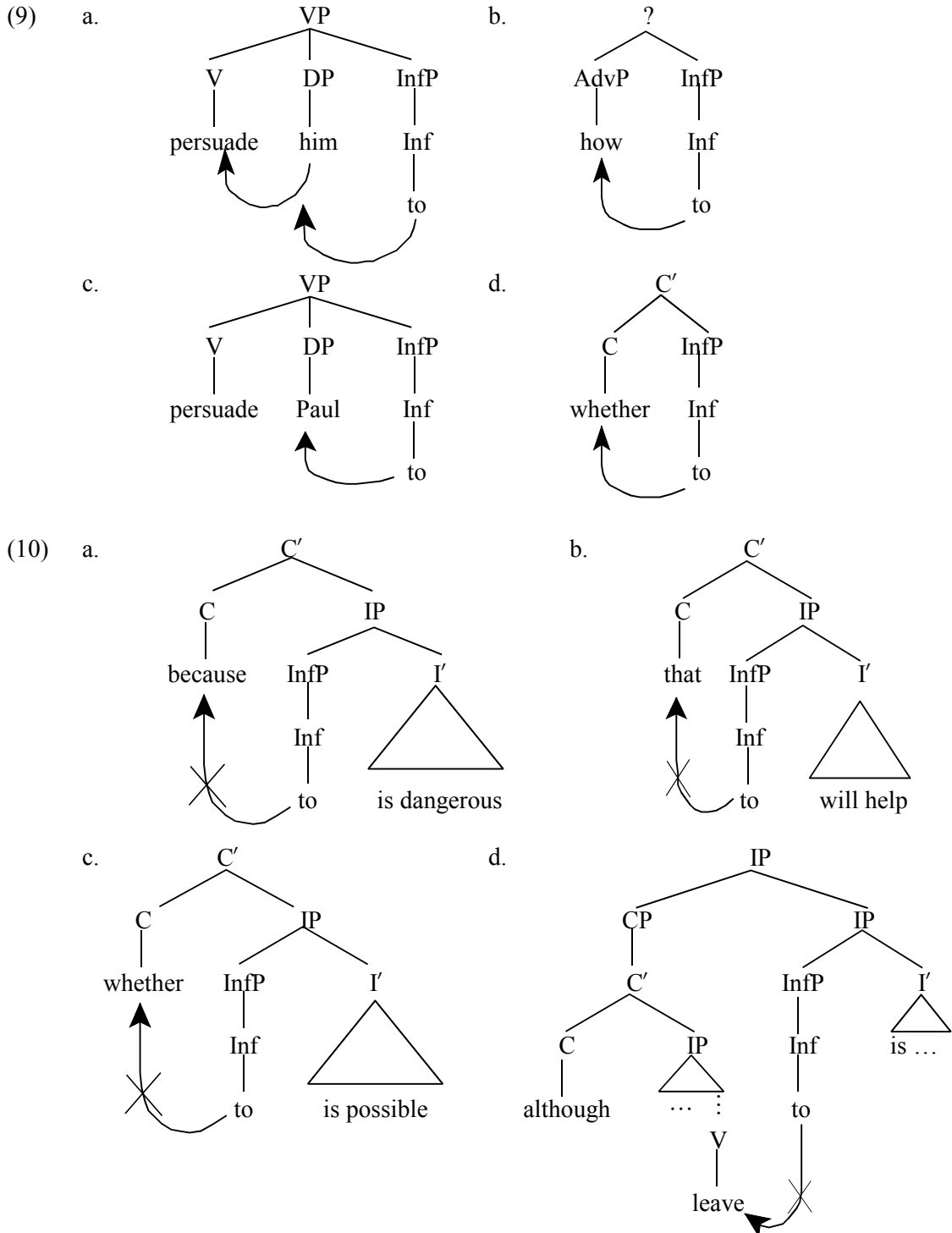
- (6) ***To* Attachment** (first approximation)
Infinitival *to* must attach to an adjacent element. It may attach
(a) to the right (the usual situation)
or
(b) to the left

This statement of *to* attachment needs to be refined somewhat. Zwicky observes that there are various constraints on the ability of *to* to attach to the left.

- (7) a. I don't know if Paul wants to buy the present, but I think we can (persuáde him to).
b. I might whittle a polar bear out of Ivory soap, but I don't know (hów to).
c. I don't know if he wants to buy the present, but I think we can persuade (Pául to).
d. I might whittle a polar bear out of Ivory soap, but I don't know (whéther to).

- (8) a. *You shouldn't play with rifles, (because to) is dangerous.
 b. *You can try to plead with him, but I doubt (that to) will help.
 c. *She'd like to surprise him, but I don't know (whether to) is possible.
 d. Although it would distress us for you to leave, to leave/* \emptyset is what I'd advise you to do.

Note the structures involved here, with arrows indicating cliticization. (For the sake of neutrality, we label the clause headed by *to* as an InfP, and label some other nodes “?”.)



In the grammatical cases, *to* attaches to the left to an element in a very specific structural configuration: the host c-commands *to* and there is no maximal-level phrasal category intervening between the host and

the InfP which *to* heads. We will refer to this as local c-command, and revise the statement of *To Attachment* accordingly.

(11) **To Attachment**

The complementizer *to* must attach to an adjacent element. It may attach

- (a) to the right (the usual situation)
- or
- (b) to the left onto a locally c-commanding host

(12) **Local C-Command**

X locally c-commands Y iff X c-commands Y and no phrasal node other than the maximal projection of Y intervenes.

As noted by Aoun and Lightfoot (1984), many of the non-long-distance-dependency structures in which *wanna* contraction is barred can be blocked if a local c-command condition is placed on *to* attachment. For example, Carden (1983) observes that the inability of *to* to contract with a *want* which is part of a coordinate structure is not mirrored by other contraction rules, such as the contraction of ...*t you* to [čə].

- (13) a. I don't need or want to hear about it. *wanna
- b. I don't expect or want you to get involved. ✓čə

The presence of a local c-command condition on the leftward attachment of *to* will rule out *wanna* contraction in this case. Unlike adjacency, which is a consequence of the concept of attachment, we take local c-command to be a rule-specific stipulation.

3.2. On *Want*

We turn now to the verb *want*. From the perspective of coming to an understanding of the nature of *wanna* contraction, there are two crucial facts: *to* cliticizes onto *want*, instead of merely leaning on it; and the /t/ of *want* deletes.

It is observed by Jacobson (1982) that deletion of /t/ is not a peculiarity of *wanna* contraction; rather, it is a general property of the verb *want*. She observes that the /t/ frequently deletes in forms like *wanted* and *wanting*. Strikingly, the phonetic sequence we are representing as *wanna* is not only a realization of *want to*; it is also a realization of *want a*:

- (14) a. I wanna play. *wanna*=*want to*
- b. I wanna toy. *wanna*=*want a*

The only situation in which the /t/ of *want* is obligatory is in the present subjunctive:

- (15) a. I demand that you want a part in the play.
- b. *I demand that you wanna part in the play.

As noted by Brame (1981: 286 fn 13), *want to* cannot contract to *wanna* in the subjunctive either.

- (16) a. The director requires that all of the actors want to give their most.
- b. *The director requires that all of the actors wanna give their most.

This confirms our view that the deletion of /t/ in *wanna* is no different than deletion of /t/ in other uses of *want*.

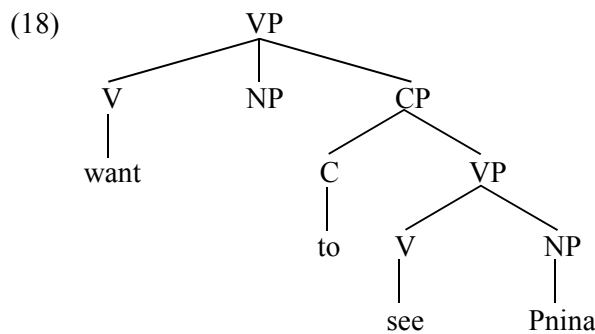
While we do not presume to propose a full phonological analysis of the deletability of /t/, a possible analysis would give *want* a phonological representation in which the last skeletal position is only optionally filled by /t/.

- (17) X X X X
 | | | |
 w a n (t)

It is possible that it is the availability of an empty skeletal position if the /t/ is not included that drives cliticization of *to* to *want*.

3.3. Empty Categories and Contraction

We hypothesize that the canonical position of the locally licensed function of a fronted element is (sometimes, at least) marked by an empty category, an unfilled phrasal node. For example, the VP headed by *want* in (2) has the following structure:



The critical question is what happens in (18) if *to* attempts to attach to the left. The local c-command condition does not block the leftward attachment, since *want* does locally c-command *to*. The condition that is relevant is adjacency, which, as already noted, we take to be an integral property of attachment phenomena. In this case, the applicability of the adjacency condition is not entirely clear. Phonologically, *want* and *to* are adjacent; no phonological material intervenes between them. They are also adjacent at the level of terminal elements in the constituent structure; there is no terminal element that intervenes between them. However, at higher levels of structure they are not adjacent: the unfilled NP intervenes. That is to say, the interpretation of the adjacency condition in a structure with an empty category depends on which part of the structure is relevant. Given the ambiguous status of adjacency in this case, one might expect variability between speakers, with some treating *want* and *to* in (18) as adjacent, and others as not adjacent. Such an expectation would be well founded.

It is well known in the literature on *wanna* contraction that not all speakers share the judgment in (2). For some, often referred to as speakers of the liberal dialect, *wanna* contraction is possible in sentences of this kind. Zwicky (1982) takes this one step further. He notes that, even in the absence of contraction, speakers differ on the acceptability of B's utterance in the following discourse (Zwicky 1982: 26):³

- (19) A: Who do you want *e* to vanish?
 B: %I don't know; who do you want *e* to?

Since *to* must attach to something, and it is stranded on the right, it must attach to the left. If it cannot at least lean, the sentence is ungrammatical. Apparently, some speakers allow *to* to lean onto *want* in this case and others do not; Zwicky reports a 50-50 split among speakers he consulted. The number of speakers who accept *wanna* in these environments is apparently less: Zwicky suggests that not all

³For what it's worth, the author of this paper finds B's response crashingly bad.

speakers who allow leaning allow cliticization.⁴

Empty categories, by their very nature, have an ambiguous status in terms of adjacency. This ambiguity leads to a situation where the Lakoff/Horn Generalization is valid for most speakers, but not all. Strikingly, while the existence of the liberal dialect has been cited by Pullum and Postal (1979) as an embarrassment for an empty-category account of the Lakoff/Horn Generalization, it actually provides exactly the right tools for explaining this inter-speaker variation.

4. Alternative Analyses

4.1. Subject Sharing

The first challenge to the Lakoff/Horn Generalization came from Postal and Pullum (1978). Much of their argument concerns not the Lakoff/Horn Generalization, but rather the implementation of an empty category analysis in the Extended Standard Theory literature of the time. Many of their arguments against this kind of analysis are valid, but irrelevant to other implementations, including ours. However, in the course of arguing against “trace theory,” they propose an alternative to the Lakoff/Horn Generalization.

As Postal and Pullum observe, *want* is not the only verb which contracts with *to*. The following is presented as a complete list by Pullum (1997: 81).

(20)	<i>want</i>	<i>wanna</i>
	prospective <i>go</i>	<i>gonna</i>
	habitual <i>used</i>	<i>usta</i>
	<i>have</i> (necessity/obligation)	<i>hafta</i>
	<i>got</i> (necessity/obligation)	<i>gotta</i>
	<i>ought</i>	<i>oughta</i>
	<i>supposed</i>	<i>supposta</i> or <i>sposta</i>

Postal and Pullum observe that, other than *want*, these are all Raising-to-Subject verbs. *Want*, when contraction is possible, is a Subject Equi verb. What unites all of these cases is that the two clauses share a subject. They therefore propose that contraction is possible when the two clauses share a subject, or, in LFG terms, functional control.

The subject-sharing alternative to the Lakoff/Horn Generalization appears to have never been subjected to critical scrutiny in the literature. A close look reveals several problems. For example, it is only with *want* that there is a contrast between environments that allow contraction and those that do not. It is therefore difficult to draw conclusions about the source of the *wanna* contraction facts from these other verbs. Therefore, contra Postal and Pullum, it is not clear that the other verbs are relevant for determining the conditions under which contraction does and does not occur. Another problem is that the analysis is essentially arbitrary; unlike the empty category analysis, which is based on the idea that an intervening element breaks the contiguity necessary for contraction, there is no inherent relation between subject sharing and contraction.

However, the biggest problem with the subject sharing analysis is that it is empirically incorrect. The context in which *want to* can contract to *wanna* includes cases in which the reference of the subject of *want* is a subset of the reference of the subject of the *to* clause; i.e. cases which cannot be analyzed as functional control.

⁴Postal and Pullum (1978) mention a different idiolectal treatment, which appears to dispense with the adjacency condition. In this idiolect, the following is grammatical:

- (i) I wanna very much go to the game tomorrow.

Without further information, it is difficult to know what lies behind judgments such as this. A very preliminary speculation would be that *to* undergoes prosodically motivated movement, much as second-position clitics do in some languages. On the movement of clitics, see Halpern (1995) and, in LFG, Kroeger (1993).

- (21) a. *I met on Sunday at 10:00
 b. I wanna meet on Sunday at 10:00.
 c. *I hafta meet on Sunday at 10:00.
 d. *I tried to meet on Sunday at 10:00.

As (21a) shows, *I* is not a possible subject for the intransitive verb *meet*, since it requires a plural subject in this subcategorization frame. Thus in (21b), the subject of *meet* cannot be *I*, but rather some group including *I*. This contrasts with a Raising verb like *hafta* (21c), which has to be functional control. This overlapping reference is also not a necessary property of Equi constructions: in the case of *try* there is functional control, resulting in the ungrammaticality of (21d). The overlapping reference that is possible in the case of *want* requires an anaphoric control analysis.

We conclude, therefore, that functional control (subject sharing) cannot be the property that licenses *wanna* contraction. Subject sharing is therefore not a possible alternative to the Lakoff/Horn Generalization.

4.2. Local C-Command

Another alternative that has been suggested in the literature (e.g. Bouchard 1984 and Barss 1995) is that the condition that *want* locally c-command *to* obviates the need for the Lakoff/Horn Generalization. While we have adopted such a condition on the leftward cliticization of *to*, the process that underlies *wanna* contraction, the claim by Bouchard and Barss is problematic.

This proposal has been made within the context of Government/Binding (or Principles and Parameters) theory, in which local c-command, under the name government, is taken to be one of the fundamental structural relations in syntax. The local domain of c-command is said to be delimited by certain nodes. One of these “barriers to government” is the CP node, and it is this which is taken to block contraction when the fronted element functions as the subordinate subject.

- (22) a. *Who do you wanna see Pnina? (=2b)
 b. ... want [_{CP} [_{IP} *e* to see Pnina]]

Under this analysis, contraction would involve the matrix verb *want* and the head of the IP embedded within the CP. The intervening CP node renders the c-command non-local, and contraction is therefore impossible. Under standard GB assumptions, however, the same configuration would obtain in the case of grammatical contraction.

- (23) a. Who do you wanna see? (=1b)
 b. ... want [_{CP} [_{IP} PRO to see *e*]]

The local c-command analysis of contraction has to therefore make an additional assumption: namely that control complements are bare IPs:

- (24) ... want [_{IP} PRO to see *e*]

In this structure, there is no CP barrier between *want* and the IP headed by *to*, and contraction is therefore possible.

The viability of the Bouchard/Barss local c-command analysis depends on the plausibility of the proposed structures, in particular the status of the controlled infinitive as a bare IP and the categorization of *to* as an infl. The latter, while standard in the transformational literature, is not obviously correct; Falk (2001: 154) argues that *to* is a complementizer. If *to* is a complementizer, there cannot be any structural difference between (22b) and (23b); ignoring the positions of possible empty categories, the structures in question would be:

- (25) a. ... want [_{CP} to see Pnina]
 b. ... want [_{CP} to see]

However, even assuming an infl analysis for *to*, the analysis of the controlled complement of *want* as a bare IP seems dubious: its distribution is that of a CP, not an IP.

- (26) a. [To see Pnina] is what I want.
 b. [_{CP} That I might see Pnina] is what I said.
 c. *_{[IP} I might see Pnina] is what I said.
- (27) a. I want very much [to see Pnina].
 b. I said very loudly [_{CP} that I would see Pnina].
 c. *I said very loudly [_{IP} I would see Pnina].

We therefore consider the Bouchard/Barss analysis to be untenable.

4.3. *Wanna* as a lexical item

It has also been proposed that *wanna* is a lexeme distinct from *want*, and that there is no actual contraction in *wanna* sentences. If this is correct, syntactic structures are irrelevant to *wanna*, and what matters is the nature of the morphological relation between *want* and *wanna*. Such analyses have been proposed by several researchers (for example, Brame 1981: 286 fn 13), but the most thorough argument for it is that of Pullum (1997). As we will show here, we find Pullum's argument unconvincing.

The heart of Pullum's argument is that *to* contraction is morphologically and phonologically idiosyncratic. In this respect, the argument mirrors the argument presented by Zwicky and Pullum (1983) that *n't* is an inflectional suffix in contemporary English, and not a contracted form of *not*. However, while the argument is quite compelling in the case of *n't*, it is more problematic with *wanna*. We will discuss the claim of morphological idiosyncrasy first, and then phonological idiosyncrasy.

Pullum's argument for morphological idiosyncrasy is that only a limited set of verbs can contract with *to*, the ones listed in (20) above. For example, while *ought to* contracts to *oughta*, *thought to* does not contract to **thoughta*. Idiosyncrasy of this kind is typical of morphology, not of syntactophonological contraction. However, as we have seen, the situation is more complicated: *to* obligatorily attaches, and when it attaches to the left, some verbs allow cliticization. So while it is true that *to* does not cliticize to *thought* the way it does to *ought*, it does lean on it. Pullum's proposal that forms like *wanna* and *oughta* are lexically derived is incompatible with the analysis of *to* presented earlier. Treating *wanna* as derived by phonological contraction (cliticization) forms a more harmonic part of an overall analysis of the phonological properties of infinitival *to*.

The argument for phonological idiosyncrasy is based on the presence of irregular phonological changes in the form of the host of *to*. For example, Pullum discusses the devoicing of /v/ in *hafta* and observes that in other cases (e.g. *Aztec*) such devoicing does not occur. However, Pullum also notes that /hæf/ appears to have become a new underlying form for the verb which forms the core of *hafta*, at least for some speakers.⁵ Similarly, Andrews (1978: 267) suggests that *usta* and *supposta* have underlying voiceless /s/. He reports the following judgments:

- (28) a. Did they [yuws(t)] not to eat pickles?
 b. *Did they [yuwz] not to eat pickles?
 c. *usen 't* = [yuwsnt], *[yuwznt]
 c. You're [səpowst] not to light the wick until it's wet.

In other words, at least some of the cases Pullum cites may involve lexical reanalysis of the host, rather than lexical attachment of *to*.

The only case which is relevant for testing the Lakoff/Horn Generalization is *wanna*. The phonological change in question is the deletion of /t/ after /n/. As Pullum notes, this is not an automatic phonological rule in English: for example, in *wont to* (Postal and Pullum 1978: 2) and *taunt* (Pullum

⁵He reports that some speakers pronounce *having to* as [hæfɪntu]. Even for those who do not, however, an underlying form /hæf/ in allomorphic alternation with /hæv/ is not inconceivable; alternatively, there are distinct verbs *haf* and *have*.

1997: 90) there is no deletion of /t/. However, as Pullum himself observes, /nt/ does sometimes at least optionally reduce to [n] word-internally (as in *Santa*, *twenty*, etc.). More importantly, as we observed above, following Jacobson (1982), this deletion of /t/ is a general property of the verb *want*. The deletion of /t/ in *wanna* is no different than deletion of /t/ in other uses of *want*, and, contra Pullum, is not an idiosyncratic phonological result of adding *to* to *want*.

A further argument against an analysis in terms of derivational morphology is provided by Hudson (2006), who observes that the infinitive following *wanna* can be coordinated with a *to* infinitive, a property not shared with bare infinitives. (It is more felicitous without the *to*, coordinating just the VPs, which is why Hudson marks the example as “?”.)

- (29) a. I wanna go to sleep and (?to) not wake up until I feel better.
 b. He let me go to sleep and (*to) not wake up until I feel better.⁶

The acceptability of the coordination with a *to* infinitive is unexpected if *wanna* is a verb that takes a bare infinitive complement. However, Hudson’s alternative does not fare much better. Under Hudson’s account, *wanna* is the morphophonological realization of the sequence *want to*, i.e. a morphological unit which realizes a sequence of two words. For most speakers, this realizational rule is limited to the variety of *want* that takes an infinitive and no object; this accounts for the Lakoff/Horn Generalization, but in a totally ad hoc manner. Hudson justifies the use of a lexical realizational rule on the grounds that phonological rules should express generalizations, not a phenomenon that is as restricted as *wanna* contraction; assuming Pullum’s entire list (as he does) it is still a small number of verbs. The place for such phenomena is in the lexicon. However, as we have seen, the facts of *wanna* contraction are the consequence of the interaction of the attachment properties of *to*, the lexical phonological properties of *want* (which are independent of *wanna* contraction), and the operation of cliticization. Placing the entire phenomenon in the lexicon leads to loss of generalization, rendering other syntactic consequences of the attachment properties of *to* a distinct phenomenon.

The evidence therefore points to a phonosyntactic contraction (cliticization) analysis. There is no support for a lexical analysis.

5. Consequences

We conclude that, when all the facts are considered, *wanna* contraction can best be described in a framework in which the canonical position of the locally licensed function of a fronted element is occupied by an empty category—a phrasal node which dominates no terminal nodes. This empty element blocks the leftward cliticization of *to* onto *want* for most speakers by breaking the adjacency between them.

5.1. Other Empty Categories

It was in the early Principles and Parameters rhetoric that the idea that *wanna* contraction provides evidence for empty categories was first raised. However, already in some of the earliest literature on the matter, it was observed that not all postulated empty categories block *wanna* contraction. The structure assumed in Principles and Parameters for (1), for example, is (30).

- (30) [who do you want [*e* to see *e*]]

The second empty category in (30) is the one we have been discussing. However, there is another empty element in this hypothesized structure, occupying the canonical position of the controlled subject of *see*.⁷ This empty element, no less than the one in (2), intervenes between *want* and *to*. If the inability of most speakers to contract in (2) is evidence for the empty category, their ability to contract in (1)/(30) is

⁶Hudson has *not* and *to* in the other order, but this appears to be what he intends.

⁷This additional empty category is usually referred to as PRO.

evidence against the presence of an empty category in that position. Similarly, as originally noted by Postal and Pullum (1978), the other verbs listed in (20) are Raising verbs: in Raising constructions yet another empty category is hypothesized in Principles and Parameters, and this one also does not block contraction.

We must also consider auxiliary contraction, in particular the contraction of *is*. As shown in the following examples from Carden (1983: 45), auxiliaries can contract over the canonical position of the subject in long-distance dependency constructions.

- (31) a. Who do you think's gonna win?
b. Jack is the man that I bet's gonna win.

On the assumption that long-distance dependency constructions always have an empty category in the canonical position of the locally licensed function, these sentences should have the following representations.

- (32) a. Who do you think [*e* is gonna win]?
b. Jack is the man that I bet [*e* is gonna win].

However, the presence of empty categories such as these ought to block contraction for the same speakers for whom *wanna* contraction is blocked by empty categories. The possibility of contraction thus indicates the absence of an empty category preceding *is* in these sentences:

- (33) a. Who do you think [is gonna win]?
b. Jack is the man that I bet [is gonna win].

The difference between these cases and the earlier ones involving *wanna* is that in the previous cases the locally licensed function of the fronted element is arguably object (assuming a Raising-to-Object analysis for *want*), while in the auxiliary contraction cases the fronted element bears no locally licensed function other than subject. The conclusion that we draw from these facts is that when the locally licensed function of a fronted element is subject, there is no empty category in its canonical position. To summarize, the motivated structures are the following:

- (34) a. Who do you want [_{CP} to see *e*]?
b. Who do you want *e* [_{CP} to see Pnina]?
c. Who do you think [_{IP} is going to win]?

The conclusion, then, is that long-distance dependency constructions involve the use of empty constituent structure in the canonical structural position of the locally licensed function, unless this function is “subject”. This conclusion converges with similar proposals made in other studies, such as Gazdar (1981) and Falk (2006). However, it clashes sharply with the view in the P&P tradition.

To summarize: Pronominal empty categories and empty categories for “NP movement” constructions do not exist, and neither do empty categories in canonical subject position even for long-distance dependency constructions.

5.2. Constraint-Based Syntax

Constraint-based theories of syntax are not inconsistent with the existence of empty categories, as can be seen by examining such studies as Gazdar (1981), Kaplan and Bresnan (1982), Zaenen (1983), Gazdar, Klein, Pullum, and Sag (1984), Pollard and Sag (1994), Bresnan (1995, 2001), Falk (2001), and Culicover and Jackendoff (2005). However, there is a natural suspicion of empty categories among people working in such frameworks. As Dalrymple (2001: 415) puts it,

Further work will reveal ... whether incontrovertible evidence exists for traces, gaps, or empty phrase structure categories. In the absence of such evidence, a simpler and more parsimonious theory of long-distance dependencies results if traces [i.e. empty

categories] are not allowed.

This suspicion stems in part from a perception that the P&P tradition shows that once empty categories are recognized, there is a tendency for the inventory of empty categories to grow in an unconstrained fashion.⁸ As with all elements that are not overt in the actual utterance, these empty categories are prone to such unconstrained proliferation, as well as posing issues of parsing and the like.⁹ For this reason, many studies in constraint-based approaches have championed non-empty-category approaches to long-distance dependency constructions (see, for example, Kaplan and Zaenen 1989, Sag and Fodor 1994, Ginzburg and Sag 2000, Bouma, Malouf, and Sag 2001, Dalrymple 2001, and Dalrymple, Kaplan, and King 2001).

The empirical evidence confirms the suspicion of empty categories up to a point. It is striking that the evidence points to the nonexistence of empty categories in so many of the contexts in which they have been hypothesized. Our proposal is that while empty categories are allowed, they are a dispreferred last resort. As noted by Bresnan (2001), the last-resort status of empty categories is a consequence of the principle of Economy of Expression, under which syntactic nodes are present only when needed to license grammatical f-structures or for their semantic content. Empty categories have no semantic content, so their only possible role is in licensing grammatical f-structures. If an alternative method is available for licensing the same f-structure, the use of an empty category will be blocked by Economy of Expression. The last-resort nature of empty categories explains the empirical evidence of their rather limited distribution.

Most of the constructions for which empty categories have been proposed are lexical constructions, for which empty categories are unnecessary. The only confirmed cases of empty categories have been in long-distance dependency constructions. Long distance dependency constructions, unlike the other constructions for which empty categories have been proposed, are not argument-realization constructions and thus not lexical. Instead, they are multifunctionality constructions in which an argument is paired with a **non**-argument function. In sentences such as the following, for example, the sole argument function of the NP *Pnina* is as subject or oblique object of the verb *spoke*, just as it would be in an ordinary non-LDD sentence. The reason that the NP appears at the beginning of the sentence is because it has an additional discourse-related function. This additional function is not a result of the argument status of *Pnina*; instead, it is related to the discourse in which the sentence is embedded.

- (35) a. Pnina, I said that you think spoke to us
b. Pnina, I said that you think we spoke to.

The fact that this is not an argumenthood-related phenomenon is reinforced by the fact that adjuncts can also be assigned this additional discourse function.

- (36) a. When do you think the plane will arrive?
b. How did Yoni say he would fix the sink?

Assimilating LDD constructions to lexically-based constructions would force us to represent adjuncts as part of the lexical selectional properties of verbs, as is done by Bouma, Malouf, and Sag (2001) in an HPSG analysis.

Given that LDD constructions are not lexical and that Universal Grammar allows them, there must be some non-lexical mechanism for licensing them. One possibility would be for the clause in which the “fronted” element is located to include an outside-in equation stating that this discourse-function-bearing element bears some grammatical function in a clause arbitrarily far down: outside-in

⁸As a student of mine once blurted out in class: “If you assume traces, you may as well be doing GB!”

⁹We will not address the question of how speakers recognize empty categories if they cannot be heard. In our view it is tautological that if it can be shown that empty categories exist, the parsing mechanism for hypothesizing them must also exist. Similarly, we do not consider it a problem that in languages with freer constituent order than English the empty category could occupy more than one linear position in the c-structure. There may be principles that restrict its linear position (such as a principle that places lighter elements before heavier ones); if not, what results is a case of innocuous structural ambiguity.

functional uncertainty:

- (37) a. $(\uparrow \text{ FOCUS}) = (\uparrow \text{ COMP* SUBJ})$
b. $(\uparrow \text{ FOCUS}) = (\uparrow \text{ COMP* OBJ})$

However, while (37a) is unproblematic, (37b) runs afoul of the theory of subjecthood proposed by Falk (2006). According to Falk's theory, functional equations are not free to reference any element in a lower clause: they are limited to referencing the subject.¹⁰ This limitation, which is justified on conceptual grounds, is responsible for the restriction of functional controllees to subjects. Just as the lower element in a control equation is limited to SUBJ, the lower element in an LDD-licensing functional uncertainty equation is limited to SUBJ.

If this line of argumentation is correct, Universal Grammar faces a problem: how to license LDD constructions in which the locally-licensed grammatical function is not SUBJ. The only possibility left is an equation associated with an element of the lower clause, an inside-out equation, which would have to be associated with the node representing the locally-licensed function. Since there is no lexical content to fill such a node, the result is an empty category: the only means available to Universal Grammar to license the construction.

Despite our endorsement above of the view expressed by Bresnan (2001) that the last-resort status of empty categories is a consequence of LFG's Economy of Expression principle, our approach differs crucially from Bresnan's. For Bresnan, inside-out functional uncertainty is necessitated in languages like English to identify the locally licensed function of the fronted element because of the lack of morphological devices such as Case; languages in which such morphological marking exists do not use empty categories. Under the present approach, all languages need empty categories to license non-subject LDD constructions because of the restriction of the lower end of an outside-in designator to SUBJ, regardless of the morphological devices available in the language.

The last-resort view of empty categories, combined with a distinction between lexical and constructional phenomena and Falk's theory of subjecthood, results in a situation in which empty categories are present only in non-subject LDD constructions. This agrees with the results of our empirical investigation into the distribution of empty categories.

References

- Andrews, Avery (1978) "Remarks on *to* Adjunction." *Linguistic Inquiry* 9: 261–8.
- Aoun, Joseph, and David Lightfoot (1984) "Government and Contraction." *Linguistic Inquiry* 15: 465–473.
- Barss, Andrew (1995) "Extraction and Contraction." *Linguistic Inquiry* 26: 681–694.
- Bouchard, Denis (1984) *On the Content of Empty Categories*. Dordrecht: Foris.
- Bouma, Gosse, Robert Malouf, and Ivan A. Sag (2001) "Satisfying Constraints on Extraction and Adjunction." *Natural Language and Linguistic Theory* 19: 1–65.
- Brame, Michael (1981) "Trace Theory with Filters vs. Lexically Based Syntax Without." *Linguistic Inquiry* 12: 275–293.
- Bresnan, Joan (1982a) "The Passive in Lexical Theory." in Joan Bresnan, ed., *The Mental Representation of Grammatical Relations*. Cambridge, Mass.: MIT Press. 3–86.
- Bresnan, Joan (1982b) "Control and Complementation." in Joan Bresnan, ed., *The Mental Representation of Grammatical Relations*. Cambridge, Mass.: MIT Press. 282–390.
- Bresnan, Joan (1995) "Linear Order, Syntactic Rank, and Empty Categories: On Weak Crossover." in Mary Dalrymple, Ronald M. Kaplan, John T. Maxwell III, and Annie Zaenen, ed., *Formal Issues in Lexical-Functional Grammar*. Stanford, Calif.: CSLI Publications. 241–274.
- Bresnan, Joan (2001) *Lexical-Functional Syntax*. Oxford: Blackwell.
- Carden, Guy (1983) "The Debate About *wanna*: Evidence from Other Contraction Rules." in John F. Richardson, Mitchell Marks, and Amy Chukerman, ed., *Papers from the Parasession on the Interplay of Phonology, Morphology, and Syntax*. Chicago: Chicago Linguistic Society. 38–49.

¹⁰More precisely, they must reference the pivot, roughly in the sense of studies of ergativity such as Dixon (1994). For the purposes of the present study, this distinction is unnecessary, so the text will refer to the more familiar concept of subject.

- Chomsky, Noam (1976) "Conditions on Rules of Grammar." *Linguistic Analysis* 2: 303–351.
- Culicover, Peter W., and Ray Jackendoff (2005) *Simpler Syntax*. Oxford: Oxford University Press.
- Dalrymple, Mary (2001) *Lexical-Functional Grammar* (Syntax and Semantics, Vol. 34). New York: Academic Press.
- Dalrymple, Mary, Ron Kaplan, and Tracy Holloway King (2001) "Weak Crossover and the Absence of Traces." in Miriam Butt and Tracy Holloway King, ed., *Proceedings of the LFG01 Conference, University of Hong Kong*. On-line: CSLI Publications. 66–82.
<http://csli-publications.stanford.edu/LFG/6/lfg01.html>
- Dixon, R.M.W. (1994) *Ergativity*. Cambridge: Cambridge University Press.
- Falk, Yehuda N. (2001a) *Lexical-Functional Grammar: An Introduction to Parallel Constraint-Based Syntax*. Stanford, Calif.: CSLI Publications.
- Falk, Yehuda N. (2006) *Subjects and Universal Grammar: An Explanatory Theory*. Cambridge: Cambridge University Press.
- Gazdar, Gerald (1981) "Unbounded Dependencies and Coordinate Structure." *Linguistic Inquiry* 12: 155–184.
- Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum, and Ivan Sag (1984) *Generalized Phrase Structure Grammar*. Oxford: Basil Blackwell.
- Ginzburg, Jonathan, and Ivan A. Sag (2000) *Interrogative Investigations: The Form, Meaning, and Use of English Interrogatives*. Stanford, Calif.: CSLI Publications.
- Halpern, Aaron (1995) *On the Placement and Morphology of Clitics*. Stanford, Calif.: CSLI Publications.
- Hudson, Richard (2006) "Wanna Revisited." *Language* 82: 604–627.
- Jacobson, Pauline (1982) "Evidence for Gap." in Pauline Jacobson and Geoffrey K. Pullum, ed., *The Nature of Syntactic Representation*. Dordrecht: Reidel. 187–228.
- Jaeggli, Osvaldo A. (1980) "Remarks on *to* Contraction." *Linguistic Inquiry* 11: 239–245.
- Kaplan, Ronald M., and Joan Bresnan (1982) "Lexical-Functional Grammar: A Formal System for Grammatical Representation." in Joan Bresnan, ed., *The Mental Representation of Grammatical Relations*. Cambridge, Mass.: MIT Press. 173–281.
- Kaplan, Ronald M., and Annie Zaenen (1989) "Long-Distance Dependencies, Constituent Structure, and Functional Uncertainty." in Mark R. Baltin and Anthony S. Kroch, ed., *Alternative Conceptions of Phrase Structure*. Chicago: University of Chicago Press. 17–42.
- Kroeger, Paul (1993) *Phrase Structure and Grammatical Relations in Tagalog*. Stanford, Calif: CSLI Publications.
- Lakoff, George (1970) "Global Rules." *Language* 46: 627–639.
- Lightfoot, David (1976) "Trace Theory and Twice-Moved NPs." *Linguistic Inquiry* 7: 559–582.
- Pollard, Carl, and Ivan A. Sag (1994) *Head-Driven Phrase Structure Grammar*. Stanford, Calif.: CSLI Publications.
- Postal, Paul M. (1974) *On Raising: One Rule of English Grammar and Its Theoretical Implications*. Cambridge, Mass.: MIT Press.
- Postal, Paul M., and Geoffrey K. Pullum (1978) "Traces and the Description of English Complementizer Contraction." *Linguistic Inquiry* 9: 1–29.
- Pullum, Geoffrey K. (1997) "The Morpholexical Nature of English *to* Contraction." *Language* 73: 79–102.
- Pullum, Geoffrey K., and Paul M. Postal (1979) "On an Inadequate Defense of "Trace Theory"." *Linguistic Inquiry* 10: 689–706.
- Radford, Andrew (1997) *Syntactic Theory and the Structure of English: A Minimalist Approach*. Cambridge: Cambridge University Press.
- Sag, Ivan, and Janet D. Fodor (1994) "Extraction Without Traces." *WCCFL* 13: 365–384.
- Uriagereka, Juan (1998) *Rhyme and Reason: An Introduction to Minimalist Syntax*. Cambridge, Mass.: MIT Press.
- Zaenen, Annie (1983) "On Syntactic Binding." *Linguistic Inquiry* 14: 469–504.
- Zwicky, Arnold (1982) "Stranded *to* and Phonological Phrasing in English." *Linguistics* 20: 3–57.
- Zwicky, Arnold M., and Geoffrey K. Pullum (1983) "Cliticization vs. Inflection: English *n't*." *Language* 59: 502–513.