Wanna-contraction and prosodic disambiguation in US and NZ English

Paul Warren (Victoria University of Wellington), Shari Speer (Ohio State University) and Amy Schafer (University of Hawai'i)

Abstract

This paper presents results from a game-based study, conducted in both New Zealand and the United States, demonstrating how naïve speakers use prosodic phrasing to disambiguate syntactically ambiguous sentences. The particular set of results discussed in this paper concerns how speakers differentiate two types of *wh*-questions, as in (i) and (ii).

(i) Which triangle$_i$ do you want _____$_i$ to change the position of the square?

(ii) Which triangle$_i$ do you want to change the position of _____$_i$ this time?

Our results show that *wanna*-contraction was more frequent for Midwestern American English speakers than for New Zealand English speakers, and more likely in both data sets when there was no syntactic gap between *want* and *to*, as in (ii). Phonetic analyses indicated that both groups of speakers consistently lengthened the word preceding the relevant gap location and the following silence in both (i) and (ii). ToBI transcriptions showed prosodic constituent breaks most often patterning with the location of major syntactic breaks. While our results need not imply that syntactic gaps are directly pronounced in the phonetic structure of an utterance, it seems likely that prosody-syntax correspondence rules may result in phonological regularities in naïve speakers' productions of gap sentences, and that these regularities are available to help comprehenders locate gaps in *wh*-questions.

Wh-gaps and prosodic marking

The ambiguous question in (1) has two readings, indicated by (a) and (b). A commonly accepted syntactic analysis of the ambiguity (e.g. Chomsky, 1977) is that the different readings correspond to movement operations of *wh*-constituents from different positions in the underlying form of the sentence. Thus in the (a) reading, a fronting operation moves the *wh*-word from an underlying subject position before *to leave*, while in the (b) reading it is moved from an object position after *leave*.

(1) Who does John want to leave?

(a) John wants *who* to leave?

(b) John wants to leave *who*?

It has also been argued that the movement operation leaves a "trace" (or wh-gap) at the position of the *wh*-word in the underlying structure, which will block the operation of phonological rules that require adjacency of the words either side of the trace. One such phonological rule might result in the contraction of *want to* to *wanna.* Note however that fast speech processes such

as wanna-contraction are optional (Fodor, 1979), and cannot therefore provide determinate cues to a gap between *want* and *to*. That is, the absence of contraction need not mean that there is a gap. Nevertheless, it has been widely argued that the presence of contracted forms in utterances of (1) would indicate that *want* and *to* are in the same constituent, with no intervening trace, thus disambiguating the utterance towards the (b) reading indicated above (Baker & Brame, 1972; Lakoff, 1970). However, it also appears that the presence of contraction does not necessarily signal a gap-less interpretation. For instance, Karins and Nagy (1993) tested listeners' interpretation of wanna-contraction. They presented a short story followed be a question with or without wanna-contraction (*Who do you {want to / wanna} help?*). In the context of the story, the wh-consituent in the question was compatible with both subject and object roles, though the materials were designed to show a preference for the subject role. The experiment showed that listeners were statistically more likely to select the subject interpretation (*you want X to help*) in the absence of contraction and the object interpretation (*you want to help X*) if there had been contraction, showing that the presence or absence of contraction does have a strong cueing function. However, the contracted form was given a subject interpretation in some 32% of cases, demonstrating that the relationship between the syntactic and phonetic forms is not categorical.

Pullum (1997) argues for a non-syntactic explanation for the incidence of various types of contraction, including wanna-contraction, and posits that forms such as *wanna*, *gonna*, *hafta*, etc. are separate lexemes synchronically related to *want to*, *going to*, *have to* etc. Pullum suggests that there are 'liberal dialects' that have a wider distribution of contracted forms than other dialects, including contexts in which more traditional analyses have argued that contraction should not be found. He argues that contraction in these dialects is constrained by phonological phrasing rather than by syntactic factors. It is possible therefore that the non-categorical responses reported by Karins and Nagy reflect such dialect (or idiolectal) differences between individual listeners. This is an explanation offered by Bley-Vroman and Kweon (2002) for over-general use of *wanna* by the native speakers tested in their study of learner knowledge of wanna-contraction.

In addition to contraction, a gapping site may also be marked by further phonetic properties of the utterance, such as may be expected in the presence of some (perhaps low-level) phonological or prosodic boundary between *want* and *to*. This was indeed found in an early study by Danly (1980), in which participants read ambiguities similar to that in (1), following disambiguating paragraphs intended to make the meaning clear. For non-reduced tokens in Danly's data set, the sequence *want to* was some 25msec longer in the (a) reading, where the syntactic analysis argues for a trace between these two words.

Recent research has disputed the conditions under which wh-gaps might be marked prosodically. A study by Nagel et al (1994) analysed words in the position of *call* in sentences like (2), where the gap locations in each sentence are indicated by Δ. Like Danly, they found longer durations of these key words in the (a) versions. They also found greater pitch movements over these words.

(2a) Which doctor did the supervisor call Δ to get help for his youngest daughter?

(2b) Which doctor did the supervisor call to get help for Δ during the crisis?

Subsequent studies by Straub et al (2001) argued that the prosodic differences reported by Nagel et al for sentence pairs such as (2) derive not from gapping, but from other factors that are known to affect intonational phrasing. They pointed out that the infinitival clause (*to get help for his youngest daughter*) in the (a) version is an adjunct to the verb *call*, whereas that in (b) (*to get help for*) is a complement to *call*. This difference can be demonstrated by replacing the infinitival clause with a purposive clause (*so that he could …*), which is acceptable for (2a) but not for (2b), as shown in (3a) and (3b).

(3a) Which doctor did the supervisor call Δ so that he could get help for his youngest daughter?

(3b) *Which doctor did the supervisor call so that he could get help for Δ during the crisis?

Straub and colleagues pointed out that the critical potential gap location (*call / to*) is at the boundary between a main clause and an adjunct phrase in (2a) but within a phrase in (2b). They argued that the different prosodic phrasings observed by Nagel et al reflect this difference in argument structure rather than gapping *per se*, and that the prosodic consequences of this difference can be accounted for in terms of Selkirk's (1984) *Sense Unit Condition on Intonational Phrasing*. This condition states that the immediate constituents (C_i, C_j) of an intonational unit should form a sense unit, with C_i either modifying or being an argument of C_j (a head). This means that the intonational groupings in (4) are allowed, where an intonational break also coincides with the gap location in each case, but those in (5) are ruled out.

(4a) (Which doctor did the supervisor call Δ) (to get help for his youngest daughter)?

(4b) (Which doctor did the supervisor call to get help for Δ) (during the crisis)?

(5a) *(Which doctor did the supervisor call Δ to get help for) (his youngest daughter)?

(5b) *(Which doctor did the supervisor call) (to get help for Δ during the crisis)?

To confirm their re-interpretation of Nagel et al's experiment, i.e. that the finding of a prosodic effect at a gap location is in fact a prosodic effect due to clause structure differences, Straub et al ran a production experiment of their own using materials such as those in (6):

(6a) *Phrase-Final Gap*:

What did you return Δ to make sure you would get a full refund?

(6b) *Phrase-Medial Gap*:

What did you return Δ to the store when you didn't expect to get a full refund?

(6c) *Phrase-Medial Control*:

Who did you return to the store with Δ when you wanted to get a full refund?

Straub et al argued that if prosodic contrasts reflect differences in gap location, then the critical region (*return to*) should show a difference between (6a) and (6b) on the one hand and (6c) on the other. However, if the prosodic contrast depends primarily on clause structure, then (6a) should differ at this region from both (6b) and (6c), since only in (6a) is there a clause boundary at *return*. Straub et al's production data support this second prediction – both the duration of *return* with its following pause and the pitch excursion across the critical region were significantly greater in (6a) versions than in (6b) and (6c), which did not differ significantly from one another (though there was a non-significant difference in the direction predicted by the gapping account).

While Straub et al's study points out the dangers of confounding gap location and structural boundary location, the earlier work of Danly (using examples with the same structure as (1) above) and related studies of ellipsis sites by Cooper and Paccia-Cooper (1980) showed that clause-internal gaps are likely to be marked by lengthening and pausing relative to comparable non-gap locations. The remaining sections of this paper present some new data relating to this issue from our SPOT speech production task.

Gaps and prosody in the SPOT project

The data to be reported in this paper come from the SPOT project, and involve the gap sentences illustrated in (7) below. Our study of differs from those discussed above in two ways: first, both of the gap sentences used in our game task have complement *to* clauses (see further below); second, our analysis involves a combination of measures: durations, transcriptions, and the incidence of *wanna*-contraction across the potential gap site.

(7a) Which triangle$_i$ do you want Δ_i to change the position of the square?

(7b) Which triangle$_i$ do you want to change the position of Δ_i this time?

The SPOT project is a collaborative research programme involving researchers in New Zealand and the United States of America. The programme goals include the collection of speech production data under more naturalistic conditions than those typically used in speech production studies of prosodic disambiguation. For instance, the gap studies referred to above use reading tasks, yet readers and talkers have different pragmatic goals, and different processing demands – unsurprisingly a number of researchers have commented on the resulting differences in the prosody of read and spontaneous speech (Ayers, 1994; Blaauw, 1994; Howell & Kadi-Hanifi, 1991). Ideally, prosodic distinctions between minimally contrasting syntactic structures would be studied using spontaneous speech. This is however hardly practicable, since it is unlikely that speakers will produce these minimally contrasting utterances without prompting. Some degree of spontaneity might be achieved using role play situations or map tasks (Anderson & colleagues, 1991). Here, though, there are also typically few constraints placed on the syntactic forms of expression. Other techniques, such as descriptions of node networks (Levelt & Cutler, 1983) or tangram shapes (Clark & Schober, 1992; Clark & Wilkes-Gibbs, 1986), have revealed much about the use of prosody in the planning and repair of utterances and in the construction of collaborative discourse, but these tasks are again not designed for eliciting specific syntactic contrasts.

In our own attempts at resolving or at least side-stepping this "laboratory speech" issue, we have constructed a task in which pairs of speakers negotiate the movement of objects around a game board. The tasks requires participants to use a fixed set of sentence frames and object names to construct instructions, requests and acknowledgements, thus producing multiple instances of a set of syntactic structures of relevance to our research. The two participants have slightly differing boards, and different roles. The "Driver" knows the ultimate goals for the objects, and issues instructions accordingly. The "Slider" moves the objects in response to these instructions, and knows the locations of bonuses and dangers, but not the goals for the objects. The Slider asks questions about which object to move and provides confirmation that the object has been moved. Both Driver and Slider are aware of basic game rules governing the types of movement for game board objects. We have found that participants quickly become familiar with the utterances available for use under the rules of the game, and learn to produce them fluently and without recourse to printed sentence lists. We believe that this game task produces a rich source of data for the study of syntactic ambiguity resolution in a more spontaneous speech situation than that used in traditional sentence reading tasks. Since the utterances include a number of syntactic ambiguities frequently studied in comprehension studies, we can also use our recordings to examine the use of prosodic information in ambiguity resolution during comprehension. What is more, by careful planning of our gameboard layouts, we can construct game situations in

which potentially ambiguous utterances are contextually disambiguated and other situations where there is a contextual bias towards one of the meanings.

As well as the gap structures in (7), the syntactic contrasts that we have investigated include early (8a) vs. late (8b) closure of constituents (Schafer, Speer, Warren, & White, 2000) and high vs. low attachment of prepositional phrases (PPs; e.g. in (9) the triangle could be an instrument used to move the square or it could be part of a description of a house-like piece, the "square with the triangle") (Schafer, Speer, & Warren, forthcoming).

(8a) When that moves the square will...

(8b) When that moves the square it...

(9) I want to change the position of the square with the triangle

We claimed above that both utterances (7a) and (7b) involve infinitive clauses as complements. Recall that Straub et al (2001) argued that Nagel et al's (1994) phonetic differences were a reflection not of a gapping difference but of the contrast between adjunct and complement roles of the infinitive clause in (2), as revealed by the purposive test in (3). While it would seem that our sentence in (7a) would allow a purposive, as in (10a) below (and contrasting with (10b)), it is clear from the context of our game task that a different sense of *want* is involved in our materials, one which is compatible with a complement interpretation of the infinitive clause rather than with an adjunct interpretation. This different meaning might be paraphrased as "I wish to change the position of the square with the triangle", which does not allow the purposive (as shown by the awkwardness of a version of (10a) with "wish" in place of "want").[1]

(10a) Which triangle do you want in order to change the position of the square?

(10b) *Which triangle do you want in order to change the position of this time?

---

[1] Note that the Slider sentence in (7a) is produced in response to the instrumental meaning of the Driver sentence illustrated in (9) (i.e. "I want to use the triangle to move the square" - the rules of the game are such that the square cannot move on its own, but must be pushed by another object). The situation of the game prior to the Driver's utterance of (9) is that both game boards (the Driver's and the Slider's) have a triangle on each of two sides of the square, and either of these triangles may legitimately move the square. The Driver's utterance of (9) with the instrumental meaning thus presupposes that there is at least one triangle capable of moving the square, but the utterance remains unclear as to which triangle should do the moving, which is why the Slider produces the question in (7a). Notice though that the purposive in (10a) means something like "of the available triangles, which one do you want to have so that you can use it to move the square?", which is distinct from the meaning intended in our game task.

Data

Our project data include recordings of eight pairs of speakers of each of Midwestern American English (MAE) and New Zealand English (NZE). Each pair completed as many as six games (not including practice games) during a session of approximately two hours, swapping Driver and Slider roles between games. Data for one speaker have been removed for each dialect, in both cases because of a high level of disfluency. The remaining data have been analysed both acoustically (duration, pausing, F0 patterns) and auditorily (using Pierrehumbert's Autosegmental Metrical framework encapsulated in the ToBI transcription system (Beckman & Ayers, 1997; Pierrehumbert, 1980)).

Across the different sentence types of the SPOT project, and for both MAE and NZE speakers, our phonetic measures and transcription data have shown that naïve speakers produce patterns of prosodic phrasing that indicate the syntactic constituent boundaries they intend to communicate. This paper presents detail of our analysis of the gap structures in (7). Utterances containing disfluencies (hesitation pauses, repeats, false starts, mispronunciations, etc.) were excluded from the analysis (this accounted for approximately 9% of the total gap structures available). The remaining data for this analysis include, for MAE, 61 early gap utterances (as in (7a), repeated below) and 193 late gap utterances (i.e. where the gap is after *of*, as in (7b)). For NZE there are 41 early and 194 late gap utterances. Both of the early and late gap utterances are Slider utterances, i.e. they are requests for more information from the participant in the role of Slider. (7a) and (7b) respectively are uttered by Sliders in response to the instructions from the Driver given as (11a) and (11b) below. (The intended meaning of the ambiguous utterance in (11a) was tracked by an experimenter, who noted the object being moved by the Driver.) As stated earlier, the rules of the game are such that the square piece cannot be moved on its own, but must be pushed by another piece. Often there is a choice of pieces that can be used to push the square, including a choice between two triangles (of differing colours), and this will lead to the Slider utterance in (7a). In some gameboard configurations, the objects that can push the square might be shapes other than triangles (e.g. cylinders), so the total count of early gap utterances includes a relatively small number in which there would be a different object name in the position of *triangle* in (7a). The total numbers of fluent triangle utterances of (7a) are 53 for MAE and 38 for NZE. The instruction in (7b) refers to the movement of objects by themselves (rather pushing a square), and can similarly relate to the movement of shapes other than triangles. The counts of fluent triangle utterances of (7b) are 155 for MAE and 151 for NZE.

(7a) Which triangle$_i$ do you want Δ_i to change the position of the square?

(7b) Which triangle$_i$ do you want to change the position of Δ_i this time?

(11a) I want to change the position of the square with the triangle. [with the intended meaning of "use the triangle to push the square"]

(11b) I want to change the position of the triangle.

The following sections present summaries of our results for these materials, first for the incidence of wanna-contractions and then for phonetic (duration) and finally phonological prosodic analyses. Each section presents both the MAE and NZE data.

Wanna contraction

Earlier research referred to above (Baker & Brame, 1972; Lakoff, 1970) claimed that the presence of a syntactic gap is likely to block the contraction of *want to* to *wanna*, but because wanna-contraction is optional (Fodor, 1979), the absence of contraction need not indicate the presence of a gap. Likewise, we noted earlier that both perceptual and more theoretical studies (Karins & Nagy, 1993; Pullum, 1997) have suggested that the presence of a contracted form may not unambiguously indicate that there is no intervening gap. However, it is clear that much of the literature in theoretical syntax that has used contracted forms such as *wanna* as evidence has espoused a strong belief that contraction will be more frequent in no-gap contexts. In line with this, and with the overall preference in Karins and Nagy's (1993) perceptual study, we predicted a greater incidence of contraction in the late gap versions (7b) of our temporary ambiguity than in the early gap versions (7a).

The contraction literature has been centred largely on US English, and research on the phenomenon in New Zealand English is virtually non-existent. Researchers at the University of Canterbury in Christchurch (Quinn, p.c., 2004) report that contractions (including wanna-contraction) are more frequent in the speech of the younger speakers in their corpora than in that of the older speakers, but that this finding may be confounded by a difference in the degree of formality of the recordings across age groups. But even in the youngest group (i.e. the most progressive group in terms of contraction) the uncontracted form is far more common than the contracted. Although there is no empirical data to support a difference between the varieties, general informal comments from informants in New Zealand suggests that wanna-contraction is more typically associated with American than with New Zealand varieties of English. We therefore also predicted that wanna-contraction would be more frequent in the MAE recordings than in the NZE recordings of our SPOT task.

Figure 1 gives the percentage incidence of wanna-contraction (i.e. contraction at the potential early gap site) in the early (7a) and late (7b) gap conditions of our experiment, separately for MAE and NZE. The figure shows quite clearly that contraction was indeed more frequent in MAE than in NZE, and that in both varieties contractions are more likely when there is no syntactic gap between *want* and *to* (i.e. in the late gap condition). The

counts from which Figure 1 was derived were analysed in the following manner. Counts of tokens with and without contraction were entered into a hierarchical loglinear analysis (Hays, 1981: 561-6), with Variety (MAE vs. NZE), Type (early gap vs. late gap) and Contraction (present vs. absent) as factors. A parsimonious model for the data was built by backwards elimination of factor combinations. The tests of this model that are of interest are the interactions of Contraction with the other factors, since these interactions indicate an effect of these other factors on the incidence of contraction. The analysis shows that both of these other factors are significant (as measured by the likelihood ratio χ^2 changes, which for the interaction of Contraction with Variety was 11.89, $p<0.001$, and for the interaction of Contraction with Type was 18.88, $p<0.001$). There was no significant interaction involving both Variety and Type.

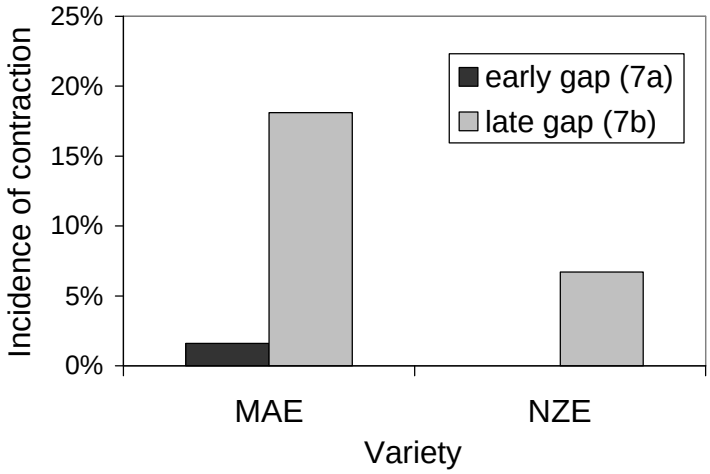

Figure 1. Incidence of wanna-contraction in early and late gap sentences (see examples (7a) and (7b) respectively) in Midwestern American English and New Zealand English. Data are from 15 speakers of each variety.

Despite the fact that contraction is – as predicted – much more likely when there is no gap, the low overall level of contraction indicates that it is unlikely to be a reliable indicator of the difference between (7a) and (7b).

Wanna-contraction is 2.5 times more likely in the MAE data (13.8% overall) than in the NZE data (5.5% overall). This would seem to reflect the informal observations that such contractions are more widespread in colloquial American English than in colloquial New Zealand English.[2] Interestingly though, the incidence of contraction is due in each case to high levels of *wanna* from a few speakers (one NZE speaker and three MAE

[2] This pattern of greater wanna incidence in MAE than in NZE is confirmed in an analysis of the Driver utterances exemplified in (11) that precede utterances such as (7). For these, there was an overall contraction rate of just under 5% for MAE, but less than 1% for NZE.

speakers have incidence levels of 35% or more, and these account for 85% or more of wanna-contraction in the data for each variety), suggesting that the use of *wanna* is something of an individual characteristic, used by Pullum's (1997) "liberal" dialect speakers, while still more likely in one variety than the other. Note however that Pullum suggests that such liberal dialect speakers are likely to contract across the board, and not just when there is no intervening syntactic gap. But this general pattern of wanna-contraction does not seem to be the case here, as shown by the difference between the early and late gap conditions. Pullum argues that contraction is constrained for liberal dialect speakers by phonological factors, such as whether the relevant words are in the same intonational phrase. For this reason, we will return to discussion of wanna-contraction after our discussion of other phonological distinctions between utterances (7a) and (7b).

Durational analyses

Recall that the research discussed above was equivocal as to whether durational differences would provide a reliable indicator of the presence of a syntactic gap. On the one hand, Danly (1980) and subsequently Nagel et al (1994) found significant durational differences. On the other, Straub et al (2001) argued that Nagel et al (1994) confused gapping and other syntactic differences; Straub et al went on to present data that showed non-significant differences. We predicted for our data that we will find similar differences to those reported by Danly for similar materials, and that these differences would obtain despite the lack of overt instruction to our participants to disambiguate the structures they produced.

We measured key durations at both the early gap site and the late gap site. At the early gap we measured the duration of *want* and *to*, and any pause between these words. In the cases of contraction, we took a single measure of the duration of *wanna*. For the following presentation we consider the duration of the whole sequence *want+pause+to*, which is the measure used by Danly (1980). At the late gap site we measured the preposition *of* and any following silence. Again, the presentation below gives the values of the entire *of+pause* sequence. For each variety, speakers' data were included only if they had two or more fluent tokens of each of the early and late gap utterances (7a) and (7b). This means that the data analysed below comes from 13 MAE speakers and 11 NZE speakers.

Durations at the early gap site

Two analyses were carried out on the data from the early gap site. One included all data from the 24 speakers in the two varieties, while the other excluded contracted (*wanna*) tokens, reflecting Danly's (1980) analysis. Figure 2 displays the results for the larger dataset.

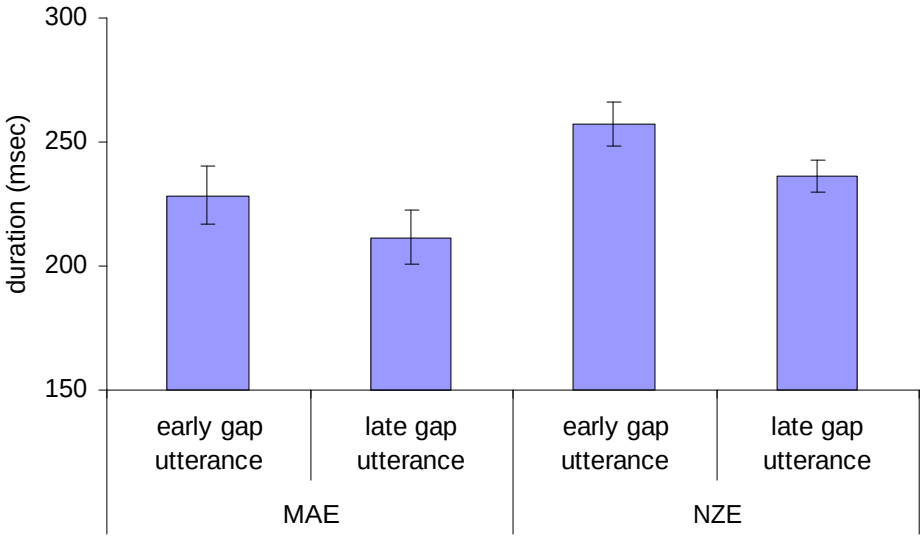

Figure 2. Duration of wanna sequence (want+pause+to, or contracted wanna) in early and late gap utterances (see examples (7a) and (7b) respectively) in Midwestern American English and New Zealand English. Data are from 13 speakers of MAE and 11 speakers of NZE.

Analysis of Variance was conducted on the subject mean values, with utterance type (early gap vs. late gap) and variety (MAE vs. NZE) as factors. The ANOVA showed a significant main effect of utterance type ($F[1,22]=12.92$, $p<0.01$) and a strong near-significant main effect of variety ($F[1,22]=4.087$, $p=0.056$). The averages of the subject means showed a 20 msec difference in the duration of the wanna sequence in early and late gap utterances (242 vs 222 msec). The difference between the means of the complete datasets (i.e. not first averaged for each subject) was slightly larger at 27 msec (246 vs 219 msec). The differences were slightly smaller when contracted *wannas* were excluded, but the difference between early and late gap utterances remained highly significant. The difference between the varieties is consistent for both early and late gap utterances (there was no significant interaction of the two factors in the analysis). While this may indicate a general difference in speech rate between the two groups, it may also reflect differences in phrasing (see Discussion below).

Durations at the late gap site

The averages of the subject mean durations of the sequence *of+pause* are shown in Figure 3. Analysis of Variance revealed a significant main effect of utterance type ($F[1,22]=25.497$, $p<0.001$), and no other effects. The difference between the two utterance types was 79 msec (123 msec vs. 202 msec), indicating that this late gap position is clearly marked by durational differences. Note that there is a tendency for the NZE data to show shorter durations here, in contrast to a tendency in the opposite direction at the early gap position (compare Figure 2). This may also reflect differences in phrasing between the two varieties (again, see Discussion below).

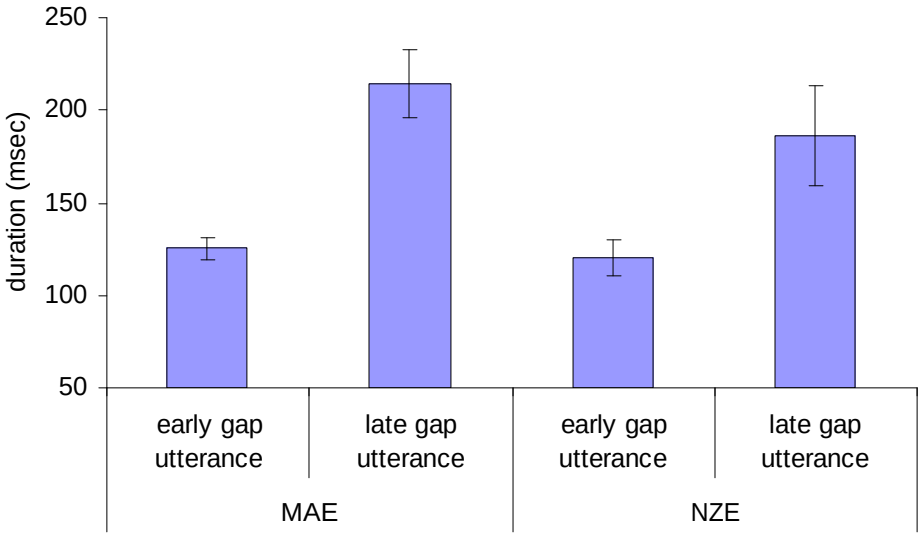

Figure 3. Duration of the sequence of+pause in early and late gap utterances (see examples (7a) and (7b) respectively) in Midwestern American English and New Zealand English. Data are from 13 speakers of MAE and 11 speakers of NZE.

Overall, the differences in duration at the two sites are similar for the two varieties – there was no suggestion of an interaction of variety and utterance type at either site. The difference at the early gap site is similar in magnitude to that reported by Danly (1980) (25msec). The difference at the later gap site is clearly larger, as might be expected given the difference between the structures (in (7a) the preposition is followed by its complement, in (7b) by a new phrase, serving as an adjunct to the verb phrase).

The data in Figures 2 and 3 focus on the gap sites in utterances (7a) and (7b). Further measurements show that there are also differences at other sites in the utterances, particularly at *change* and *position*. These two sites are summarised in Figure 4. The clearest differences are for NZE, though Analysis of Variance shows an interaction of variety and utterance type only for the *change*+pause data ($F[1,22]=5.156$, $p<0.05$). In analyses for both *change*+pause and *position*+pause there was a significant main effect of utterance type (with longer durations in early gap than in late gap utterances in each case: *change*: $F[1,22]=11.236$, $p<0.01$; *position*: $F[1,22]=5.255$, $p<0.05$).

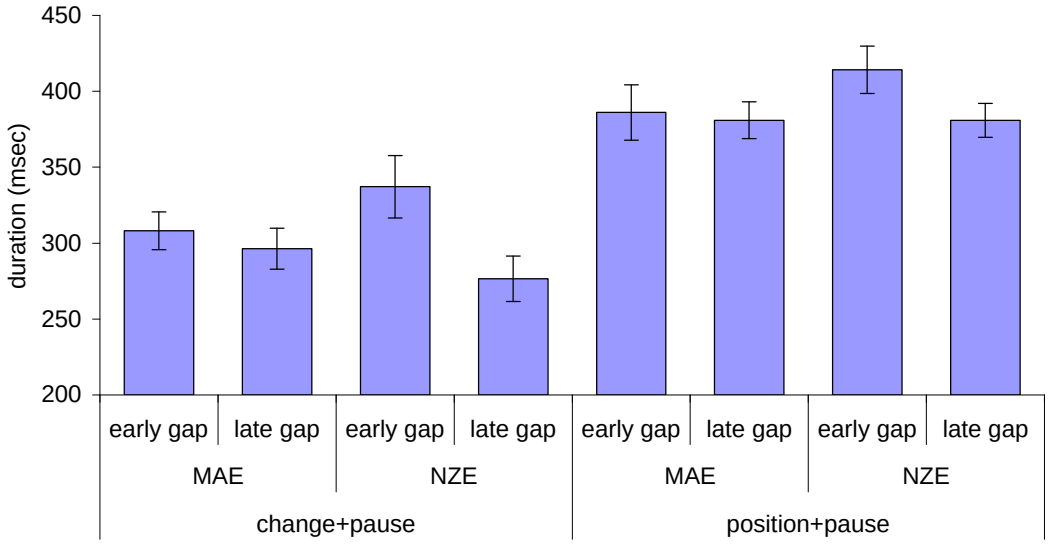

Figure 4. Duration of the sequences change+pause and position+pause in early and late gap utterances (see examples (7a) and (7b) respectively) in Midwestern American English and New Zealand English. Data are from 13 speakers of MAE and 11 speakers of NZE.

At issue here is whether this pattern of differences at *want*, *change* and *position* is indicative of a general lower rate of speech in early gap vs. late gap utterances, or whether it reflects a tendency to place intonational phrase breaks at these positions in the early gap utterances. Note that this need not mean that speakers break at each of these positions in any one utterance, but that there are breaks at these three locations in various combinations across the utterance set as a whole. An explanation in terms of speech-rate could in fact argue for some general speeding-up through "want to change the position of" in late gap utterances, while an explanation based on boundary locations may relate to how a constraint such as the Sense Unit Condition may operate in the case of early and late gap utterances. The two explanations are of course not mutually exclusive. The phonological analysis discussed in the next section is relevant both to this issue and to the question of whether possibilities for wanna-contraction reflect intonational phrasing.

Phonological analysis

The phonological analysis of our wanna-materials in (7) consists of ToBI-style transcriptions of the utterances. These transcriptions are available for 159 of the fluent MAE recordings (49 early gap utterances and 110 late gap utterances), and for a subset of 40 of the NZE recordings (a random selection of 20 early and 20 late gap utterances). Each transcription was carried out by at least one expert ToBI transcriber. Utterance files were cut after "of" so that transcribers did not know from the lexical content whether the early or late gap utterance had been produced. Transcribers marked pitch accents (H*, L* or bitonal e.g. L+H*), phrase accents (at the ends of intermediate phrases, i.e. L- or H-) and boundary tones (at the ends of full intonational phrases, i.e. L%

or H%), and indicated break index values at each word boundary. (Possible break index values are 0 for where words are run together, as in *wanna*, 1 for phrase-internal word boundaries, 3 for intermediate phrase boundaries and 4 for intonational phrase boundaries. 2 is used to indicate an ambiguous boundary that carries perhaps tonal but not rhythmic phrase boundary markings.)

The following presentation focuses on the locations of phrase boundaries in the utterances, i.e. corresponding to break index values of 3 or 4 and the tonal markings of intonational and/or intermediate phrase boundaries. Panels (a) through (d) in Figure 5 show, for each word location from *which* through to *of,* the proportion of utterances that contain such a phrase boundary at that location, for MAE early and late gap, and NZE early and late gap utterances.

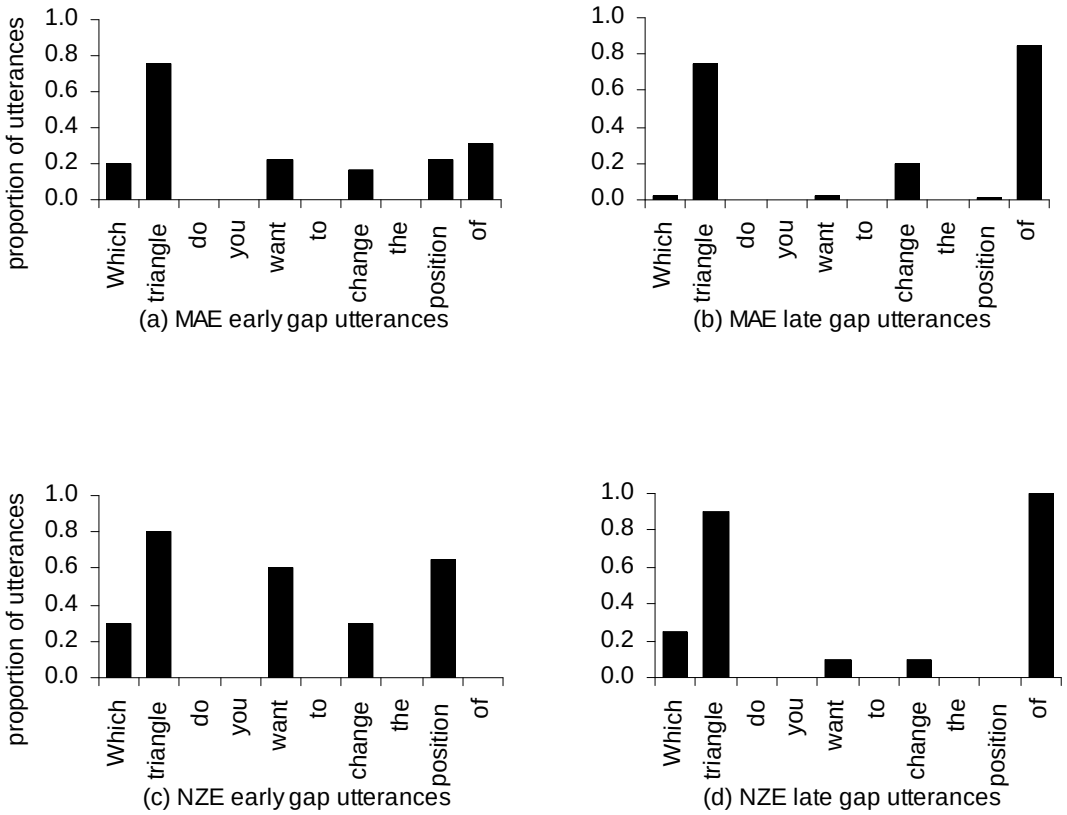

Figure 5. Incidence of phrase breaks (intonational and/or intermediate) across the sequence Which triangle do you want to change the position of, in early and late gap utterances (see examples (7a) and (7b) respectively) in MAE (panels (a) and (b)) and NZE (panels (c) and (d)).

The counts from which Figure 5 was derived were analysed in the following manner. For each location where there were sufficient data (i.e. excluding *do*, *you*, *to*, *the*) counts of tokens with and without breaks were entered into a hierarchical loglinear analysis, with Variety (MAE vs. NZE), Type (early gap vs. late gap) and Break (present vs absent) as factors. For each analysis a

parsimonious model was built by backwards elimination of factor combinations. The tests of this model that are of interest are the interactions of Break with the other factors, since these interactions indicate an effect of these other factors on the incidence of intonational breaks. Table 1 shows the significant interactions for each location. The differences between breaks reported for MAE and NZE occur at *which*, *want* and *position*, with, in each case, more breaks in NZE. At *which* and *of* there was a main effect of Type, with more breaks at *which* in the early gap utterance, and more at *of* in the late gap utterance. The interactions of Type with Variety at these locations simply reflect differences in the size of these effects – the Type difference at *which* is greater for MAE, the Type difference at *of* is greater for NZE. Two further main effects of Type, at *want* and *position*, reflect the greater likelihood of breaks at these locations in the early gap utterances.

	Variety	Type	Variety*Type
which	**	**	*
triangle			
do			
you			
want	**	**	
to			
change			
the			
position	**	**	
of		**	**

*Table 1. Results of statistical analysis of the incidences of phrase breaks shown in Figure 5. For each factor (combination) shown by the column headings, significant interactions with the Break factor at each word location are indicated by ** (for $p<0.01$) or * (for $p<0.05$), blanks indicate that no significant difference was found, shaded cells show comparisons that were not possible because of small datasets. (See text for further details of the analysis.)*

Thus, while there are some differences between MAE and NZE in the level of incidence of breaks, both varieties show similar differences between the early and late gap utterances in how these are phrased. The analyses and the pattern of breaks shown in Figure 5 indicate that late gap utterances are likely to be marked by two main breaks, after *triangle* and *of*. Further, while early gap utterances are just as likely as late gap utterances to have a break after *triangle*, they are more likely to also have breaks at *want* and *position*, and are unlikely to have one at *of*.

Discussion

The above analyses have shown that patterns of prosodic marking of gap constructions demonstrated in a range of previous studies are found also in data from the SPOT project. Phonological (transcription) and phonetic (duration) measures reflect a greater tendency to break at *want* in the early gap utterance (7a) than in the late gap utterance (7b). In addition, contraction

of *want to* to *wanna* was more likely in the late gap utterances, i.e. when there is no gap site at *want*.

The differences between the two varieties at the early gap site (i.e. at *want to*) are consistent. NZE has fewer contractions (see Figure 1), which is in agreement with general informal observations on the two varieties. At the same time, however, NZE has greater durations across the *want+pause+to* sequence (Figure 2). It is unlikely that this difference is a simple result of slower speech in the NZE because the opposite pattern (shorter durations in NZE) is found at the late gap site (Figure 3). It is also the case that these early gap sites are more frequently marked by prosodic breaks in NZE, both in early gap utterances and in late gap utterances (Figure 5 and Table 1). There may therefore be a general difference between the two varieties in the extent to which these utterances are broken into intonational and intermediate phrases. The greater incidence of such phrasing at the early gap site in NZE would then correlate with the lower incidence of wanna-cotnraction.

We also found that the prosodic phrasing of the utterances is highly variable in both varieties. The distribution of breaks at different locations in Figure 5 shows little evidence that any particular word boundary is consistently marked by an intermediate or intonational phrase break, with the exception of *triangle* in both utterance types and *of* in the late gap utterances. The variability of phrasing is revealed in the counts of different phrasings for the transcribed materials. For instance, from the 20 transcriptions completed for each of the utterance types in the NZE data we obtained 10 different phrasings (i.e. different organisations into intermediate and/or intonational phrases) for early gap and 7 different phrasings for late gap utterances. From the MAE transcriptions, we get 20 distinct phrasings for 49 early gap utterances, and 12 phrasings for the 110 late gap utterances. Despite this degree of variability, there are some areas of consistency. The break after *triangle*, for instance, is found overall in approximately 77% of cases (across both varieties). This break appears to place the phrase *which triangle* into focus as the element for which further information is being sought. This is appropriate to the context of the game, in which a Slider utterance from (7) follows one of the Driver utterances in (11).

What is of interest is whether the subsequent phrasing of the Slider utterances is constrained in any way, in particular in order to indicate the possible location of a gap to which the filler *which triangle* is linked. Constraints from such phrasing might operate over and above any explicit marking of a gap site through the prosodic and other phonological means investigated in our analyses of the gap locations. Research into the analysis of filler-gap dependencies has considered a range of factors that might affect the linking of fillers with potential gap sites (Clifton & Frazier, 1990). In one study of prosodic constraints, Broderick (1996) reports that listeners prefer to

link a filler to a gap within the same intonational phrase. As Straub et al suggest, in connection with Broderick's work, "it is conceivable that the comprehension and the production of filler-gap dependencies interact with the comprehension and production of prosodic structure, even in the absence of prosodic cues that exist solely to mark the location of syntactic gaps" (Straub et al., 2001: 392). If speakers are sensitive to the apparent listener preferences indicated by Broderick (1996), then this would suggest that the structure for the utterances would be as in (12), where (/) indicates (possible) break locations. Linking of the filler *which triangle* to the gap after *want* is possible with any of the phrasings in (12a) (as well of course as with the phrasing in (12b)), since the filler will attach in each case to a gap within the phrase started at or before *do*. On the other hand, the presence of such a phrase break before *of* would make the late gap interpretation less likely, since the gap location would no longer be in the phrase started with or immediately after the filler.

(12a) Which triangle (/) do you want (/) to change (/) the position (/) of the square?

(12b) Which triangle (/) do you want to change the position of (/) this time?

In the transcribed data from our NZE speakers, full intonational phrase breaks are found after *triangle* in 6 (of 20) early gap utterances, and in 3 (of 20) late gap utterances. All of the remaining utterances have no internal intonational phrase breaks, except for a single early gap utterance, which has a break after *position* (and is therefore compatible with the postulated break locations in (12a) above). The majority of the transcribed NZE utterances do not therefore distinguish early and late gap structures on the basis of the location of full intonational phrase breaks. If we include intermediate phrase breaks (while noting that Straub et al's discussion of Broderick's work refers to intonational phrase breaks.), then we find that nineteen of the twenty transcribed early gap utterances have a break at either *want*, *change* or *position* (as predicted in (12a)), while only 5 of the late gap utterances have breaks before *of* (3 at *want* and 2 at *change*).

In the transcribed MAE data, the locations of full intonational phrase breaks again do not fully distinguish early and late gap structures – for example, 40 (of 49) early gap utterances have no intonational phrase break between *triangle* and the end of the utterance, and this is true of 87 (of 110) late gap utterances. If we include intermediate phrase breaks, then there are still some that have no break later than *triangle* in both gap types (12 early and 7 late). There are nevertheless some differences in phrasing – some 55% (27) of early gap utterances have a break at one or more of *want*, *change* or *position*, while only 25% (27) of the late gap utterances have breaks at one or more of these positions. Conversely only 20% (10) of the early gap utterances have a break at *of* without an earlier break at *want*, *change* or *position*, while the figure for late gap utterances is 69% (76). These analyses show that there are clear tendencies towards the patterns of phrasing suggested in (12) for early and

late gap utterances, but also some exceptions. Closer inspection of pitch accent and boundary tone patterns over the utterances might reveal the use of tonal cues to structure not exposed by this analysis.

In summary, the data reported in this paper provide support for intonational distinction of the utterances in (7a) and (7b). However, they also reveal considerable variation in intonational realizations and in the incidence of wanna-contraction. While some of this variation may be due to differences between English varieties, it is clear that much more of the variation is a result of individual choice in how to phrase the utterances. Importantly, however, the choices are not unconstrained, and an important constraint is the distinction between early and late gap structures. Further experimentation will determine the significance of the observed distinctions for the comprehension of these utterances, and the immediacy with which prosodic cues are interpreted as indicators of the different structures.

References

Anderson, A. H., & colleagues. (1991). The HCRC map task corpus. *Language and Speech, 34*, 351-366.

Ayers, G. (1994). Discourse functions of pitch range in spontaneous and read speech. *OSU Working Papers in Linguistics, 44*, 1-49.

Baker, C. L., & Brame, M. K. (1972). Global rules: a rejoinder. *Language, 48*, 51-75.

Beckman, M. E., & Ayers, G. (1997). *Guidelines for ToBI labelling*.Unpublished manuscript, Columbus, OH: Ohio State University

Blaauw, E. (1994). The contribution of prosodic boundary markers to the perceptual difference between read and spontaneous speech. *Speech Communication, 14*, 359-375.

Bley-Vroman, R., & Kweon, S.-O. (2002). *Acquisition of the constraints on wanna contraction by advanced second language learners: Universal grammar and imperfect knowledge*.Unpublished manuscript.http://www.sls.hawaii.edu/bley-vroman/ (accessed 9th Feb 2004)

Broderick, K. (1996). The role of prosody in Wh-movement: Intonational phrasing and adjunct construal. Poster presented to the *Ninth Annual CUNY Conference on Human Sentence Processing*, New York, March 21-23, 1996.

Chomsky, N. (1977). On wh-movement. In A. Akmajian, P. Culicover & T. Wasow (Eds.), *Formal Syntax*. New York: Academic Press.

Clark, H. H., & Schober, M. F. (1992). Understanding by addressees and overhearers. In H. H. Clark (Ed.), *Arenas of Language Use* (pp. 176-197). Chicago: University of Chicago Press.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition, 22*, 1-39.

Clifton, C., & Frazier, L. (1990). Comprehending sentences with long-distance dependencies. In G. N. Carlson & M. K. Tanenhaus (Eds.), *Linguistic structure in language processing* (pp. 273-318). Dordrecht: Kluwer.

Cooper, W. E., & Paccia-Cooper, J. (1980). *Syntax and Speech*. Cambridge, MA: Harvard University Press.

Danly, M. (1980). *The wh-movement transformation and speech timing*.Unpublished manuscript

Fodor, J. D. (1979). Superstrategy. In E. Walker & W. Cooper (Eds.), *Sentence Processing*. Hillsdale, NJ: Lawrence Erlbaum.

Hays, W. L. (1981). *Statistics* (3rd ed.). New York: Holt, Rinehart & Winston.

Howell, P., & Kadi-Hanifi, K. (1991). Comparison of prosodic properties between read and spontaneous speech. *Speech Communication, 10*, 163-169.

Karins, K., & Nagy, N. (1993). Developing an experimental basis for determining grammaticality. *The Penn Review of Linguistics, 17*, 93-100.

Lakoff, G. (1970). Global rules. *Language, 46*, 627-639.

Levelt, W. J. M., & Cutler, A. (1983). Prosodic marking in speech repair. *Journal of Semantics, 2*, 205-217.

Nagel, H. N., Shapiro, L. P., & Nawy, R. (1994). Prosody and the processing of filler-gap sentences. *Journal of Psycholinguistic Research, 23*, 473-485.

Pierrehumbert, J. B. (1980). *The Phonology and Phonetics of English Intonation.* Unpublished Doctoral dissertation, MIT.

Pullum, G. K. (1997). The morpholexical nature of English to-contraction. *Language, 73*(1), 79-102.

Quinn, H. (p.c., 2004, p.c.). Modal contractions in the ONZE copora.

Schafer, A. J., Speer, S. R., & Warren, P. (forthcoming). Prosodic influences on the production and comprehension of syntactic ambiguity in a game-based conversation task. In J. C. Trueswell & M. K. Tanenhaus (Eds.), *Approaches to Studying World-Situated Language Use: Bridging the Action and Product Traditions*.

Schafer, A. J., Speer, S. R., Warren, P., & White, S. D. (2000). Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research, 29*, 169-182.

Selkirk, E. O. (1984). *Phonology and syntax: the relation between sound and structure*. Cambridge, MA: MIT Press.

Straub, K., Wilson, C., McCollum, C., & Badecker, W. (2001). Prosodic structure and wh-questions. *Journal of Psycholinguistic Research, 30*, 379-394.