

Stranded *to* and phonological phrasing in English*

ARNOLD M. ZWICKY

Abstract

*The English infinitive marker to usually forms a phonological phrase with its following VP, but sometimes it is stranded from its VP and phrases instead with preceding material, as in I'd hate to. However, in some circumstances this reattachment is blocked: *To is what I want. Speaker acceptability judgments and the distribution of stranded to in conversational texts together suggest two general conditions on reattachment (one blocking the leftward movement of to out of its S, the other blocking the attachment of to to forms of be) and several conditions subject to complex variation among speakers (a requirement that to attach only to a predicate construction, plus refinements of this requiring close construction between to and what precedes it and requiring that the predicate be interpreted personally); these latter variable constraints do not apply when the preceding word is a monosyllabic nonlexical item like not, how, or him. It is speculated that this array of conditioning arises in acquisition, during which some projection must be made from the three exemplary constructions with any frequency (want to, have to, used to).*

1. Phrases

There are at least four places in linguistics where something like the notion of a PHRASE is called for. There is, first, the phrase of syntax, a group of words that functions as a distributional unit, that is, a CONSTITUENT.¹ Then there is the corresponding unit in semantics, a collection of material upon which an interpretive principle works. Obviously, these semantic units will ordinarily also be syntactic phrases — but not invariably, or otherwise the task of compositional semantics would be much easier than it appears to be. Next there is a lexical unit: a multi-word lexical item like *shoot the breeze* is not just any combination of words, but has internal structure and a kind

of integrity as a whole. Again obviously, lexical items will ordinarily also be syntactic phrases and function as units for the purposes of semantic interpretation, but not invariably, for there are discontinuous lexical items, asyntactic idioms, and idioms (like *make light of*) composed of pieces of adjoining constituents.

Finally, there are multi-word constructs that function as units phonologically. These PHONOLOGICAL PHRASES serve as the domains within which external sandhi rules operate, and their boundaries are locations for the operation of phonological rules conditioned by hiatus or pause; they also serve as the domains for the assignment of stress and intonation patterns, and as units of timing; and their boundaries mark locations where parenthetical interruptions can occur. It seems unlikely that a single type of construct will be able to serve all functions at once, but in at least some simple cases we should be able to expect that the various criteria demarcating phonological phrases will coincide.

In any event, phonological phrases, however delimited, are not invariably syntactic constituents. The divergence between the syntactic and phonological organizations of the same material has long been recognized as a problem in analysis and a challenge to theorizing, finding recognition in works as diverse as Kahane and Beym (1948), Pulgram (1970), Bing (1979), Cooper and Paccia-Cooper (1980), and the writing of the 'metrical phonologists', in particular Selkirk (1981). In this paper I will add to this literature by exploring in great detail the principles governing the way one single morpheme in one language — namely, the English infinitive marker *to* — attaches to neighbouring material to form phonological phrases. As it turns out, this initially rather barren-seeming topic unfolds into matters of considerable subtlety and complexity, from which, at least tentatively, lessons of some generality can be drawn.

One fundamental assumption informs most of this work on phonological phrasing: unless otherwise specified, phonological phrases ARE syntactic constituents. That is, the principles governing phonological phrasing have usually been viewed as reshapings of syntactic constituent structure to meet the demands of prosodic and segmental phonology. I accept this assumption in what follows — on *a priori* grounds, because this assumption has the potential of being considerably more restrictive than the hypothesis that syntax and phonology independently organize the same material in different ways; and on *a posteriori* grounds, because phonological phrasing agrees so often with syntactic constituent divisions (while being less articulated, 'flatter' in the current terminology, than syntactic structure). What this assumption means formally is that there are principles reorganizing and reducing the structures of syntax, to yield the divisions appropriate for phonology — the READJUSTMENT RULES of

Chomsky and Halle (1968: 9–11) (especially as elaborated on in Langendoen, 1975).

2. Leaners

The particular subtopic, within the general area of phonological phrasing, that interests me here concerns the analysis of words that are prosodically subordinate to neighbouring material. LEANERS, as I shall call these words, form a rhythmic unit with the neighbouring material, are normally unstressed with respect to this material, and do not bear the intonational peak of the unit. English articles, coordinating conjunctions, complementizers, relative markers, and subject and object pronouns are all leaners in this sense.

2.1. Clitics

A special case which must be distinguished at the outset is that of the CLITIC, a morpheme that attaches to a neighbouring word to form a word-like unit with it. The phonological phrase/phonological word (Pulgram's *cursus/nexus*) distinction is not always easy to make, although there are reasonably clear examples of each in familiar languages: the English prepositions form phonological phrases with their object NPs, but these combinations do not exhibit word-internal phonology; the English possessive morpheme 's, in contrast, always combines phonologically with the final word of the NP with which it is in construction, as in *the governor of New York's cabinet*. It can also happen that a morpheme which is ordinarily a leaner acts as a clitic in certain circumstances, as when the *to* that can phrase with a preceding verb (*I'd hate to, you'll need to*) contracts with a handful of these verbs to form the phonological words [wʌnə, hæftə, jüstə] etc. that have been the object of so much attention in the recent syntactic literature (more on this in section 5.2 below). However, it is the leaner *to* that I want to focus on here, not the much more restricted clitic *to*. To this end, I will examine instances of *to* with an unreduced vowel, [tu], thus avoiding cases of [tə] which are potentially clitics. And since my interest is in *to* as a leaner, I will also avoid instances of [tu] serving as phrasal stress peaks, that is, cases of *to* bearing contrastive stress (as in *I'd hate TO go*); in any event, it is known (Zwicky and Levin, 1980) that contrastively stressed *to* is subject to special restrictions on its occurrence.

2.2. *General observations*

I now set the stage for this discussion of *to* with some general observations about leaners. In my examples phonological phrasing is indicated by square brackets.

First, recall that phonological phrasing is ordinarily supplied by syntactic constituent structure, as in (1), but that on occasion special readjustments are required, as in (2).

- (1) a. [the líon]
 b. [wént] [to África]
 c. [not Súsan]
 d. [big enough]
- (2) a. The rat [that Tára] found was fierce.
 b. You know [that Sám] will soon be here.
 c. It was Africa my brother [wént to].
 d. I would [háte to].

In either case, attachment can be to the right, as in (1a–c) and (2a–b), or to the left, as in (1d) and (2c–d).

Some readjustments are optional, corresponding to alternative phrasings and patterns of sentence stress. Thus the sentence *I saw him* can have its pronoun object independent and stressed as in (3a), or attached to the verb and stressless as in (3b).

- (3) a. [I sáw] [hím].
 b. [I sáw him].

(The traditional wisdom on the relationship between stress and phrasing has it that differences in sentence stress lead to different phrasings, that is, to different readjustments. But there is much to recommend the reverse assumption, which was first explored in a generative framework by Selkirk (1980), where attachment leads to destressing. For my immediate purposes, it is sufficient to note that attachment and relative stresslessness go together.)

Sometimes if a leaner is STRANDED (separated from its natural coconstituent) on one side, it reattaches on the other. This happens when a preposition, as in (1b), loses its object NP — compare (2c) — and when the infinitive marker *to*, as in (4a), loses its VP — compare (4b).

- (4) a. [We're nót] [to léave].
 b. [We're nót to].

Indeed, some leaners **MUST** have material to attach to. Here, the stranded preposition *to* contrasts with the stranded infinitive marker *to*, because only the former can stand alone:

- (5) Q. Are you travelling to, or from, Africa?
A. To.
- (6) Q. Do you want to leave, or not to?
A. {Not to.}
{*To. }

Stranded articles and coordinating conjunctions will not reattach and are consequently as unacceptable as the stranded *to* in (6):

- (7) Q. Do you just want a sandwich, or do you want some particular one?
A. *I just [wánt a].
- (8) *Who did you see [Pául and]?

(Note that various descriptive devices have been suggested to account for the unacceptable examples. Whether these are to be excluded by the proper formulation of phrase structure rules, by conditions on particular transformations, by such general constraints on transformations as Ross' Coordinate Structure Constraint, or by special conditions referring to the offending surface configurations is not my concern at this point. The analysis developed below for stranded *to* will not necessarily carry over to other leaners; it might, for instance, be sufficient for a description of English to rule out examples like (8) via the CSC or its effective equivalent, without mentioning the fact that *and* is an obligatory leaner.)

Most important for what is to follow, even those leaners that reattach, like the infinitive marker *to*, won't attach to just anything. Stranded *to* attaches to a preceding verb in (2d), to *not* in (4b) and (6), to a verb-object combination in (9a), and to interrogative *how* in (9b), but refuses to phrase with the conjunctions and complementizers in (10); in what follows I will add further types of constructions to those illustrated here.

- (9) a. I don't know if Paul wants to buy the present, but I think we can [persúade him to].
b. I might whittle a polar bear out of Ivory soap, but I don't know [hów to].
- (10) a. *You shouldn't play with rifles, [because to] is dangerous.
b. *You can try to plead with him, but I doubt [thát to] will help.
c. *She'd like to surprise him, but I don't know
 { [whéther to] }
 { [if to] } is possible.
d. *If you want to finish today, [thén to] you're going to have to work fast.

The problem now is to state the conditions that rule out the examples in (10) while permitting the earlier acceptable examples, to do this in a

general way, and to connect the analysis with what is known about phonological phrasing in other cases.

2.3. *Research strategies*

There are a number of ways of obtaining relevant data about phonological phrasing in a language. There is the strategy of the psycholinguist, who asks people to read carefully selected material and then uses instruments to measure timing, pitch, loudness, and segmental characteristics of the subjects' productions. There is the strategy of the sociolinguist, who tapes conversations, narratives, descriptions, and the like, later extracting the required information by ear or by instrument and analysing it statistically. There is the strategy of the text linguist, who examines printed representations of conversations, interviews, speeches, narratives, or whatever is available, looking for cases where the phrasing can be determined from what is on the page. And there is the strategy of the generative grammarian, who asks people for their judgments on the acceptability of particular renditions of sentences.

Each strategy has its strengths and weaknesses (the psycholinguist's strategy, for instance, provides an enormous amount of information about a very small data world, and it is always troubled by the incalculable unnaturalness of experimental settings); ideally, one would want to pursue a given topic on all fronts at once. My own preference has been to select problems that can be pursued fruitfully via the strategies of the text linguist and the generative grammarian. The cliticization of English auxiliary verbs to preceding material is one such case, and the phrasing of infinitival *to* is another. What makes *to* so amenable to non-instrumental investigation is the fact that stranded *to* has (in general) nowhere to attach BUT to the immediately preceding material, so that if we find an instance of stranded *to* in the text, or if speakers accept an instance of stranded *to*, we can assume that the *to* does in fact phrase in these cases with the immediately preceding material.

3. **Digressions on method**

I have now presented sets of acceptable and unacceptable examples, which, though chosen by me with some thought, have no guarantee of being representative instances of the full set of data to be accounted for. It would be natural at this point to ask for a list of all relevant types of examples. Unfortunately, assembling such a list involves (a) providing a

good taxonomy of the types of English constructions involving marked infinitives, (b) enumerating the potentially relevant properties of the linguistic contexts in which each construction-type occurs (a quick glance at the data provided so far suggests that only the context immediately preceding an instance of *to* is relevant, but this context must still be categorized in an appropriate way), and (c) listing the ways in which *to* can be stranded — and of these tasks, only (c) is easy. Task (a) is challenging, and I will argue that task (b) is impossible. Let me take them up in this order.

First, the means by which *to* can be stranded. I believe that there are only two ways: through the separation of *to* and its VP by parenthetical material, and by the rule of Verb Phrase Deletion (VPD).² The first way is illustrated in (11) with collected examples; the second has already been illustrated with invented examples, which are supplemented in (12) by some collected ones.

- (11) a. I made a decision whether to — and he made a decision, too — have him in the group.

(John Lennon, *Rolling Stone Interviews (RSI)*, p. 150)

- b. ... we gotta — I have to personally — get down to so-called reality.

(John Lennon, *RSI*, p. 130)

- c. ... like, if I want to or could use a man ...

(Hareven/Langenbach, *Amoskeag*, p. 234)

- d. It was evident that his mood did not quite match mine, and that I should have to — as women nearly always must — damp down my own exuberant happiness ...

(Pym, *A Glass of Blessings*, p. 187)

- (12) a. ... but I didn't go around agitating — I had no reason to.

(*Amoskeag*, p. 312)

- b. I'd love to eventually, I feel I could write lyrics someday.

(Elton John, *RSI*, p. 296)

- c. I can put my dukes up now if I have to in life.

(Joni Mitchell, *RSI*, p. 389)

- d. "How come you've never called me in all these years?" Tasha asked.

"I wanted to, but I was afraid to ..."

(Davidson, *Loose Change*, p. 293)

- e. I don't really like to talk about my work with my friends. They don't really seem to, either.

(Terkel, *Working*, p. 710)

It is important that the examples in (11) all involve true PARENTHETI-

CALS (in the sense of Rotenberg, 1978, Part I), rather than hesitations, expletives, or irrelevant interruptions like coughs or sneezes; in particular, the interruptions in (11) are all syntactically structured and semantically related to the surrounding sentences.³ Hesitations, including pauses filled with *uh* and the like, are well known to occur at points of high unpredictability in an utterance — between an article and a following adjective or noun, or indeed between *to* and a following verb — rather than at phrase boundaries, and it follows that the distribution of hesitations will tell us little about phonological phrasing. Insofar as true parentheticals are not distributed in this way, they can be used (cautiously) as indicators of phrase bounding.

Next, the constructions in which *to* occurs. It isn't easy to list these. (The *OED* has four pages of exposition and examples under TO B. 'To before an infinitive', with 22 notional subdivisions, some with up to five further subdivisions.) Marked infinitive constructions function as subjects, direct objects, objects of prepositions, adjectival modifiers (relative clauses), adverbial modifiers, degree modifiers, not to mention several hard-to-classify types. Table 1 provides a rough outline of these cases, with no implied claim to completeness.

Table 1. *Marked infinitive constructions*

-
- I. Subject complements
 1. (*for/of*) *to*
 - a. In subject position: (*For you*) *not to finish your food is irritating.*
 - b. Extraposed: *It was nice of you to come. It's hard (for us) to understand this.*
 2. *Wh to*
 - a. In subject position: *What to do puzzles me.*
 - b. Extraposed: *It is obvious what to do.*
 - II. Predicate complements
 1. Equi
 - a. Subject control: *I try to be nice. We promised Matilda to behave ourselves.*
 - b. Modals (*ought/dare/need to*) and quasimodals (both with subject control): *He has to exercise daily. I used to read Sanskrit. You are to leave quietly.* (Similarly, *be (un)able to, be going to, be about to, be supposed to, be obliged to.*)
 - c. Object control: *We asked Stan to leave.*
 2. 'Mental Equi': *Surely you knqw to jump if you see a snake.*
 3. Indirect questions
 - a. Subject control: *I realized where to go.*
 - b. Object control: *I told Georgia who to see.*
 4. SS raising: *I happen to like kumquats.*
 5. SO raising: *I expect him to explode soon.*
 6. OS raising: *Noam is hard to convince.*
 - III. Nominalizations of constructions in II: *his intention to leave, her ability to see into dark places.*

Table 1. *Continued*

IV.	Infinitival relatives
	1. Subject readings: <i>I was the first (person) to enter. He's not the kind of man to do that.</i>
	2. Object readings: <i>Vienna is the place to visit. This is the gift to give.</i>
	3. Adverbial readings: <i>Ten is the time (for us) to go. I need someplace to live.</i>
V.	Degree modifiers
	1. <i>Too</i>
	a. Subject readings: <i>I'm too tall to get in.</i>
	b. Object readings: <i>It's too heavy (for me) to lift.</i>
	2. <i>Enough</i>
	a. Subject readings: <i>It's large enough to be troublesome.</i>
	b. Object readings: <i>It's large enough (for Tamara) to have trouble with.</i>
VI.	Adverbial subordinate clauses
	1. Condition (expressed subject): <i>(In order) for you to understand what I am saying, a miracle is apparently necessary. If you were to sing, I'd join you.</i>
	2. Reason (subject control): <i>He stood up (in order) to see better. The mayor, (so as) to see better, stood up.</i>
	3. Purpose
	a. Subject control: <i>I've come to take the furniture. He rose to go.</i>
	b. Double control: <i>I bought this to amuse myself with. I bought you this to amuse yourself with.</i>
	4. Cause (subject control): <i>I was distressed to see the state of the manuscript.</i>
	5. Result (object control): <i>He fixed it to explode in three hours.</i>
	6. Manner (subject control): <i>She moved as if to take a gun from her purse.</i>
	7. Speech act (speaker control): <i>To be frank with you, this soup stinks. Your ideas, to draw a familiar analogy, are like gold.</i>
VII.	Objects of prepositions
	1. Expressed subject: <i>I can't think of a solution, other than for Kevin to commit suicide.</i>
	2. Subject control: <i>He did nothing except to shake his head.</i>
VIII.	Absolutes (subject control): <i>The yetis jumped into the crevasse, never to be seen again. Herbert ran out, only to return a moment later.</i>
IX.	Exclamations: <i>Oh, to be myself again! Never to see himself in the newspaper!</i>
X.	Assorted unclassifiables: <i>Books are for people to read. It has something to do with George. Nothing was to be heard. It's mine to keep. John is to blame. This idea is to be approved of.</i>
XI.	Idioms containing infinitives: <i>(And) now to finish dinner; to tell the truth; to be sure; etc.</i>

The task of listing the relevant linguistic contexts in which *to* occurs is intimately connected with the task of constructing something like Table 1. Is a subject complement in subject position a different construction from

an extraposed subject complement, or are these a single construction occurring in two different contexts? Similarly for verbs like *intend (to)* and nouns like *intention (to)*; for reason and speech act adverbials located at different places within a sentence; and for subjectless indirect questions with subject control and with object control. These I have listed separately in Table 1, but there are an indefinite number of constructions I haven't listed, for instance, passives of II.3.b cases (13a), wh-clefted versions of these passives (13b), inversions of these wh-clefts (13c), or these inversions appearing as objects (13d); nor have I listed the predicates of II.6 turned into prenominal modifiers (*several hard-to-classify types*); nor have I differentiated subtypes of II.1.a, although nothing guarantees that *try to*, *want to*, *expect to*, *promise someone to*, *get to*, and *bring oneself to will* show identical behaviour with respect to the stranding of *to*.

- (13) a. Georgia was told who to see.
 b. What Georgia was told was where to go.
 c. Where to go was what Georgia was told.
 d. I know that where to go was what Georgia was told.

The possibility that different lexical items might behave differently and the fact that constructions can be compounded together mean that a COMPLETE list would be very unwieldy. The recursiveness of language ensures that a complete list is impossible. The best we can hope for is a large and varied sampling.

3.1. *Variation*

Armed with Table 1, or an expansion of it, we now attempt to strand *to* by inserting parenthetical material between it and its VP, and by supplying a context in which the VP can be deleted. Several problems arise.

To begin with, most people find it very difficult to judge the acceptability of sentences with parenthetical interruptions in them, so that VPD is the only usable device for stranding *to*. The fact that we must rely so heavily on one rule leads to several complications in interpreting judgments; I discuss these in section 4.1 below.

Then there are disagreements about the acceptability of particular examples. I believe that the judgments in (1)–(10) are shared by nearly everyone, but when we go beyond these examples speakers differ in their judgments. Moreover, these differences manifest themselves as DISAGREEMENTS, one set of speakers maintaining they find a certain example utterly beyond reproach, others asserting that they can't imagine anyone uttering such a thing. I have set off shouting matches by unwittingly including

such examples in my data sets at lectures. I will argue in the next section that these differences are real, and that for the most part a speaker's judgment on one type of example is correlated with that speaker's judgments on a number of other types. These claims can be illustrated by a brief reconsideration of the (truly awful) sentences in (10), which involve *to* as a remnant of a subject complement in subject position (10a–c) or as a remnant of a preposed reason adverbial (10d). There are extraposed variants of the former construction, and postposed variants of the latter:

- (14) a. You shouldn't play with rifles, because it's dangerous to.
 b. You can try to plead with him, but I doubt that it will help to.
 c. She'd like to surprise him, but I don't know $\left\{ \begin{array}{l} \text{whether} \\ \text{if} \end{array} \right\}$ it's possible to.
 d. If you want to finish today, then you're going to have to work fast to.

I've given these examples without asterisks — my own judgments — but a substantial number of speakers reject them all, and also reject variants of (9) with slightly different wording:

- (15) a. I don't know if he wants to buy the present, but I think we can persuade [Pául to].
 b. I might whittle a polar bear out of Ivory soap, but I don't know [whether to].

In the face of this sort of variation it is difficult to discern even the gross outlines of the correct analysis; the universally acceptable and unacceptable examples underdetermine the analysis. Consequently, I have had to go beyond the small-scale collection of judgments in my work on stranded *to*, in two ways: I have collected judgments, on sentences exemplifying a few types of infinitive constructions, from a large group of speakers, and I have searched a large body of texts for instances where stranded *to* has been reattached to preceding material, so as to gain some sense of the distribution of this phenomenon in natural conversation. I now briefly describe these two activities.

3.2. Judgment data

The questionnaire included 21 sentences, some indisputably acceptable, some indisputably unacceptable, some involving constructions that I believed to be the source of dialect differentiation. They are listed, in their order of presentation, in Table 2. This order is not random; I did not want

to begin with clearly unacceptable examples, or to have unacceptable examples clustered together, or to have examples that were minimally different (like 4 and 7, or 9 and 13) or similar in structure (like 10 and 12, or 11 and 21) next to one another. The questionnaire was short, because most subjects find the task tiring.

I presented the questionnaire orally (after the first few administrations, the presentations were on a tape). This enabled me to control the prosodic pattern associated with each example. The typed versions, even with *ad hoc* indications of stress maxima, are consistent with a number of different readings, some of which are less natural than others. I tried to give each example the most favourable reading I could.

There were 74 subjects whose responses were usable (I excluded nonnative speakers and one subject who rejected the two very best examples, 8 and 16, and who was incidentally the only subject to skip responses). These divided into two groups, 15 academics (some linguists, some nonlinguist Fellows at the Center for Advanced Study in the Behavioral Sciences) and 59 Stanford University undergraduates.⁴ Responses on particular items did not appear to differ between the two subgroups, although the academics were significantly more positive and

Table 2. *Stranded to questionnaire*

-
1. You can try it, but I doubt that it would help to.
 2. We needed someone to buy the present, so Paul was persuaded to.
 3. Marge thinks I can't lift this weight, but I think it would be easy to.
 4. You want Máry to cook, but I would prefer for Jóhn to.
 5. A. Who do you want to vanish?
B. I don't know; who do you want to?
 6. It's not easy to justify your attitudes, and besides you'd have to talk fast to.
 7. It would be nice for Jóhn to sing, but not for Máry to.
 8. Polly doesn't like to eat meat, but I prefér to.
 9. I didn't want to go, but what he wanted was to.
 10. Maybe you should take some aspirin; I think it would help to.
 11. I'm thinking of moving to Hawaii, but I can't figure out whether to.
 12. A. Do you think you could move this piano?
B. I don't think it's possible to.
 13. I didn't want to go, but what hē wanted was to.
 14. Did he sing?! He didn't do anything excépt to!
 15. Don't ask questions. I believe to will cause trouble.
 16. He didn't clean his room, and I don't think he inténded to.
 17. Herb is easy to convince, but Eve is hard to.
 18. He ought to know his times tables; certainly he's old enough to.
 19. She wanted to ask his age, but she realized it would be rude to.
 20. They had intended to work several more hours, but they were too exháusted to.
 21. I want to calculate the bill, but I don't know hów to.
-

less variable in their responses overall, indicating probably that they understood the nature of the judgment task better. Consequently, I lumped the two groups together for statistical analysis. Preliminary data collection suggested no strong effects associated with regional dialect, age, or sex (social class and ethnic group did not vary enough in this particular group for me to assess possible contributions from these variables), and this information was not collected for the full sample. If there are effects associated with these variables, a much larger sample would be required to detect them.

The subjects were asked to judge the examples on a four-point scale — *ok*, *?ok*, *?no*, *no* — with *ok* assigned to sentences that ‘sound perfectly acceptable’, *no* to those that don’t ‘sound like ANYTHING you would say’, and the intermediate judgments assigned to those they were inclined to accept or reject, respectively, but were unsure of. The intermediate judgments were provided because many people are extremely reluctant to give sharp yes/no responses.

The data set resulting from a procedure like this one is full of irrelevant noise. Some subjects are remarkably positive in general (one gave only two *?no* responses and only one *?ok*, all the rest being *oks*), some are remarkably negative (one gave three *oks*, one *?ok*, and all the rest *no*), some stick to the extreme values on the scale (one with twelve *oks* and nine *nos*), some avoid the unqualified judgments (one with two *?nos* and nineteen *?oks*). The task is understood differently by different subjects (some appeared to accept an example only if it was the BEST of the alternative constructions available for expressing the idea), and notions of correctness are always ready to intrude (one student subject commented marginally on the questionnaire sheet, with somewhat shaky punctuation, ‘prepositions are not permitted at the end of sentences You loser, learn your grammar!’).

Finally, each example contains material that can interfere with judgments. Comments by subjects have led me to select the following as potential contributions to negative judgments in this data set: the *prefer for* construction in 4; the fact that the *to* at the end of 6 could be left out; the fact that when it is first heard the *it* in 10 could be understood as anaphoric to *some aspirin* rather than as an impersonal pronoun; the relative formality of *whether*, versus the informality of *figure out*, in 11; the expression *times tables* in 18, which is unnatural for some; and, in 20, the combination of the jingle *too ... to* with the choice of *exhausted* rather than the more natural *tired*. The statistics for these six examples cannot necessarily be taken at face value.

3.2.1. *Clear judgments.* The first step in analysis is to distinguish those

example sentences for which the subject group gives unquestionably positive or unquestionably negative judgments. The relevant raw data are supplied in Table 3, where the tabulation of the four responses is given for all 21 example sentences. The table also provides the total number of positive (*?ok* or *ok*) responses for each sentence (the total number of negative responses is then 74 minus this figure) and its χ^2 statistic, a measure of how far the positive-negative distribution of responses on any item is from an even 37-37 split.

The sentences in Table 3 have been divided into five groups. Group A contains unquestionably acceptable sentences, those with a preponderance of positive responses and with a χ^2 significant at the 0.001 level (indicating a deviation from an even split which would occur by chance no more than once in a thousand times); the sentences are listed in descending order of response level, with the most positive response (for sentence 6) first. Group B contains unquestionably unacceptable sentences, those with a preponderance of negative responses and with a χ^2

Table 3. *Judgments on example sentences*

Sentence	no	?no	?ok	ok	Total positive	χ^2
A 16	1	0	4	69	73	70.05
8	3	4	10	57	67	48.65
*18	6	1	14	53	67	48.65
21	5	4	18	47	65	42.38
3	5	5	17	47	64	39.41
2	6	5	9	54	63	36.54
1	11	8	21	34	55	17.51
19	11	9	21	33	54	15.62
B 15	68	3	1	2	3	62.49
14	50	19	3	2	5	55.35
9	52	12	7	3	10	39.41
13	40	19	11	4	15	26.16
C *10	21	14	18	21	39	0.22
*11	19	15	21	19	40	0.49
5	13	18	19	24	43	1.95
17	10	20	22	22	44	2.65
D *20	16	10	19	29	48	6.54
*4	13	13	16	32	48	6.54
7	11	14	15	34	49	7.78
12	14	9	19	32	51	10.59
E *6	24	22	13	15	28	4.38

significant at the 0.001 level; these are again listed in descending order of response level, with the most negative response (for sentence 15) first.

My own judgments on the sentences in groups A and B agree with those of the subjects.

Group C contains sentences with response patterns not significantly different from an even split, even at the 0.05 level; these are listed in ascending order of χ^2 , with the most even response (for sentence 10) first. Group D contains sentences with predominantly positive responses, though not reaching significance at the 0.001 level; these are listed in ascending order of χ^2 . Group E contains the one sentence (6) with predominantly negative responses, though not reaching significance at the 0.001 level. In any case, the 'middle' groups C, D and E all contain examples for which the judgments were neither crashingly positive nor crashingly negative, hence for which some amount of 'dialect' variation can be suspected.

The asterisked sentence numbers in Table 3 are those of the six for which there might be negative effects from other material in the sentences. Clearly for example 18 there is no such effect, since it falls in group A. The assignment of examples like 4, 10 and 11 to middle groups is supported by the text data below. Example 20, however, is probably misassigned to group D; in any case the analysis that emerges from the other data taken together predicts the acceptability of 20, so that I must treat it as an anomaly in some way.

3.2.2. The middle group examples. This leaves examples 10, 11, 5, and 17 in Group C, examples 4, 7, and 12 in group D, and example 6 in group E. The question now is how to interpret the distribution of responses among the four categories *no*, *?no*, *?ok* and *ok*. In what follows I will treat these sets of responses as graphs, with 0, 1, 2, and 3 on the horizontal axis corresponding to the four responses from *no* to *ok* (in that order), and with the vertical height at each point determined by the number of responses of the appropriate type. The graph for example 10 thus comprises the four points (0, 21), (1, 14), (2, 18), and (3, 21) and is roughly V-shaped.

But what should we expect these graphs to look like? Suppose that the subjects were simply unsure of their responses; we should then expect these responses to cluster in the two central categories, producing an inverted U graph. Not one of the eight examples I am considering here has a response graph of this shape; so I will discard this possibility without further analysis.

Another possibility — which does not strike me as very plausible — is that the responses are randomly distributed, that the subjects were in

effect mentally tossing two coins to decide on their judgments for certain examples. In this case the response graph would be a horizontal straight line of height 18.5. This possibility can confidently be discarded for the three examples in group D (4, 12, and 17); when their distribution of responses into four categories is compared with the even distribution with 18.5 responses in each category, all three examples have a χ^2 (3 df) significant at the 0.05 level or better (13.50, 17.78, and 15.84, respectively).

What we really want to test for, however, is the extent to which the distribution of responses into four categories for groups C-E deviates from the corresponding distribution for groups A and B. I chose as a measure of deviation/closeness the coefficient of correlation r and had calculated the 190 coefficients of correlation for each pair of the 20 example sentences under consideration (to permit comparison between the negatively oriented groups B and E and the rest, the responses for examples in these two groups had their polarity reversed, so that the distribution 15-12-7-3 for example 9 was treated as 3-7-12-15, and so on.)⁵ I could then compare the r s relating clear cases (groups A and B) to one another with the r s relating the middle group examples to the clear cases. The means (and standard deviations) of these sets of coefficients of correlation are reported in Table 4, the first line of which gives the mean for the 66 correlations of one clear-case response pattern (for a sentence in group A or B) with another.

The following three lines give means for the correlations of group D response patterns with clear-case response patterns; these are high, indicating that the response patterns for sentences 4, 7, and 12 are much like those of clearly acceptable sentences (although there are enough negative responses on 4, 7, and 12 to merit special attention). The remaining comparisons, for the response patterns to sentences in groups C and E versus the clear-case response patterns, all show significant differences.

Table 4. *Closeness of response patterns*

Sentence	Mean r	s
all A and B	0.968	0.030
4	0.981	0.021
7	0.970	0.030
12	0.956	0.026
5	0.863	0.041
6	0.827	0.085
17	0.522	0.085
10	0.500	0.065
11	0.306	0.129

I take this statistical analysis as indicating that the examples in groups C and E have response patterns involving a significant minority judgment. The examples in group D show a weaker effect, not apparent in Table 4. I will treat all these examples as exhibiting a majority judgment at one pole of response and a minority judgment (in the case of the group D examples, the judgment of a small minority) at the opposite pole.

3.2.3. *Clusters of responses.* Next, I wanted to determine to what extent the (relatively) polarized judgments on one example correlated with such judgments on another. For this purpose I compared examples 4, 5, 6, 7, 10, 11, 12 and 17 with one another, comparing now the 74 responses on one example with the 74 on another. The highest coefficients of correlation linked examples 6, 7 and 11 with one another, and 10 with 12 (in all cases, r is significant at the 0.001 level). The next highest (with the same level of significance) link 7 and 11 with 10 and 12. The next highest (r significant at the 0.01 level) link 5 to 6, 7 and 11. There is no significant correlation between the judgments on 5 or 6 and those on 10 and 12. It is clear that 10/12 form a set, as do 7/11, and the two sets are closely tied to one another; 5/6 together are linked to 7/11, but are essentially independent of 10/12.

I am not claiming here that examples 10 and 12, say, are linked in the sense that anyone who accepts one will accept the other. There is a moderate amount of incidental fluctuation in judgments, even on such similar examples as 10 and 12. Nevertheless there is a strong tendency for the judgments on 7, 11, 10 and 12 to hang together and similarly for those on 5, 6, 7 and 11.

(The judgments on example 4, which is similar in construction to example 7, are correlated with judgments on all the other examples in this set except 6 and 12; in what follows I will take example 7 as the representative case, leaving for some future occasion the exploration of possible differences between 4 and 7. Example 17 showed judgments correlating well with the judgments on ALL SIX of the other examples analysed, a fact which is consistent with several states of affairs, including the possibility that 17 is simply acceptable; for the moment I have nothing further to say about this case.)

From the independence of 5/6 and 10/12, I conclude that the relationships are implicational: if you accept 5/6, you are likely to accept 7/11; and if you accept 10/12, you are likely to accept 7/11. You might accept NONE of the middle group examples, in which case I will refer to your variety as 'narrow'. Or you might accept them ALL, in which case I will refer to your variety as 'wide'. Or you might have an intermediate variety (accepting only 7/11, or 7/11 in combination with 5/6 or 10/12).

The interpretation of the results so far is not immediately obvious, except for the examples 10 and 12, both of which involve extraposed subject complements, 10 having the verb *help*, 12 the adjective *possible*. What unites 7 (with *for Máry to*) and 11 (with *whéther to*), or 5 (with *want to* and a *wh* trace intervening between *want* and *to*) and 6 (with a postposed subjectless reason adverbial) does not leap to the eye. As it turns out, the data collected from texts will help in understanding these judgment data.

3.3. Text data

I begin by noting that although VPD in general is not registrally or stylistically limited, applications of VPD leaving *to* as a remainder are so limited. It is possible to read hundreds of thousands of words of narrative or expository prose without coming across a single example like those in (12) above, although other instances of VPD will be plentiful. VPD with *to* is distinctly conversational; its natural habitat is in informal two-person discourse. The *OED* agrees, saying that *to* with ellipsis of the following infinitive is 'rare before 19th c.; now a frequent colloquialism'.

Consequently, my research for examples in published material has relied primarily on edited informal interviews, with some examples from conversation in fictionalized autobiography, and some from dictionary entries, plus a few fortuitous finds.⁶ I collected 157 examples, all but five involving VPD. During the time since the exhaustive searches that resulted in the data of Table 5, I have collected individual instances of items not in that table: from spoken conversation, *and not be too rigid to* followed by a parenthetical interruption; from reading, *be about to*, *wish to*, *deserve to*, *begin to*, *urge me to*, *have a chance to*, *aspire to*, *mean to*, *be meant to*, *be ready to*, *intend for him to*, *expect to*, *choose to*, and *how was he to*, all with VPD.

It was clear from the outset that there were no large effects involving tense, aspect, mood, or person/number, so that the listing in Table 5 lumps together, for instance, *he wants to*, *I wanted to*, *I would have wanted to*, *I didn't want to*, and *you want to* all as *want to*; if there had been instances of *to want to* or *wanting to*, they would have also have been counted as *want to*.

In interpreting these figures, I must stress that most of the frequencies are quite small, so that the nonappearance of a particular combination cannot be taken as indicating its unacceptability (for English speakers in general, or for some substantial subgroup). *Like to* occurs four times, *love to* twice, but *hate to* not at all; aspectual *start to* occurs once, but *begin to*

Table 5. Stranded to in texts

-
- a. Subject-controlled (Equi or SS Raising) predicate constructions (129)
- i. Extremely frequent (89):
 - want to* (43, including one *wanna*)
 - have to* (24)
 - used to* (22)
 - ii. Infrequent but attested (38):
 - seem to, like to, try to* (4 each)
 - need to, got to* (3 each, including one *gotta*)
 - ought to, love to, refuse to; be able to, be going to, be supposed to* (2 each)
 - can't afford to, consent to, intend to, start to, be allowed to, be obliged to, be crazy to, be afraid to* (1 each)
 - iii. *have N to* (2): *have no reason to, have time to*
- b. Impersonal (extraposed subject complement) constructions (3): *it ... be hard to, it ... be stupid to, it ... be easier to*
- c. Indirect questions (3): subject-controlled *how to* as object of *know*; subject-controlled *whether to* as complement of *decision*; *how best to*, understood with generic subject, as complement of *question* (*it's a question of how best to*)
- d. Object-controlled (Equi or SO Raising) predicate constructions (13): *want Pro to* (7), *tell Pro to* (2), *ask Pro to, expect Pro to, get Pro to, allow Pro to* (1 each)
- e. *not to* with various types of control (9), e.g. *I should have told her not to, it was an effort not to, you're too good not to, I try not to.*
-

and *continue to* not at all; there is *consent to* once and *refuse to* twice, but no *agree to*, and one *be allowed to* but no *be permitted to*, and one *be obliged to* but no *be required to*, and four cases of *seem to* but none of *happen to*; and *be likely to* doesn't occur in this sample, nor does *prefer to*, which was one of the two most favoured constructions in the judgment data. I would not want to conclude that any one of these nonoccurring lexical items was actually UNACCEPTABLE with stranded *to*. On the other hand, the absence of whole construction types in the data, or gross disparities in the frequency of construction types, should be given some weight in my analysis.

In this light, several properties of Table 5 are striking. Starting at the bottom of the table, I note that (a) *not* seems to be the only preceding 'adverb' to which *to* attaches with any frequency; (b) *to* attaches fairly freely (13 times) to preceding V + Pro combinations, but very infrequently (no instances in this sample) to combinations of V and a full NP object; (c) impersonal constructions are much less receptive to stranded *to* than personal constructions (contrast the 3 impersonal examples in Table 5b with the 128 personal examples in Table 5a, a disparity that only increases if the remaining data are added in); (d) the attested examples are concentrated in just a few of the construction types listed in Table 1, with only one degree modifier (*too good not to*), only one example perhaps

classifiable as an adverbial subordinate clause (postposed, in *we'd be crazy to*), only one nominalization (*a decision whether to*), and only one object of a preposition (*a question of how best to*); and (e) absolutes, exclamations, and idioms are entirely missing, as are subject complements in subject position, 'mental Equi' examples, and OS Raising cases.

4. Analysis

I now attempt to draw all these observations together. There are three large groups of effects: INTERFERING FACTORS, according to which restructured *to* is unacceptable, or very infrequent, for reasons not directly relevant to the restructuring of *to* itself; TWO MAIN EFFECTS, conditions specifically on the restructuring of *to* which I believe to hold for speakers of English in general; and several SECONDARY EFFECTS, conditions specifically on the restructuring of *to* which hold only for certain groups of speakers.

4.1. *Interfering factors*

Particular examples of stranded *to* can be unacceptable, or not very acceptable, for a number of essentially irrelevant reasons: the source construction itself might be unacceptable; conditions specifically on VPD might apply; VPD might clash stylistically with other features of the examples; the surface configuration resulting from VPD might on its own be less than fully acceptable; VPD might be in competition with other elliptical constructions; or universal conditions on phrasing might explain the unacceptability, without any need for a statement covering one English morpheme. I illustrate these possibilities in turn.

4.1.1. *Unacceptable sources.* For the first case, observe that the unacceptability of (16) follows directly from the unacceptability of its putative source, (17). And the relative unacceptability of (18):

- (16) A: I've been trying very hard to avoid crowds.
 B: *Oh, has to been a problem for you?
- (17) A: I've been trying very hard to avoid crowds.
 B: *Oh, has to avoid crowds been a problem for you?
- (18) *Well, I would jump away from a snake, but *George* just doesn't know to.

is surely related to the fact that the 'mental Equi' construction is unnatural for many modern American speakers.

4.1.2. *Conditions on VPD.* For situations where it is natural to refer to conditions applying to VPD, recall cases like (19), discussed first in Kuno, 1975:

- (19) A: With whom did John want to go to Paris?
 B: a. He wanted to go there with Mary
 b. *He wanted to with Mary.

Levin (1980, ch. 3) reviews the fairly extensive literature on this constraint, which is general for VPD, affecting not only VPs with *to*, but also all other types; the proposed analyses have involved either a constraint specifically on VPD or a more general constraint on the applicability of transformations, but in neither case does the infinitive marker *to* merit special attention. Moreover, the unacceptable (19b), with VPD, contrasts with the acceptable (20), where the *to* of (19a) has been separated from its VP by parenthetical material:

- (20) He wanted to, as I understood the plan, go there with Mary.

Similarly, various idiomatic V + *to* combinations resist VPD for many speakers, as a special case of the well-known (though, imperfect) generalization that idioms resist deformations, including ellipsis; cases that are unacceptable for many speakers are illustrated in (21), while examples like (22) indicate that VPD is specifically at fault.

- (21) a. *First Helen started to leave, and then Helen *máde to*.
 b. *Isaac thinks green demons are out to get him, but I don't think they're *óut to*.
 c. *At 10, Janet went to see who was at the door, and a few moments later, Kevin *wént to*.
 d. Do I think you're silly? ?*Well, recently I've *cóme to*.
 (22) Sometime later, Helen *made to* — it was entirely unexpected — leave the party without further word.

4.1.3. *Stylistic discord.* Next, the informal, conversational character of VPD with *to* may clash with the formal style of the construction in which or with which it occurs. I believe that such stylistic discord is sufficient explanation for the absence, in the data collected from texts, of VPD within subject complements in subject position (as opposed to extraposed subject complements), in absolutes, and in exclamations, all three constructions being formal. Thus, I would claim that examples like those in (23) are grammatical, but somewhat awkward because of the stylistic clash in them, while those in (24) and (25) are beyond the pale, thanks to the great formality of infinitive absolutes and infinitive exclamations in modern English.

- (23) a. ?You *must* buy a hunting license; *not* to is illegal.
 b. ?For *adults* to curse isn't surprising; for *children* to usually is.

(Speakers who don't much like sentences 4 and 7 in Table 2 will tend not to like (23b), either, but this effect is independent of the relatively mild stylistic clash at issue here.)

- (24) *We hoped that Herbert would return; but he ran out, *not* to.
 (25) *They asked if I wanted to be myself again — Oh, if *only* to!

Once again, VPD is specifically at fault; compare (24) with

- (26) Herbert ran out, *not* to — as it turned out — return for several days.

4.1.4. *To to*. Next, VPD (and parenthetical interruption, for that matter) can give rise to unacceptable surface configurations, such as two occurrences of the infinitive marker *to* in succession. In general, two identical occurrences of leaners in succession are either unacceptable or (as for English *that that* sequences) stylistically disfavoured. In the case of *to*, such sequences would arise if VPD applied to an infinitive phrase serving as the object of a SO Raising verb like *believe*. That is, from a structure like that in (27a), SO Raising would derive the (relatively) acceptable (27c), and VPD would then yield the unacceptable (27d). Example (27d), in fact, has several strikes against it.

- (27) a. I believe (that) to do so would be troublesome.
 b. *I believe to would be troublesome.
 c. I believe to do so to be troublesome.
 d. *I believe to to be troublesome.

If (27b) rather than (27c) is treated as the immediate source of (27d), then the unacceptability of (27d) follows from the unacceptability of its source. However, the proposals I will make below rule out (27b) and the examples in (10) on the basis of their surface configuration (*to* representing the subject of a clausal object of *believe*), but do not rule out (27d) (with *to* representing a phrasal object of *believe*), so that I would not want to press an explanation based on the unacceptability of (27b). In any case, (27c) is rather awkward; SO Raising applied to complex or heavy NPs is in general less acceptable than when applied to simple NPs. And SO Raising with *believe* is distinctly formal, so that the colloquial VPD in (27d) is in conflict with it. On top of all this, (27d) has the infelicitous *to to* in it.

Matters are improved slightly if we use a verb like *expect*, for which SO Raising is not so stylistically marked as for *believe*. Note that (28a) is very much better than (28b).

- (28) a. I expect him to.
 b. ??I believe him to.

Now consider the parallels to (27b) and (27d): (29a) is just as hopelessly unacceptable as (27b), but (29b) improves somewhat on acquaintance — though there is no getting round the awkwardness of *to to*.

- (29) a. *I expect to would be troublesome.
 b. ??I expect to to be troublesome.

4.1.5. *Alternative constructions.* Next, judgments on VPD stranding *to* are confounded by the availability of various alternative forms of expression, each of which has its own subtle pragmatic and stylistic concomitants. In particular contexts, some of these will seem more appropriate than stranded *to*, and consequently stranded *to* might seem less than perfectly acceptable in comparison, though the reasons for these delicate judgments will be extremely difficult to explicate. Consider a real-life example from the Hareven/Langenbach Amoskeag interviews (p. 286):

- (30) By the time I was sent upstairs, where my sisters were working, I knew how to spin, but I learned how to do it with a stranger.

The speaker could have repeated *spin* (*I learned how to spin*), or used *do that* or *do so* instead of *do it*, or applied VPD either to strand *to* (*I learned how to*) or to delete the entire marked infinitive phrase (*I learned how* — my own favourite among these alternatives, though I cannot say why), or deleted the entire object of *learn* (*I learned*). If the original VP had been copular, for instance (*I knew how*) *to be a spinner's helper*, then VPD stranding *be* would have been a further possibility: *learned how to be*. In the particular context at hand, it seems to me that *I learned to* is certainly possible, but it is not one of the best of the available alternatives, and some judges might award it a question mark rather than a clear *ok*.

4.1.6. *Universals of readjustment.* Finally, there are universal conditions on readjustment, in particular those following from Rotenberg's (1978) proposed universal structural conditions on phrase-internal phonological phenomena, like the English *a/an* alternation and French liaison. According to Rotenberg, three situations block such phenomena: sentence edges, parentheticals, and traces. The same three situations should also block the attachment of stranded *to* to preceding material. These predictions are straightforwardly borne out for sentence edges (boundaries between two adjacent Ss) and parentheticals, as illustrated in (31) and (32) below.

- (31) Although it would distress us for you to leave, $\left\{ \begin{array}{l} \text{to leave} \\ *to \end{array} \right\}$ is what I'd advise you to do.
- (32) a. You should leave now, because not to, if you must know, would be rude.
 b. *You should leave now, because not, if you must know, to would be rude.

The trace case is more difficult to test, given the great variety of proposals that have been made for the location of traces. It is clear that some EMPTY nodes, as in (33a), do block attachment of *to* to the left, even when the preceding word is a verb, as in (33b).

- (33) a. A: Why did he do it?
 B: I suppose *e* to impress his parents.
 b. A: Why did he do it? To impress his parents?
 B: *I suppose *e* to.

Not all traces of moved constituents invariably have a blocking effect, however. Example 5 of Table 2, repeated here as (34), involves a *wh* trace, but responses on this item were polarized, with a substantial number of subjects accepting it.

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

4.2. *First main effect*

The main effect to be observed is a structural one, closely related to Rotenberg's 'sentence edges' case. The most relevant contrasts are illustrated in (35) and (36), with various types of complements.

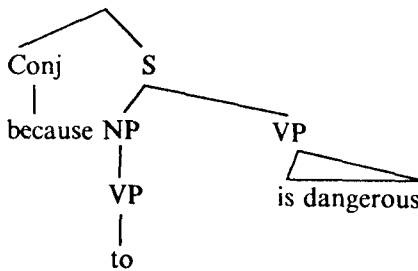
- (35) a. *You shouldn't play with rifles, because *to* is dangerous.
 (= (10a))
 b. You must write a thank-you note, because not to would be impolite.
 c. She wanted to ask his age, but she realized it would be rude to.
 (= 19 in Table 2)
- (36) a. *She'd like to surprise him, but I don't know whether *to* is possible. (= (10c))
 b. *You can try it, but I believe *to* will cause trouble.
 c. I want to calculate the bill, but I don't know how to. (= 21 in Table 2)
 d. %I'm thinking of moving to Hawaii, but I can't figure out whether *to*. (= 11 in Table 2)

In (35a), *to* as a remnant of a VP serving as a subject complement fails

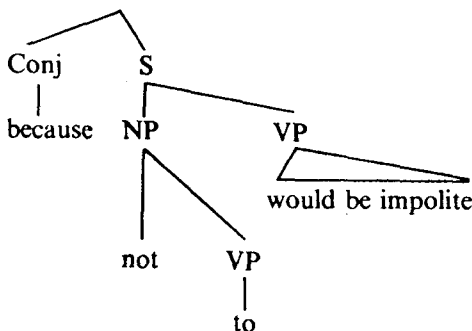
to attach to a preceding subordinating conjunction, *because*; while in (35b), *to* as a remnant of a VP in a subject complement attaches to a preceding adverb, *not*; and in (35c), the extraposed version of a *to* as in (35a) attaches to a preceding adjective, *rude*. In (36a), *to* as a remnant of a VP serving as a subject complement itself in an object complement fails to attach to a preceding complementizer, *whether*, and in (36b), *to* to a preceding verb, *believe*; while in (36c), *to* as a remnant of the VP in an object complement attaches to a preceding complementizer, *how*, and in (36d) — with some variation in acceptability — to *whether*.

These verbal descriptions scarcely bring out the relevant factors. Consider now schematic constituent structures for the relevant subparts of the examples in (35) and (36). In these structures I have labelled a constituent as an S only if it has the surface components of an S, namely, both some remnant of a subject NP and some remnant of a VP.⁸ This decision will be used in distinguishing the different *whether to* structures in (36a) and (36d), although these structures are distinguishable under other assumptions about surface structure (for instance, under the assumption that all surface VPs *to* are Ss). I have also suppressed \bar{S} nodes, since they turn out to lie above the level of constituent structure that is relevant. And I have suppressed the internal constituent structure for most VPs, since this too turns out to be irrelevant in most cases.

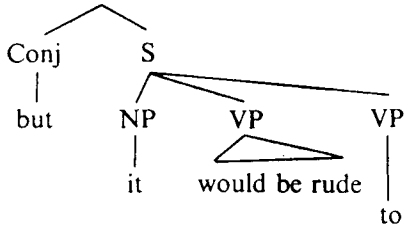
(37) a. *



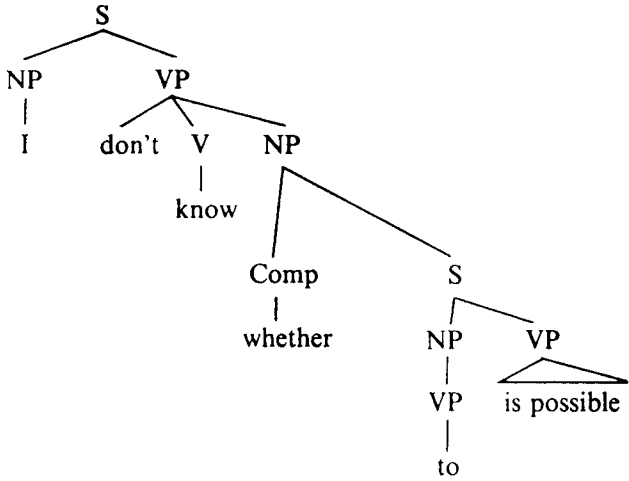
b. ok



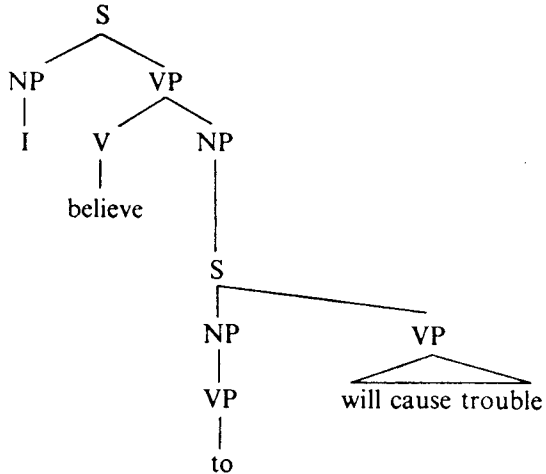
c. ok



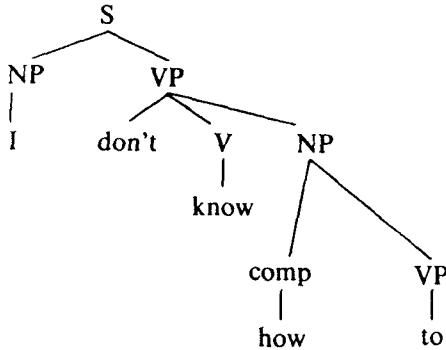
(38) a. *



b. *



c. ok



((36d) has exactly the same structure as (36c), with *whether* in place of *how*.)

What, then, blocks *to* from attaching to the left in (37a), (38a) and (38b), but permits it to in (37b), (37c) and (38c)? In the blocked cases, *to* would have to move out of its S (the constituent dominated by the first S node above *to*), whereas in the permitted cases *to* attaches to material within its S. Thus, I propose that the readjustment rule in (39a) is subject to the condition in (39b), as well as to Rotenbergian universal conditions:

- (39) a. (*To* Reattachment, or 2R) When it does not form a VP constituent with an immediately following VP, the English infinitive marker *to* attaches to the constituent immediately to its left, to form a phonological phrase with it,
 b. (The Own-S Condition, or OSC) except that it cannot move out of its surface-structure S.

The analysis is not complete without an account of the details of this attachment operation: what is the 'constituent immediately to the left' of *to*, and what is the constituent structure resulting from the attachment? The appendix to Zwicky, 1980 proposes an answer to the first question, and the operation of Chomsky-adjunction can at least plausibly be appealed to for the second. These issues do not, however, bear on the formulation of the conditions in which attachment takes place.

The analysis must also include a statement equivalent to the assertion that *to* is an obligatory leaner; that is, if 2R is unable to alter the phrase affiliation of an instance of stranded *to*, then the result is simply unacceptable. Formally, this is a surface filter, applying (like surface filters in general; see Zwicky, 1982 for a summary discussion) after all operations of cliticization and readjustment:

- (40) (The Stranding Filter, or SF) Structures containing VP exhaustively dominating the infinitive marker *to* are unacceptable.

Note that SF is required independently of OSC, for SF applies also in situations where there is nothing whatsoever to the left of *to*, in which case the readjustment 2R cannot apply. This happens when *to* is in absolute sentence-initial position, as in (6) above, with **to* as a one-word answer to an alternative question.⁹

The analysis in (39) and (40) makes a number of predictions about structures which turn out to be subject to significant variation. In discussing these cases, I will momentarily abstract away from this variation. That is, I would treat (39) as falsified only if it classified as acceptable a construction rejected across varieties or as unacceptable a construction accepted in at least some varieties. This is equivalent to testing (39) with respect to what I earlier called the WIDE variety of English.

First, (39) predicts a crucial difference, for subject complements in subject position and for preposed adverbials, with respect to whether or not the subject of the construction is expressed: if the subject is expressed, then a stranded *to* will have something in its S for 2R to attach it to, namely this subject; but if the subject is not expressed, then *to* is truly stranded. These are the contrasts in (41) and (42).

- (41) a. %I don't think you children should play with rifles; in particular, for George *to* would be very dangerous.
 b. *I don't think you children should play with rifles; however, *to* would probably amuse you.
- (42) a. %It's not easy for anyone to justify such attitudes, and for Helga *to* she'd have to be exceptionally persuasive.
 b. *It's not easy for Helga to justify her attitudes, and in any case *to* she'd have to be exceptionally persuasive.

Then, (39) predicts a difference between postposed and preposed adverbials, parallel to the difference between extraposed subject complements, as in (35c), and such complements in subject position, as in (35a): a sentence-final *to* will have something in its S for 2R to attach to it, namely whatever comes at the end of the VP, but a sentence-initial subject *to* is truly stranded. This is the contrast in (43).

- (43) a. %It's not easy to justify your attitudes, and besides you'd have to talk fast *to*. (= 6 in Table 2)
 b. *It's not easy to justify your attitudes, and besides *to* you'd have to talk fast.

I believe that the examples considered so far exhaust the relevant types

of constructions in which OSC, which refers implicitly to S-initial *to*, would apply. In all cases the condition correctly predicts unacceptability. On the other hand, there are a number of cases in which OSC does not apply but which are nevertheless more or less unacceptable to everyone, against prediction. Most of these involve interfering factors already considered, but at least one further case calls for an additional condition, independent of OSC; this I discuss in section 4.4 below.

4.3. Alternative proposals

Though OSC covers the main facts presented so far, it is not the only analysis consistent with many of them. Several alternative proposals for the first main effect have been suggested to me, and I should like to consider each of them briefly.

4.3.1. Prosodic weight. One very attractive idea builds on two facts: that the phonological units at issue are (at least) units for the assignment of stress patterns, and that the clearest acceptable cases involve *to* attached to a stressed word immediately preceding it. The proposal is then that *to* can attach to the left if, and only if, the word in that position bears sufficient prosodic weight. One sort of datum that suggests very strongly that prosodic effects must be taken into account (if only as one of the 'interfering factors') is the relative unacceptability of sentences like (44) with *for* Pro *to* — this despite the fact that *for* NP *to* is acceptable for many (see 41a) and that V Pro *to* is acceptable for virtually everyone (on the evidence of its frequency in the text data).

(44) ?I don't think George should play with rifles; for him to would be very dangerous.

The problem with (44) is a kind of prosodic standoff: *to* wants to lean on (anaphoric) *him*, which in turn wants to lean on the preposition, *for*, of which *him* is the object; but at the same time *for* in this construction is itself a leaner, serving in fact to mark the function of its NP object as the subject of the infinitive verb that follows; and the result is that NONE of the words *for*, *him*, or *to* can naturally bear the stress for the phrase they constitute.

But such accounts cannot serve as general explanations for the failure of *to* to attach to preceding words. On the one hand, conjunctions like *because*, which easily take phrasal stress, cannot attract stranded *to*. Similarly, proper names like *Mary* (7 in Table 2) and indeed non-pronominal NPs in general, which certainly bear phrasal stress, are (for

some speakers) unacceptable with *to* attached to them. On the other hand, unstressed pronouns (themselves leaning on preceding verbs) attract stranded *to* for everyone; in such cases phrasal stress falls two words, and up to three syllables, before *to*, as in

(45) a. ... it didn't quite sound like we wanted it *to*.

(Jerry Garcia, *RSI*, p. 196)

b. ... most bachelors have felt the same way, when they have allowed themselves *to*.

(Lopate, *Bachelorhood*, p. 278)

In any case, this proposal fails to distinguish different occurrences of *whether to* (unacceptable for everyone in (36a), acceptable for some in (36d)) and different occurrences of V *to* (universally unacceptable cases like *believe to* in (36b), universally acceptable cases like *prefer to* in 8 of Table 2).

4.3.2. *Clause-final position.* A second proposal is incorporated into the descriptions of stranded *to* in both the *OED* and *WNI3*: 'Used absolutely at the end of a clause, with ellipsis of the infinitive, which is to be supplied from the preceding clause' (*OED*); 'often used by itself at the end of a clause in place of an infinitive suggested by the preceding context' (*WNI3*). Now a quick glance at Table 1 shows that most infinitive constructions are in fact clause-final, so that stranded *to* will in the nature of things also be clause-final most of the time. Nevertheless, the text data demonstrate that insofar as other material can occur after infinitive constructions, it can occur after stranded *to* as well; (12b) has *I'd love to eventually*, (12c) *if I have to in life*, and of course all the examples involving parenthetical interruptions have material following *to*. Most speakers also accept subject complement constructions like *not to* in (35b), and many accept *for George to* in (41a); in both cases the stranded *to* is part of a subject.

Indeed, Dwight Bolinger has suggested to me that material following the *to* might actually IMPROVE the acceptability of stranded *to* constructions (he speculates that this material in some sense replaces the VP that is missing after the *to*). He offers contrasts like those in (46) and (47) below; I find the judgments delicate, but I certainly find the (a) examples acceptable.

(46) Q. Have you seen him?

A. a. I'm off to right now.

b. ?No, but I'm off to.

(47) Q. You planned to deliver the stuff, didn't you?

- A. a. Yeah, I came to but Joe was gone so I had to give up the idea.
 b. ?Yeah, I came to.

4.3.3. *Immediately preceding category.* A third proposal is based on the fact that the class of constructions to which stranded *to* can attach is really quite limited. Of course, OSC, in combination with restrictions on the positions in which marked infinitives can occur in the first place, rather reduces this class of constructions. The questions are then: just how limited is this class? And which factor is fundamental, OSC or a condition on the category of material immediately preceding *to*?

Table 5 suggests that the acceptable immediately preceding categories can be enumerated very briefly: V, V + Pro, *be* + Adj, *not*, Comp containing *wh*, a small class of idiomatic *have* + N combinations. However, this is not a list of the ACCEPTABLE preceding categories, but instead a list of the FREQUENTLY OCCURRING ones, which is not the same thing. In the wide variety, stranded *to* can appear in complements with subjects (*for George to*) and in postposed adverbials (*talk fast to*), so that it can occur following any category whatsoever.

Moreover, the short list of categories extracted from Table 5 is too inclusive, since it does not distinguish the different *whether to* constructions or the different V *to* constructions.

Even if this short list were adequate, however — as it nearly is, for the narrowest varieties — I would still want to reject it as an analysis. To begin with, the list is a grab bag of categories, in no ways united, and the complement of the list is no more a natural class than the original. More important, I would want to reject ANY analysis that attached *to* to a preceding constituent entirely on the basis of the category to which that constituent belonged, without reference to the structures in which these elements find themselves. I am not rejecting the idea of categorial constraints on readjustment — in fact, my second main effect and the secondary effects involve reference to categories — but merely proposing that readjustment rules should be at root structurally conditioned. Whether or not this general proposal can be maintained, of course, OSC is a better account of the first main effect than a list of categories to which *to* can attach.

4.3.4. *Dominating mode.* Another alternative, suggested to me by Christopher Longuet-Higgins, is drawn directly from the tree configurations in (37) and (38). Longuet-Higgins observed that in the unacceptable configurations, the VP exhaustively dominating *to* is itself exhaustively dominated by another node, namely NP, while this is not true in the

acceptable configurations. In these, the VP mother node is one of several subconstituents of an S; as in (37c), or it is one of two subconstituents of an NP (as in (37b), where this VP has *not* as a sister, or in (38c), where it has the complementizer *how* as a sister). the proposal is then that there is a surface filter, roughly as in (48), which I will call the L-H FILTER, after its deviser.

- (48) Any surface structure in which some node other than VP exhaustively dominates *to* is unacceptable.

The L-H Filter makes the same predictions as (39) plus (40), not only for the cases in (35) and (36), but also for a number of others.

S-initial *to* will be part of either a subject NP or an adverbial construction, and *to* beginning an answer fragment will be part of either an object NP or an adverbial construction, so that a stranded *to* in any of these positions will be dominated not only by VP but presumably also by NP or Adv. The L-H Filter then predicts that all such constructions are unacceptable, and so does SF, since the stranded *to* has nothing preceding it to attach to. These predictions are not altered, in either analysis, if a conjunction or complementizer precedes the S of which the *to* is the first word, as in our original unacceptable examples in (10) above; in the L-H Filter analysis these are unacceptable by virtue of their offending NP-over-VP-over-*to* or Adv-over-VP-over-*to* configurations, and in my analysis they are unacceptable because OSC prevents readjustment, causing the resultant surface structures to fall prey to SF.

Note that both analyses must include a readjustment rule like 2R. This is required in any case to relate phonological phrasing to the constituent structure provided by the syntactic component. My analysis posits in addition a structural condition on readjustment and a surface filter, SF, whereas the L-H analysis adds only a surface filter, (48). With respect to the complexity of the analysis, then, the L-H formulation seems to have the advantage over mine.

This advantage is illusory, however. Consider how the readjustment rule and the filter must interact in the L-H analysis. The readjustment rule cannot apply first, for if it did it would go about blindly attaching *to* to preceding conjunctions, complementizers, adverbs, anything, thereby destroying the offending structures. So the filter must do its work before the readjustment rule. But this order of application is exactly the reverse of the order predicted ON UNIVERSAL GROUNDS. There is considerable evidence (summarized in Zwicky, 1982) that surface filters as a group always apply after readjustment and cliticization, indeed after morphemes have been assigned phonological shapes and morphophonemic rules have applied. The L-H analysis cannot even be entertained, if we are to

maintain these general hypotheses about the interactions among components of a grammar.

There is, in addition, evidence from stranded *to* facts against the L-H analysis. First, if any predicate complements (group II in Table 1) at all are NPs in surface structure, the L-H Filter will reject *to* as a remnant of these complements. An NP structure is quite natural for the complements of some Equi-type verbs with subject control — for instance, *prefer*, which occurs with NP objects generally — and is at least arguable for the complements of Equi-type verbs with object control, like *persuade*. But stranded *to*, as in *prefer to* and *persuade him to*, is uncontroversially acceptable in these constructions. The L-H analysis also does not extend easily to the wide variety of English. If extraposed subject complements are dominated by NP, and if postposed adverbials are dominated by Adv, then the L-H Filter rejects both, and I see no natural way to amend the filter so as to reject *to* as a preposed adverbial but accept it when postposed, or to reject *to* in subject position but accept it when extraposed. Finally, the L-H Filter permits some constructions that are unacceptable in the narrow variety. In particular, it does not reject *for George to* (here *to* is simply a remnant of a VP), although some speakers find it unacceptable; and if it does not reject *I want to* it will also not reject *Who do you want t to?*, although some speakers reject the latter while accepting the former.

4.3.5. *Conditions on binding.* Finally, several colleagues have suggested to me that the work of OSC might be accomplished instead by conditions on the binding of anaphoric elements. The proposal would be to extend the treatment in such works as Chomsky (1980, 1982) to cover VP anaphors, presumably by treating *to* itself as an anaphor. No specific analyses have been suggested to me, and I am reluctant to make proposals of my own within such a rapidly changing theoretical framework. In addition, this line of inquiry does not seem particularly promising to me. Especially problematic are (a) the contrast in acceptability between *not to* and plain *to* as subject complements in subject position (I can imagine conditions-on-binding treatments of the other facts in (35) and (36)); (b) the contrast in acceptability between subject complements with expressed subjects (*for George to*) and those without them, as in (41) and (42); and (c) the contrast in acceptability between postposed and preposed adverbials (as in (43)). Here it is hard to see how binding conditions can be strong enough to do the job.

In at least one case, it appears that a conditions-on-binding analysis would be too strong. I have already pointed out a contrast, for some but by no means all speakers, between *how to* and *%whether to*, as in (36a, d).

Informal judgment collection suggests that *when to* and *where to* behave much like *how to* (or are somewhat less acceptable); if so, we appear to be faced with a contrast between the acceptable moved wh word *whether* and the dialectally unacceptable unmoved wh word *whether*, a contrast for which a suitable binding condition might be devised. But it seems to me that any set of binding conditions ought to rule out *who(m) to*, with *who(m)* representing an NP within the (null) VP, as in

(49) I know I ought to see *some* psychiatrist, but I don't know *who* to.

I find such examples much less acceptable than the *whether to* cases, but Nomi Shir reports that the *how-when/where-whether-who(m)* cline of decreasing acceptability is just that; for some informants, in her experience, examples like (49) are as unquestionably acceptable as the other *who + to* constructions.

4.4. Second main effect

The analysis so far does not cover two examples already presented but not discussed, sentences 9 and 13 in Table 2, repeated here in (50).

- (50) a. *I didn't want to go, but what he wanted wás to.
 b. *Í didn't want to go, but what hě wanted wás to.

Both are universally unacceptable, despite the fact that neither OSC nor any universal condition on readjustment blocks the attachment of *to* to a preceding form of *be* in wh clefts. I now claim that this unacceptability is in no way specific to wh clefts, but in fact arises whenever the word preceding *to* is a form of *be*. I will call this the B-2 CONSTRAINT.

The constraint can also be observed in *it* clefts of adverbials, as in (51); in 'distributed nominalizations' (Culicover, 1977), as in (52); and in other equational sentence types, as in (53):

- (51) *You might have heard that it was to search for the elixir of life that Harold left home as a teenager — well, in fact, it wás to.
 (52) *Their decision is not to spend the million immediately, but ours is to.
 (53) a. *Kevin thinks that to visit Philadelphia is to experience big city life, but I can't imagine how anyone can believe it is to.
 b. *The reason I went was to please the Smiths, and apparently the reason you went was to, as well.

In fact, the constraint accounts for a very peculiar gap in the attested V *to* constructions in Table 5: though such idiomatic combinations as *have*

to or *used to* are very frequent, or (in the case of *got to* 'must'; *be able to* 'can'; *be going to* 'will'; *ought to* and *be supposed to* 'should') are attested at least twice in my data, the verbal idiom *is to* is completely absent. Constructed examples do not strike the ear well:

- (54) a. *I am to arrive with a pick-axe, that is so, but you are to, also.
 b. *He was not to speak Danish, but you were to.

There is some question about the status of the B-2 Constraint. Is it a condition on readjustment, like OSC; or a surface filter, like SF; or simply a condition on VPD? This last possibility hinges on whether parenthetical interruptions yield the same sort of unacceptability as VPD or not, a judgment I find hard to make. What is at issue is whether examples like those in (55) are acceptable, in which case the B-2 Constraint is a condition specifically on VPD; or not, in which case it either is or invokes a surface filter.

- (55) a. ?It was to, as I believe you know, search for the elixir of life that Harold left home as a teenager.
 b. ?You are to, you know, arrive with a pick-axe.

These examples are probably bad enough to warrant treating the constraint as parallel to OSC or SF.

To choose between a further condition on 2R and a further surface filter, I observe first that unlike SF, a B-2 filter would have no independent justification, since the only source for a constituent consisting entirely of a form of *be* and the infinitive marker *to* is 2R. Moreover, the various secondary effects I am about to consider are just like the B-2 Constraint in this respect, and some of them have no natural formulation as surface filters. Consequently I opt for an analysis with a second condition on 2R:

- (39) c. And except that *to* will not attach to a form of *be*.

I do not know the range of judgments on *be to* examples; the sentences in (50) were tested with my subjects, but those in (51)–(55) were not, and the judgments above are my own. Not everyone finds all the examples equally bad. Jacques Transue and Dwight Bolinger have independently pointed out to me that stranded *be to* is somewhat more acceptable in conditional clauses than elsewhere, and that in one conditional construction it is perfectly acceptable; this is the counterfactual infinitival *were to* illustrated in VI.1 of Table 1:

- (56) I don't think you've read this book; if you were to, it would change your life.

I agree that such examples stand in sharp contrast to cases like (50)–(54). It follows that the B-2 Constraint must be relaxed in some way, but the details are not clear to me at the moment.

I have found no examples of stranded *be to* in print, nor have I heard any during the three years I have been listening with an ear attuned to marked infinitives. The closest example I have found is one that actually supports the constraint as stated in (39c). This is a piece of interior dialogue in a novel:

- (57) But how was he to?
 (Andrew Holleran, *Dancer from the Dance* (New York: Bantam Books, 1979), p. 69.)

Here the *to* attaches not to a form of *be*, but rather to the unit *was he* (itself created by cliticization of the pronoun subject *he* to the preceding verb *was*). The phrase *was he to* ought then to contrast with the phrase *he was to*, where no subject-verb inversion has taken place; the expected contrast is between the direct questions in (58) and the corresponding indirect questions in (59).

- (58) a. But how was he to?
 b. But was he to?
 (59) a. *He wondered how he was to.
 b. *He wondered if he was to.

The judgments indicated, which are mine, are as predicted by (39c).

4.5. Secondary effects

I now return to the messy data of sections 3.2 (judgments on the examples of Table 2) and 3.3 (examples gathered from texts, summarized in Table 5). There are three effects that can be extracted from these data. One of these, the VP Condition, is central; the other two are further specializations of it. My formulations of these conditions will appear to be much too strong, in that they will rule out acceptable, even reasonably common, constructions. But these constructions can be seen to belong to a natural class, which I will treat (in section 4.7) as a systematic set of exceptions to the VP Condition.

I begin with examples 7 (%for *Mary to*) and 11 (%whether *to*) of Table 2, since I argued in section 3.2 that the acceptability of either of the other groups of middle examples I considered there, 10/12 and 5/6, tends to imply the acceptability of 7/11. In conjunction with 7/11, consider also the very skewed occurrences in Table 5, in particular the complete absence of *to* attached to a nonpronominal NP and the small number of cases (three)

of *to* attached to a *wh* complementizer. Most of the attested cases, in fact — 130 out of 157 — have *to* attached to V or *be* + Adj. One of these (*we'd be crazy to*) involves what is perhaps a postposed adverbial *to*. In any case, the condition formulated for 7/11 should not also rule out 6(%*you'd have to talk fast to*) or 10/12 (%*it would help to*, %*I don't think it's possible to*). What unites these cases we do NOT want to rule out, as opposed to 7/11, is the character of the constituent immediately preceding *to*: in the cases we do not want to rule out, this constituent is a verbal construction, either a whole VP (for extraposed subjects and postposed adverbials) or V or *be* + Adj (the remaining cases, all various types of predicate complements). My proposal is then that 7 and 11 of Table 2 are blocked (for some speakers) by something like the following condition on readjustment, the VP Condition:

(60) The constituent preceding *to* must be a VP or a predicator in a VP.

For varieties in which (60) applies, *to* cannot attach to *for/of* + NP in any position, to complementizers, to adverbs, or to objects of verbs, so that the constructions in (61) are unacceptable as well as the ones already considered.

- (61) a. %I really needed someone to fix the radiator, so it was awfully nice of Helen to.
 b. %I realized that you sometimes stay out past midnight, but it seems excessive *always to*.
 c. %When they wanted someone to make dinner for 30, they persuaded Quentin to.

My reference to 'predicator' in (60) is designed to pick out V or the Adj of a *be* + Adj construction, and also to encompass verbal idioms of various other shapes, such as *have reason (to)*, *be about (to)* and *get a chance (to)*. Perhaps the particular formulation of the VP Constraint I've given here is imperfect, but something along these lines seems to be needed.

Next, examples 5/6 from Table 2. Example 6 was not well received by my subjects in general, but it is in several respects a poorly chosen instance of its type; something like *you'd have to be a saint to* or *you'd have to be out of your mind to* would have been better. Nevertheless, despite its poor reception, I have argued that sentences of its type should be included in the set of examples that are acceptable for some speakers. Sentence 5 is the example of an intervening *wh* trace: *want t to*. What unites these two cases is the structural chasm between *to* and the preceding constituent, quite reminiscent of Rotenberg's conditions: in the case of postposed adverbial *to*, the readjustment rule has to move it into a preceding S, and

in the other case, the readjustment rule has to move *to* over a *wh* trace. These are muted versions of Rotenberg's 'sentence edges' and 'trace' conditions, respectively. It seems premature to formalize this strengthened version of the VP Condition, given the small data base I am working with at this point.

Finally, examples 10/12 from Table 2. The judgment data are muddled by a number of factors. Examples 10 and 12 are clearly rejected by one group of subjects, and the responses on these two examples are correlated. However, example 1 (*You can try it, but I doubt that it would help to*) was generally accepted, despite the fact that it includes the same impersonal construction, *it would help to*, as example 10. Example 1 also has, alas, two *its*, and their separate and joint reference is not at all clear; in any case, the first three examples were originally constructed as lead-in to the others, and subjects might well have been better disposed to the first half-minute of their experience than to what followed. Example 3 was generally accepted, but it is truly, and for my purposes fatally, ambiguous: *Marge thinks I can't lift this weight, but I think it would be easy to can be followed either by lift this weight (with it understood impersonally) or lift (with it understood 'personally', referring to this weight)*. I am left with examples 10 and 12, both with impersonal subject *it*, versus the clearly acceptable cases like 8 (*I prefer to*), 18 (*he's old enough to*), and 19 (*it would be rude to*).

The trend to be seen in all of this is that everyone accepts *to* attached to a preceding predicator, V or Adj, when this predicator is PERSONAL, that is, when it asserts (*I prefer to*) or implicates (*it would be rude to*) that it holds of a concrete subject. Some speakers reject *to* attached to a preceding predicator, V or Adj, when this predicator is IMPERSONAL, neither asserting nor implicating that it holds of a concrete subject; this is the case with *I don't think it would help to [take some aspirin]* and *I don't think it's possible to [move the piano]*, where the notional subjects of the predicates *help* and *be possible* are sentential. Again, my formulation is made on rather slim evidence, and its details are open to discussion, but I think that some reference to the personal/impersonal distinction in predicators will be required.

4.6. Goal-directedness

Dwight Bolinger has suggested to me an alternative account for many of the data that are subject to dialect variation. This proposal is summarized, in my own words, as (62) below.

(62) The more a construction describes an action or attitude directed

towards some goal, the more acceptable it is when followed by stranded *to*.¹⁰

This hypothesis is especially successful with pairs like those in (63)–(66), in which the (a) sentences express more goal-directedness than the (b) sentences.

- (63) Take that hill in high?
 a. He doesn't have the guts to.
 b. ?He doesn't have the gas to.
- (64) a. He was afraid to undertake it because he didn't have the authorization to.
 b. ?He was afraid to undertake it because he didn't have the letter of authorization to.
- (65) a. I would open it if I had the right tool to.
 b. ?I would open it if I had the tool to.
- (66) a. Enough time to was never available.
 b. ?The time to was never available.

I find these judgments subtle (and some speakers will reject them all), but on the whole I agree with them. The conditions in the preceding sections make no predictions about the cases in (63)–(66), so that for those speakers who agree with the judgment, there appears to be a further semantic condition on 2R (whether this is correctly formulated in terms of goal-directedness is another matter).

However, it seems to me that this condition does not REPLACE the VP Condition and its further elaborations, but rather that it is orthogonal to them. It is true that (62) could be appealed to to account for the greater acceptability of *expect to* than of *believe to* (examples (28) in section 4.1.4), and for the greater acceptability of *how/when/where to* than of *who(m) to* (example (49) in section 4.3.5), and perhaps for several other contrasts already seen. But I cannot see that it will serve in general.

To begin with, it is hard to see *used to*, one of the stranded *to* constructions with the highest text frequency, as expressing goal-directedness in modern English. Consider these quotations from Ronald Blythe's *Akenfield* (New York: Pantheon Books, 1969):

- (67) a. The young men don't hate the land any more — they used to but they don't now. (p. 133)
 b. I am thirty-one and I don't get as excited about things as I used to. (p. 260)

Similarly with some occurrences of *deserve to*, *afford to*, and *be about to*:

- (68) Rooney can make that kind of joke sound fresher than it deserves to ... (Robert MacKenzie, 'Review' in *TV Guide* 30.12.40, 12 March 1982)
- (69) By 1970, the Western world at least appeared to have accepted contraception as a personal matter. But the Western world could afford to, after almost a century of declining birth rates and increasing food supplies. ... (Reay Tannahill, *Sex in History* (London: ABACUS paper ed., Sphere Books, 1981), p. 402)
- (70) ... *well* ... is placed ... between the question that just occurred and the answer that is about to. (Lawrence Schourup, Ohio State University Ph.D. dissertation draft, March 1982)

All these examples should be relatively bad, given the proposal in (62).

On the other side of the issue, the condition in (62) predicts that *to* as a remnant of a reason adverbial should be relatively acceptable, since the clauses with which such adverbials occur usually describe goal-directed activities. Yet cases like *You'd have to talk fast to* — recall example 6 in Table 2 — appear to be acceptable only to a minority of speakers.

4.7. *The exception clause*

It is all very well to propose a dialect subject to the VP Condition, with (independent) subdialects subject to further 'close construction' and 'personal interpretation' conditions. But there are abundant counter-examples in the data already cited. Nearly everyone accepts *to* attached to *not*, which is certainly not a verb or an adjective (the closest major word class is that of adverbs, like *always*); nearly everyone accepts *to* attached to *persuade him*, though *him* is an NP like *George*; and nearly everyone accepts *to* attached to *how*, which is a *wh* complementizer just like *whether*. The conclusion is inescapable: the VP Condition must be hedged by a clause exempting certain types of preceding constituents from its strictures.

As it happens, there is some precedent for exception clauses of just the type I need here. Rotenberg (1978), having motivated a structural condition on liaison in French, finds himself obliged to amend his treatment with the stipulation that liaison is ALWAYS made by certain words, namely nonlexical items — members of very small closed classes in the language. Similarly, Kaisse (1983), examining auxiliary reduction (AR) in English, finds that her predictions are too exclusive, and that she must add a stipulation that 'AR may also apply if the element preceding the auxiliary is a (monosyllabic) nonlexical item'.

My immediate problem is to treat *not*, pronouns like *him*, and *how* — as

opposed to *always*, *George*, and *whether* — as a natural class. Rotenberg and Kaisse supply the common thread: *not*, *him*, and *how* are monosyllabic nonlexical items, as opposed to the lexical items *always* and *George* and the disyllabic nonlexical item *whether*. I thus propose (71) as an exception clause to the VP Condition, (60).

(71) But readjustment is always possible if the word preceding *to* is a monosyllabic nonlexical item.

5. Reflections

Three matters still need discussion. First, I do not believe that it is sufficient to describe an aspect of a language without trying to provide some rationale for it — an elucidation of the function it serves in the language and/or a tie to other phenomena of the language and/or an account of how it might have arisen historically. In section 5.1 I discuss my analysis in these terms.

Section 5.2 picks up a thread dropped back in section 2, namely the relationship of the celebrated *wanna* facts to the properties of the learner *to*.

Finally, in section 5.3 I comment briefly on a number of broader issues raised by this study.

5.1. Why these conditions?

Here I begin with the secondary effects. These were formulated largely on the basis of judgment data, but they also predict many of the text data: constructions are missing from, or very infrequent in, texts because they are acceptable only for some subset of speakers (or because of one of the many interfering factors of section 4.1). But this coin has another side. A few types of examples — *to* attached to a preceding form of *want* or of the quasimodals *have* and *used* — are enormously more frequent in texts than all the other types.

Now consider the task facing a child acquiring English. The child hears a respectable number of cases of *want to*, *have to*, and *used to* (and not irrelevantly, the child also hears a huge number of cases of *wanna/wantsta*, *hafta/hasta/hadda*, and *useta* WITH VPs following them), from which it is natural to conclude that the language permits reattachment of stranded *to* to at least certain sorts of preceding constituents. But which ones? At the very least, verbal constructions, in particular those for which *to* is the complement (thus, those with which *to* is in close construction), and in

particular verbal constructions of the class to which *want* and quasimodal *have* and *used* belong (thus, those which are almost invariably understood personally; the exceptions are 'weather and time *it*' constructions, as in *It used to be muggy in Washington* and *It has to be 3 o'clock*, and extended constructions with root interpretations of the quasimodals, as in *It just has to be (the case) that someone cares* and *It used to be (true) that the Christmas season was only twelve days long*). Presumably, the child has also experienced some instances of *not to*, *V Pro to*, and *wh to*, so that it is clear that an exception clause is required. But the child's first task is to induce a generalization from the class consisting of *want* and quasimodal *have* and *used*, to project (in the sense of Peters, 1972 and Baker, 1979), from these data to some rule(s).

However, these three items do not constitute a whole class. True, they all occur with CLITIC *to*, rather than merely the leaner *to*. But other verbal constructions, much less frequent than the core three in construction with stranded *to*, also form phonological words with *to*: *oughta*, *gotta*, *(be) gonna*, *(be) supposta*. And other verbal constructions, much less frequent than the core three in construction with stranded *to*, are quasimodals; *ought*, *be*, *need*, *be about*, *be supposed*, and *be able*, for instance, all express modal notions and are syntactically notable. The points here are that any natural generalization from the three core instances must stretch across morphological, syntactic, and semantic classes, and that SOME natural generalization is called for, because any normal, generalization-loving child in the later stages of learning English¹¹ would (unconsciously, of course) reject the idea that a structural readjustment applied only to *to* after the verb *want* and the quasimodals *have* and *used*. And these points hold even for a child whose entire experience of stranded *to* was with the core instances; the case is stronger for a child who has been confronted with occasional instances of other types.

The child, then, MUST generalize. How far? The generalization could be very conservative, sticking as closely as possible to the salient properties of the three core instances; or it could ignore some of them; or it could ignore them all, out to whatever boundaries are supplied by universal grammar. These three possibilities have now been amply illustrated in my analysis. The very conservatively generalizing child is the adult with the narrow variety; the somewhat more adventurous child is the adult with an intermediate variety; and the bold child is the adult with the wide variety.

I have, of course, no great stake in maintaining that the effects characterizing the intermediate varieties in my analysis — the VP Condition, alone or supplemented by a 'personal interpretation' or a 'close construction' clause or both — are the only ones operative in modern English, or even in asserting that I have got the formulations of

THESE quite right. But I do want to maintain, with almost all the literature on variation, that generalizations speakers make fall along natural dimensions, given the input available to those speakers.

A very similar account can be suggested for the historical development, once VPD has been extended to at least some VPs in construction with *to*. This triggering event itself depends on a reanalysis in which such instances of *to*, historically prepositions, are interpreted as belonging to the same class as other English lexical items in construction with following VPs, that is, as falling in a class with the auxiliary verbs (Pullum, 1981 argues that the infinitive marker *to* simply is an auxiliary verb in modern English, and Schachter, 1978 argues that stranded auxiliaries and stranded *to* are proforms of the same type). Once *to* appears without its VP in some contexts, the deletion can spread to other contexts, again along natural dimensions. Very high frequency verbs occurring with marked infinitive complements, especially those whose sentential complements are always marked infinitives — *want* and the quasimodals are such verbs — are particularly likely first instances for VPD after *to*, since for them it would be easy for speakers to interpret the *to* as forming a unit with the preceding verb instead of with the following VP. But all that is required as background for the modern situation is that VPD after *to* began with a verb or verbs with the salient properties of *want* and the quasimodals.

The extension of readjustment to include nonlexical monosyllables as conditioning context can also be seen as natural. Nonlexical items are those that must be specially listed in a description of the language, rather than chosen from large classes of items in the lexicon. If a nonlexical item frequently occurs in combination with some word, especially a word like *to* with very idiosyncratic distribution, then the combination can easily be interpreted as a unit on its own. Then, when unquestionable combinations of word-plus-*to* become available via the application of VPD after verb-*to* sequences, other closely knit word-*to* sequences are natural targets in an extension of the domain of readjustment. The result is that *to* has the ability to attach to nonlexical items even in the narrow variety. Nonlexical items are typically monosyllabic, and a conservative extension of the domain of readjustment might permit it only for core members of the class of nonlexical items, in particular monosyllables.

Finally, the main effects. The second main effect, the B-2 Constraint, probably results from several contributing factors, including the marginal status of *be* as a verb and the prosodic characteristics of *be*. The fact that forms of *be* bear the main stress in their phrases only in emphasis or contrast makes them poor candidates as supports for leaning *to*. And in any case *be* generally functions as an AUXILIARY verb (note that it contracts with a preceding word, even in copular and quasimodal uses:

He's too short, You're to call at 10), so that it ought to be one of the last candidates for extension of a process that prefers to apply to verbs like *want*, obligative (and stressed) *have*, *try*, and *like*.

The first main effect represents a natural outside boundary of extension for readjustment. I am inclined to believe (but see below) that this boundary is given by universal grammar, and not induced specifically by acquirers of English. That is, I would like to propose a strengthening of Rotenberg's universal 'sentence edges' condition, from one having the effect of prohibiting phrase phonology from operating across]_S s[, to one having the effect of prohibiting it from operating left across s[or right across]_S. For readjustment rules, the strengthened condition says that an item cannot be moved out of its S; this is simply condition OSC generalized to all reattached items in all languages, rather than holding only for English *to*.

To be more precise, the revised condition should be stated as follows:

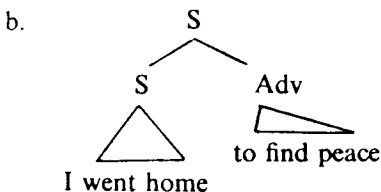
- (72) Readjustment rules cannot move an item so that it would no longer be dominated by its S.

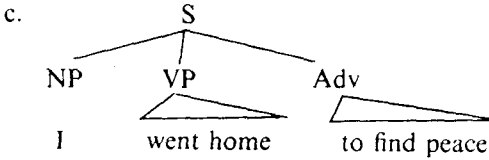
This version has the effect of prohibiting movement left over s[or right over]_S, both of which are movements up into an embedding S or over into an adjacent S. It does NOT have the effect of prohibiting movement right over s[or left over]_S, both of which are movements down into an embedded S. And (72) shouldn't prohibit such movements, for they are attested. For instance, it has long been known that the simple cliticization of English auxiliaries to preceding words can move an auxiliary down into a relative clause:

- (73) This guy who came in from Detroit's been complaining about our weather.

In fact, depending on the surface structure of postposed adverbial clauses in English, they might also provide an example. If the structure of (74a) is as in (74b) rather than (74c), then those who accept the readjustment in (74d) move the *to* down into a preceding S.

- (74) a. I went home to find peace





d. %I went home to.

Condition (72) holds absolutely for 2R, but (as Ellen Kaisse has pointed out to me) can be lifted for the simple cliticization of auxiliaries in English. Thus, although auxiliaries normally will not leave their S to attach to a preceding *wh* complementizer, they will do so if the complementizer is a monosyllabic nonlexical item; (75a) is out because of (72), but (75b) is possible despite the fact that *is George doing* constitutes a surface S.

- (75) a. *What act's George doing?
 b. What's George doing?

But *to* won't attach even to monosyllabic nonlexical items if they are outside its surface S: **that to* and **if to* in (10) cannot constitute phrases, nor can **since to* or **for to*. I have no proposal to offer at the moment for this difference between a simple cliticization rule and a phrasing readjustment rule (certainly more cases must be examined in detail), beyond pointing out that (72) must be restricted to some degree.

To take stock: I have now claimed that the first main effect is not language-particular but universal. The B-2 Constraint, I have suggested, is the product of several facts about *be* in English and probably has no close parallel in other languages. The 'personal interpretation' and 'close construction' conditions might be expected to have parallels, though. And I have already pointed out that the 'monosyllabic nonlexical items' exclusionary cause seems to be required elsewhere in English and in French. What we need now are detailed studies of readjustment, cliticization, and phrase phonology in enough languages to determine what sorts of structural, categorial, prosodic, and semantic conditions grammatical theory must make available for the description of these phenomena. The English *to* data at least seem to be well within the range of phenomena studied by the other students of phrase phonology already mentioned and by others (for instance, Napoli and Nespor, 1979 and Napoli, 1981); I return to these questions in section 5.3.3 below.

5.2. *Readjustment and cliticization*

The recent outpouring of literature on contractions of verbs with *to* (for bibliography, see Postal and Pullum 1979, 1982) is taken up with the analysis of facts like those in (76) and (77); acceptability judgments are for the majority of speakers.

- (76) a. Teddy is the man you want to vanish.
 b. *Teddy is the man you wanna vanish.
 c. Who do you want to vanish?
 d. *Who do you wanna vanish?
 e. Teddy you want to vanish.
 f. *Teddy you wanna vanish.
- (77) a. You want to vanish.
 b. You wanna vanish.

In all these sentences *you* functions as the subject of *want*, but only in (77) does it serve as the subject of *vanish*. One line of analysis invokes traces to distinguish (76b, d, f) from (77b): the movement rules in the former cases must leave traces blocking contraction, while Equi, the operative rule in the latter case, must not. The Postal/Pullum analysis appeals to 'subject sharing' as the relevant difference between the acceptable latter case and the unacceptable former ones.

An important point for my purposes is that there is a 'liberal dialect', discussed by Postal and Pullum, in which ALL the sentences in (76) are acceptable, as well as those in (77). This observation is important because it recalls a similar type of variation in the stranded *to* data: a wh trace blocked readjustment for some speakers in the VPD analogue of (76d), and I assume that on the whole these speakers would also reject the VPD analogues of (76b) and (76f). These analogues are supplied in (78), where suitable preceding context is to be assumed.

- (78) b. %Teddy is the man you want to.
 d. %Who do you want to?
 e. %Teddy you want to.

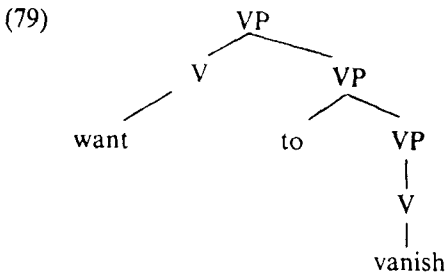
Thus it appears that the traces interfering with 'close construction' for the readjustment of *to* are the same ones blocking the contraction — that is, cliticization — of *to*.

What this parallel suggests (though it does not demonstrate it) is that the cliticized *to* cases are simply a subset of the readjusted *to* cases and are subject to the same constraints. The relationship is strengthened by the facts (already noticed) that the verbs that contract with *to* include the three superhigh-frequency readjustors *want*, *past used*, and *obligative have*, and that the remaining contractors *ought*, *got*, *supposed*, and *going* are among the readjustors with some frequency.

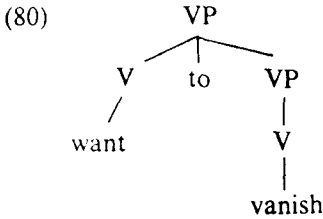
This proposal is testable, but only with considerable labour (which I have not undertaken). The prediction is that if a *wh* trace blocks readjustment of stranded *to* to some verb, then a *wh* trace will block cliticization of *to* to that verb. Counter-evidence to this prediction would consist of speakers for whom a *wh* trace blocks readjustment but not cliticization. But it seems clear that only a small percentage of speakers permit cliticization over a *wh* trace (speakers of the liberal dialect are hard to find, certainly no more than 5–10% of the populations I've dealt with), and these constitute the crucial group; they should reliably and generally accept (76d) and its analogues. The labour comes in assembling a sufficiently large group of indubitable speakers of the liberal dialect, on whom my prediction can be checked.

One potentially sticky point in this proposal is the readjustment rule required. It must be optional, to allow for cases like (76a, c, e), where the phrasing is [*want*] [*to vanish*]. And it must apply to a small set of verbal constructions. It must consequently be some new rule, since it cannot be collapsed with 2R, for 2R applies obligatorily, to stranded *to*, and quite generally (subject to the conditions I have already discussed). Pullum, 1981 has suggested that the required readjustment is nothing other than the Clause Union that has been studied in a number of languages (not including English), for instance in the *faire* + infinitive construction of French. His proposal does not eliminate or explain the fact that the English readjustment is conditioned by a small number of lexical items, which must be listed. In any event, whether the special readjustment can be subsumed under the heading of Clause Union or whether it is simply to be stated as a rule peculiar to English, it seems that something like it is needed if the close relationship between the readjustment and contraction of *to* is to be expressed.

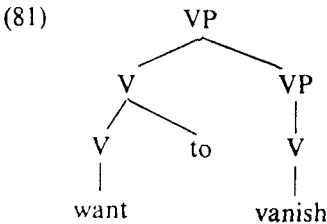
The details of the proposal are fairly straightforward. The special readjustment operates on structures like the one in (79), to detach *to* from its VP.



There are two possible derived constituent structures (see Aissen, 1974 for some discussion of this issue). In one, the *to* becomes a sister of V under VP —



in which case our original readjustment rule 2R will apply, Chomsky-adjoining *to* to the immediately preceding constituent, here V (I assume that 2R works as sketched in section 4.2):



In the other, the special readjustment rule directly Chomsky-adjoins *to* to the preceding V; but this simply yields the structure of (81) in one step rather than two. The ultimate derived constituent structure is (81), with *want to* constituting a phonological word, since it is dominated by the lexical category node V.

5.3. Broader issues

The preceding discussion has opened up at least as many questions as it has answered. There are numerous hypotheses and speculations that have not been tested, but in summary: I have claimed that the readjustment of *to* is subject to a structural condition, OSC, which I defended against several alternative formulations, and which I eventually proposed as part of a universal condition on readjustment. Aspects of these alternatives appear in various forms in the B-2 Constraint, the VP Condition, two further extensions of the VP Condition, and a clause exempting monosyllabic nonlexical items from the VP Condition and its extensions, all three of which are variable. In conclusion, I will take up four broader issues that emerge from these proposals (in addition to the projection problem and the connection between readjustment and cliticization, which I have just discussed).

5.3.1. *Dialect variation.* Syntacticians have been inclined to idealize variation away, treating differing judgments within a group of informants

as essentially random and using either a 'majority vote' or a 'my judgments are the right ones' principle to assign asterisks to examples; or they have atomized variation, treating every differing judgment as evidence of a dialect distinction. This line of criticism is familiar from variationist scholars with an interest in the social and geographical concomitants of variation. But it is not necessary for an investigator to share these interests, or even for there to BE any notable social or geographical concomitants, for certain types of variability to be amenable to analysis.

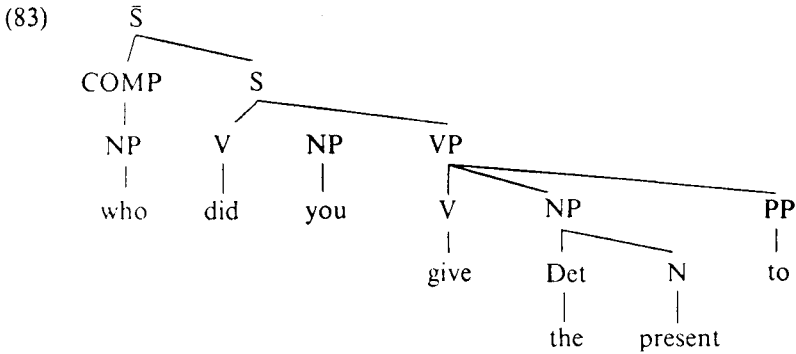
Recall that I detected, and spent some time interpreting, variation in my judgment data without finding any accompanying pattern of social or geographical variation. Indeed, if my speculations in section 5.1 are on the right track, this variation could arise without any significant concomitants. My suggestion there was that the analysis is underdetermined by the positive data available to the child acquiring English. Speaking metaphorically, each child must decide in what direction, and how far (up to OSC), to extend readjustment beyond the very frequent clear cases. The child gets no help from others in this decision, and whatever decision is made will have so infrequent a reflex in the child's own productions that individual differences will go unobserved by others, though they can be observed in adults when a linguist specifically asks them for acceptability judgments on a set of crucial examples. The result is a complex pattern of majority and minority 'dialects'.

5.3.2. *The cliticization/readjustment component.* One of the major claims in Zwicky, 1982 (a concise presentation of joint work by G. K. Pullum and me) is that processes of cliticization and readjustment together constitute a component of grammatical description in any language, a component related to others by strict principles of precedence. In particular, we advance the proposals that syntactic rules, as a set, can feed or bleed rules of cliticization/readjustment (but not vice versa), and that rules of cliticization/readjustment, as a set, can feed or bleed prosodic, morphophonemic, and phonological rules (but not vice versa). For my present purposes, the most important consequence of this view is that rules of cliticization and readjustment are in fact viewed as constituting a component of grammatical organization.

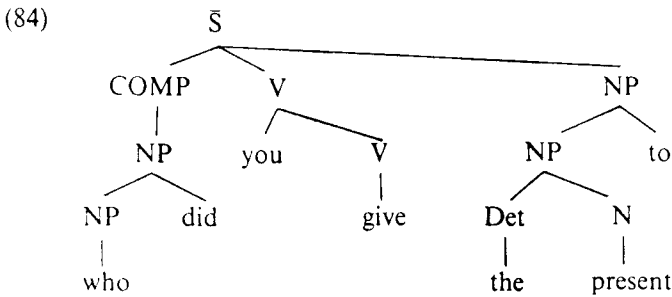
Much remains to be discovered about the nature of the rules in this component and their interactions with one another. Consider even such a simple example as *Who did you give the present to?*, phrased as in (82) below (of course, other phrasings are possible).

(82) [whó did] [you gíve] [the présent to]

The structure provided by the syntactic component is roughly:



At least four readjustment/cliticization rules apply to (83) to give the phrasing in (82): two cliticizations (of an auxiliary verb, here *did*, to the last word of the syntactic phrase immediately preceding it; and of a personal pronoun, here *you*, to a verb or preposition adjacent to it and governing it) and two readjustments (one attaching a preposition, here *to*, to an immediately preceding phrase if that preposition is not in construction with an immediately following NP; and a 'flattening' rule phrasing separately the constituents of a VP or S). The effect of all four together is to alter the structure in (83) to something like:



The rules are all optional, and can interact in various ways to yield a series of different phrasings. The details have yet to be worked out, even for such familiar phenomena in such a well-studied language as English.

5.3.3. The repertoire of conditions on cliticization/readjustment. A closely related line of research concerns the conditions available (in universal grammar) for constraining the operation of rules of cliticization and readjustment and of rules of phrase phonology. The few cases that have been studied in great detail show interesting similarities in the relevant aspects of syntactic structure and properties of lexical items. S

boundaries, left branches of major constituents, and the structural relationship of c-command recur in descriptions of phonological phrase phenomena. The head/modifier distinction probably also plays a role. Monosyllabic nonlexical items are clearly important. And the semantics of individual lexical items and of the constructions in which they occur must routinely be taken into account. Finally, there is the relationship between phrasing and stress, however that is to be described.

I assume that the full set of conditions is not simply the set of all logically possible conditions on rules, but rather is limited in an interesting way; just how it is limited is not yet known.

It is worth pointing out here that conditions like those on stranded *to* in English are not at all uncommon. I suspect that nearly any example, in any language, of readjustment, 'external sandhi', or 'optional pronunciation' would, if examined carefully, turn out to be equally complex and fascinating.

5.3.4. *The universality of conditions.* A question that has arisen at several points is whether some condition is peculiar to English or is universal. I have already treated the OSC in these terms, but there is a bit more to be said. And the monosyllabic non-lexical item exemption deserves some attention as well.

Several people have suggested to me that the OSC be extended from the domain of readjustment rules to some larger domain including other phrase-phonological rules. This proposal makes sense only if some distinction is made like that drawn by Rotenberg, 1978: Part IV, between prosodic [phonological] rules (conditioned ONLY by prosodic organization) and other phonological rules applying between words (rules which are subject to lexical and syntactic conditioning). The problem here is that prosodic rules, in Rotenberg's sense, do not respect syntactic boundaries, even S boundaries.

The English palatalization of alveolar stops (in particular, the shift of *t* to *č*) before *y* will serve as an example of a prosodic rule in Rotenberg's sense. The palatalization occurs freely before a *y* beginning a surface S, as in (85), and it affects a *y* ending a surface S, as in (86).

- (85) a. What [_S you must realize] is that your hair is on fire.
 b. The part [_S you don't understand] is in Dutch.
 c. I can't believe that [_S you like *Gravity's Rainbow*].
- (86) a. The one [you just hit_S] used to be mine.
 b. [What you caught_S] you have to clean.

In each case, the italicized *t* can palatalize under the influence of the immediately following *y*, regardless of the syntactic structure. Clearly, no extension of the OSC is applicable to this palatalization.

Another candidate for universality is the exemption of monosyllabic lexical items from conditions on readjustment and cliticization. The hypothesis of universality here is attractive because this exemption presents a projection problem: how does a child learn that *how*, *not*, *him*, and so on can all support stranded *to*? By example, one by one? Or by natural induction, from a few models (or even from none)? If the latter, then to some extent the exemption is universal.

Moreover (as Ellen Kaisse has pointed out to me), the general framework of components mentioned in section 5.3.2 above would require that the involvement of monosyllabicity be universal rather than language-particular. In that framework, the application of cliticization/readjustment rules can affect the application of phonological rules, but phonological properties cannot condition or constrain cliticization or readjustment. It follows in this framework that monosyllabicity is not a possible conditioning factor for a readjustment rule like 2R. If, however, the involvement of monosyllabicity is given universally, rather than requiring statement in the description of some languages but not others, the essential theoretical claims associated with the framework can be maintained: the English rule 2R would not be stated with a condition on it mentioning number of syllables.

Received 23 February 1982

Revised version received

3 June 1982

Department of Linguistics

Ohio State University

Columbus, Ohio 43210

USA

Notes

- * The earliest version of this paper appeared as Zwicky, 1980, essentially the text of a paper I gave at the LSA annual meeting in Los Angeles in December, 1979. Further versions were presented at Sussex University (June, 1980), Ohio State University (November, 1980), the Washington (D.C.) Linguistics Club (November, 1980), University of Kansas (April, 1981), and Stanford University (November, 1981), CASBS (January, 1982), University of Washington (April, 1982), and University of British Columbia (April, 1982). At this point I cannot possibly thank everyone in these audiences whose suggestions and criticisms have shaped, or simply provided, my ideas; if this were a joint paper it would have several dozen authors. The version immediately preceding this one was distributed to about 50 colleagues, many of whom responded with useful critiques. In addition to those specifically named in the text, Brian Butterworth, Donna Jo Napoli, and an anonymous reader for *Linguistics* deserve thanks, even though I have not always followed their advice. The initial spark was provided by Polly Jacobson at her conference on syntactic representations in the spring of 1979. The final synthesis, such as it is, was made possible by a sabbatical year from

Ohio State and financial support from the Spencer Foundation, which together allowed me to spend a year at the Center for Advanced Study in the Behavioral Sciences.

1. Section 4 of Zwicky, 1978 surveys the types of distributional evidence for syntactic phrases.
2. In the discussion that follows, it is irrelevant whether this construction is derived by a transformational deletion or whether it is an elliptical base structure whose semantic interpretation involves the reconstruction of missing material. See Sag and Hankamer, 1980 for some discussion of these alternatives. Since they favour the latter, they refer to the construction as Verb Phrase Ellipsis, thus avoiding any implication that there is a deletion rule. Another possibility is that *to* (also auxiliaries) without a complement should be analysed as a proform, as Schachter (1978: section 8) argues. I have nonetheless maintained the name Verb Phrase Deletion, following most of the recent literature on the subject.
3. One case might be seen either as a very lengthy parenthetical interruption or as a complete change of thought (and syntactic construction) in mid-sentence, with a later return to the original:

And there was a hotel across the street. A black hotel. And [an Indian girl] was there occasionally and she started to — well for one thing, I really didn't know her right age. I met this black girl ... So I brought the Indian girl to St. Louis and I gave her the job. And after a while she would go around to the hotel to make some extra money. I never knew she was no prostitute, but she was a prostitute.

(Chuck Berry, *RSI*, p. 234)

Here the uncompleted *she started to* is eventually completed by the VP *to make some extra money*.

4. My thanks to Ivan Sag for allowing me to use the students in his introductory linguistics class (autumn, 1981) for this purpose.
5. My thanks to Lincoln Moses for statistical advice and to Lynn Gale for computer services.
6. My sources for the interview materials: Studs Terkel, *Hard Times*, 1970 paperback edition (New York: Pocket Books), 6 examples; *Working*, 1975 paperback edition (New York: Avon), 40 examples; *American Dreams Lost and Found*, 1980 (New York: Pantheon), 11 examples; Tamara Hareven and Randolph Langenbach, *Amoskeag*, 1978 (New York: Pantheon), 12 examples; and *The Rolling Stone Interviews*, 1981 (New York: St. Martin's Press), 62 examples. For fictionalized autobiography: the final chapter of Sara Davidson, *Real Property*, 1981 paperback edition (New York: Pocket Books), 2 examples; and Davidson, *Loose Change*, 1978 paperback edition (New York: Pocket Books), 8 examples. Five further examples were taken from the *OED* entry for *to*, and five from the *WNIS* entry. To these add one example each from the following: Barbara Pym, *A Glass of Blessings*, 1980 paperback edition (Harmondsworth, Middlesex: Penguin); Philip Lopate, *Bachelorhood*, 1981 (Boston: Little, Brown); Elizabeth Perenyi, *Green Thoughts*, 1981 (New York: Random House); a UPI story in the 28 April 1981 *Ohio State Lantern*; the instructions in the 1980 U.S. Income Tax return; and a Chinese fortune cookie supplied by Gregory Stump. My thanks to Margaret Thurston, Jacques Transue, and Elizabeth Zwicky for searching the Terkel books for me.
7. From this point on, a % prefixed to an example indicates that it is of a type subject to genuine dialect variation.
8. From the point of view of surface structure, this is the most conservative labelling scheme available. However, from the point of view of the operation of transformations

deriving these surface structures, this scheme presupposes an elaborated theory of tree pruning. But I would like my analysis to stand, insofar as possible, whether the constructions at issue are transformationally or base derived, and consequently I adopt this scheme here.

9. It is instructive at this point to compare the analysis of (39) with my earlier analysis of essentially the same set of facts in Zwicky, 1980. This earlier, quite baroque, analysis, involving reference to 'original' rather than surface structure and to boundary symbols (indeed to the NUMBER of boundary symbols in particular locations), was a consequence of my not taking structure seriously. In the first place, I did not take seriously the structure provided by the syntax, creating as a result the pseudoproblem of when *to* attaches to the left, as in *not to*, and when to the right, as in VPs like *to do it*. In addition, I did not heed the lesson of Pyle, 1972 and Rotenberg, 1978 that phonological rules (and, by extension, readjustment rules creating phonological phrases) should be sensitive to the type and extent of particular structural domains, not to some set of boundary symbols marking the borders of those domains, and as a result I reified the boundaries S_f and S_s and counted how many of them intervene between *to* and constituents flanking it.
10. Bolinger also proposes that (62) results from an identification of the preposition *to* and the infinitive marker *to* during language acquisition, that is, that (62) involves a generalization from the semantics of sentences like *Give it to me* and *Walk to the door*, and especially from the semantics of sentences with stranded prepositional *to*, such as *Where did you drive to?* I am very dubious about the proposition that children generally assume the identity of homophonous morphemes having (vaguely) relatable meanings, given the many acquisitional anecdotes in which children treat as separate lexical items morphemes that are obviously related (Annette Karmiloff-Smith reports, for instance, her daughter's sudden dinner-table insight: 'Mommy, chicken is CHICKEN!').
11. Marilyn Vihman points out to me that the children I am talking about here are probably considerably older than the customary subjects of language acquisition studies. Certainly the generalization cannot be made until after the child has analysed at least some of the chunks *wanna*, *hafta*, *useta*, *gotta*, *gonna*, etc. as consisting of a verbal element plus the infinitive marker *to*.

References

- Aissen, J. (1974). Verb raising. *Linguistic Inquiry* 5, 325–366.
- Baker, C. (1979). Syntactic theory and the projection problem. *Linguistic Inquiry* 10.4, 533–581.
- Bing, J. (1979). Aspects of English prosody. Ph.D. dissertation, University of Massachusetts.
- Chomsky, N. (1980). On binding. *Linguistic Inquiry* 11.1, 1–46.
- (1982). *Lectures on Government and Binding: The Pisa Lectures*. 2nd rev. ed. Dordrecht: Foris.
- Chomsky, N. and Halle, M. (1968). *The Sound Pattern of English*. New York: Harper and Row.
- Cooper, W. and Paccia-Cooper, J. (1980). *Syntax and Speech*. Cambridge, Mass.: Harvard University Press.
- Culicover, P. (1977). Some observations concerning pseudoclefts. *Linguistic Analysis* 3, 347–375.

- Kahane, H. and Beym, R. (1948). Syntactical juncture in colloquial Mexican Spanish. *Language* 24, 388–396.
- Kaisse, E. (1983). The syntax of auxiliary reduction in English. *Language* 59, 1.
- Kuno, S. (1975). Conditions for verb Phrase Deletion. *Foundations of Language* 13, 161–175.
- Langendoen, D. T. (1975). Finite-state parsing of phrase-structure languages and the status of readjustment rules in grammar. *Linguistic Inquiry* 6, 533–554.
- Levin, N. (1980). Main-verb ellipsis in spoken English. (1979 Ph.D. dissertation, Ohio State University.) *Ohio State University Working Papers in Linguistics* 24, 65–165.
- Napoli, D. (1981). Semantic interpretation vs. lexical governance. *Language* 57.4, 841–887.
- Napoli, D. and Nespor, M. (1979). The syntax of word-initial consonant gemination in Italian. *Language* 55.4, 812–841.
- Peters, S. (1972). The projection problem: How is a grammar to be selected? In S. Peters (ed.), *Goals of Linguistic Theory*, 171–188. Englewood Cliffs, N.J.: Prentice-Hall.
- Pulgram, E. (1970). *Syllable, Word, Nexus, Cursus*. The Hague: Mouton.
- Pullum, G. (1981). The category status of infinitival *to*. University of Washington: *Washington Papers in Linguistics* 6, 55–72.
- Pullum, G. and Postal, P. (1979). On an inadequate defense of “trace theory”. *Linguistic Inquiry* 10, 689–706.
- (1982). The contraction debate. *Linguistic Inquiry* 13.1, 122–138.
- Pyle, C. (1972). On eliminating BM’s. *Chicago Linguistic Society* 8, 516–532.
- Rotenberg, J. (1978). The syntax of phonology. Ph.D. dissertation, MIT.
- Sag, I. and Hankamer, J. (1980). Toward a theory of anaphoric processing. *Stanford Working Papers in Semantics* 1, section E.
- Schachter, P. (1978). English proppredicates. *Linguistic Analysis* 4.3, 187–224.
- Selkirk, E. (1980). *The Phrase Phonology of English and French*. (1972 Ph.D. dissertation, MIT.) New York: Garland.
- (1981). On prosodic structure and its relation to syntactic structure. In Thorstein Fretheim (ed.), *Nordic Prosody II: Papers from a Symposium*, 111–140. Trondheim: Tapir.
- Zwicky, A. (1977). *On Clitics*. Bloomington: Indiana University Linguistics Club.
- (1978). Arguing for constituents. *Chicago Linguistic Society* 14, 503–512.
- (1980). Stranded *to*. *Ohio State University Working Papers in Linguistics* 24, 166–173.
- (1982). An expanded view of morphology in the syntax-phonology interface. Preprints for the XIIIth International Congress of Linguists, Tokyo.
- Zwicky, A. and Levin, N. (1980). You don’t have *to*. *Linguistic Inquiry* 11.3, 631–636.