"That Sounds Unlikely": Syntactic Probabilities Affect Pronunciation

Susanne Gahl (gahl@uchicago.edu)
Beckman Institute for Advanced Science and Technology, 405 N. Mathews Ave
Urbana, IL 61801 USA

Susan M. Garnsey (sgarnsey@psych.uiuc.edu)
Department of Psychology, 603 E. Daniel St
Urbana, IL 61801 USA

Cynthia Fisher (clfishe@uiuc.edu)
Department of Psychology, 603 E. Daniel St
Urbana, IL 61801 USA

Laura Matzen (lmatzen@psych.uiuc.edu)
Department of Psychology, 603 E. Daniel St
Urbana, IL 61801 USA

Abstract

The probability of encountering a particular syntactic configuration, given a particular verb (i.e. "verb bias" or "subcategorization preference") affects language comprehension (Trueswell, Tanenhaus, & Kello, 1993). A recent study (Gahl & Garnsey 2004) of sentences with temporary direct object / sentential complement ambiguities shows that such probabilities also affect language production, specifically pronunciation. In this paper, we extend that finding to a new sentence type – sentences with initial subordinate clauses with temporary closure ambiguities, such as *If the tenants beg [, the landlord will let them stay | the landlord, he will let them stay*]. We show systematic differences in the pronunciation of sentences with high vs. low probability structures, given their verbs. We briefly discuss the implications of this finding for research on sentence comprehension and pronunciation variation.

Keywords: Probability; pronunciation; prosody; verb bias; subcategorization; closure ambiguities; early closure; late closure; variation; probabilistic parsing; language production.

Introduction

The probability of encountering a particular syntactic configuration, given a particular verb is often referred to as "verb bias" or "subcategorization preference" and has been shown to affect language comprehension (Trueswell et al., 1993). Verb biases are normally thought of as generalizations over usage. Do such probabilities affect the language production system during the production of individual utterances? A recent study shows that they do: Examining sentences with temporary direct object / sentential complement (DO/SC) ambiguities, such as *The divorce lawyer argued the issue* [*was irrelevant | with her colleague*], Gahl and Garnsey (2004) found systematic differences in the pronunciation of sentences with high vs. low probability structures, given their verbs.

The only other studies, to our knowledge, to check for effects of verb bias on pronunciation (Blodgett, 2004; Kjelgaard & Speer, 1999) did not find such effects. This discrepancy could be due to the sentence types examined in those studies, viz. sentences with initial transitive or intransitive subordinate clauses, such as *When Roger leaves* [*the house is empty | the house, it's empty*]. Such sentences – sometimes referred to as "Closure sentences" – are typically pronounced with a prosodic boundary at the end of the subordinate clause. This boundary tends to be more salient than boundaries in DO/SC sentences. Warren (1985), for example, reports that the duration of verbs at clause boundaries is lengthened, on average, by a factor of 1.08 in DO/SC sentences, but by a factor of 1.51 in Closure sentences. Perhaps the salient boundaries in Closure sentences eclipse microvariation of the sort described in Gahl and Garnsey (2004). On the other hand, the absence of probability-based variation in Kjelgaard and Speer's and Blodgett's materials may be due to the fact that these studies used recordings made by a trained speaker intentionally producing particular boundary types in the ToBI standard (Beckman & Hirschberg, 1994; Beckman, Hirschberg, & Shattuck-Hufnagel, 2005; Silverman et al., 1992). Such deliberate, tightly controlled pronunciation may not show the same range of variation found in naïve productions. Perhaps less controlled pronunciation of Closure sentences does reflect structural probabilities.

Many more verbs participate in transitivity alternations than in the DO/SC alternation (cf. Levin, 1993). Therefore, closure sentences allow us to study the workings of a larger part of the lexicon than do DO/SC sentences, as well as to examine pronunciation variation in strongly marked prosodic boundaries.

How might structural probabilities affect the pronunciation of Closure sentences? Gahl and Garnsey (2004) predicted that phonetic characteristics of boundaries

would tend to be exaggerated for low-probability boundaries, and minimized for high-probability boundaries. This prediction was confirmed. Thus, Gahl and Garnsey found greater degrees of pre-pausal lengthening near low-probability clause boundaries than near high-probability clause boundaries.[1] These predictions were motivated by earlier findings that high lexical frequency and high lexical transitional probability – the probability of a word given a neighboring word - promote phonetic reduction (Bell et al., 2003; Gregory, Raymond, Bell, Fosler-Lussier, & Jurafsky, 1999; Jurafsky, Bell, Gregory, & Raymond, 2001).

Clause boundaries in Closure sentences occur either immediately following the verb (Early Closure, i.e. intransitive: *When Roger leaves, # the house is empty*), or following the direct object (Late Closure, i.e. transitive: *When Roger leaves the house, # it is empty*). In both cases, speakers tend to insert pauses at the boundary and lengthen the words immediately before the boundary (the verb in Early Closure sentences, the direct object in Late Closure sentences), compared to its baseline duration (Kjelgaard & Speer, 1999; Schafer, Speer, Warren, & White, 2000; Warren, 1985; Warren, Grabe, & Nolan, 1995). If structural probabilities have similar effects in these sentences as in DO/SC sentences, we should expect words near low-probability boundaries to be lengthened more than those near high-probability boundaries.

In this paper, we offer evidence showing that this is the case. Our results support the notion that verb biases affect language production, as well as comprehension. Our results have implications for research on sentence processing and probabilistic pronunciation variation.

Method

Participants

Twenty undergraduate students (ten male, ten female) at the University of Illinois participated in the experiment for payment. All were native speakers of English without reported hearing problems.

Materials and design

The materials consisted of Early/Late Closure sentence pairs, such as *When the python escaped the zoo (it) had to be closed to the public,* some of which were based on the sentences in Kjelgaard and Speer (1999). The main consideration in selecting the verbs for the subordinate clause was strong verb bias, either towards transitive (Late Closure) or intransitive (Early Closure). Estimates of verb bias were based on corpus counts (Gahl, Jurafsky, & Roland, 2004). Ten Transitive Bias verbs and ten Intransitive Bias verbs were selected.

The two sets of verbs did not differ significantly in frequency or length in letters, phonemes, or syllables (all $t(18) < 1$, $p > .5$), but did differ in transitivity bias ($t(18) = 18.8$, $p < 001$.). Each of the verbs appeared in two sentences, one with Early Closure and one with Late Closure, e.g. *When the python escaped, the zoo had to be closed to the public* and *When the python escaped the zoo, it had to be closed to the public*. The complete set of sentences appears in Table 1. All participants read all sentences.

To prevent confounds due to presentation order, two presentation lists were constructed, each with two blocks of twenty sentences. On List 1, half the verbs of each bias type appeared in their bias-conforming syntactic context in block 1, and in their bias-violating context in block 2, while for the other half of the verbs, the opposite order was used. On List 2, the relative order of bias-conforming and bias-violating environments was reversed. Within blocks, the order of sentences was randomized. The same random order of sentences was used in each block, thus maximizing the distance between sentences containing the same verb. The participants were randomly assigned to two groups, one receiving List 1, the other List 2.

The subject noun phrases were different for each verb, but were the same for the two sentences each verb appeared in. The nouns used in the subject noun phrases for the two sets of verbs did not differ significantly in frequency, length in letters, phonemes, or syllables (all $t(18) < 1.6$, $p > .15$), or in plausibility as subjects for the verbs they appeared with, as estimated by the method described in (Keller, Lapata, & Ourioupina, 2002).

The ambiguous noun phrases (e.g. *The python escaped the zoo…*) did not differ in frequency, length in letters, phonemes, or syllables, or in estimated plausibility as objects of the verbs they appeared with (all $t(18) < 1.5$, $p > .15$).

The forty experimental sentences were pseudorandomly interleaved with 177 filler sentences of various syntactic structures, which represented stimuli for two other experiments.

Procedure

Sentences were recorded in a sound booth as 16-bit digital sound files at a sampling rate of 44.1 kHz and resampled to 22.1 kHz. Participants were told to read each sentence silently first, until they felt confident that they understood it and could say it without difficulty. They were also told that there was no limit on the amount of time they could take to read and record the sentences, and that the recordings would be used as stimuli in a comprehension experiment. If participants felt that they had not said a sentence in a natural manner, they were asked to repeat it. When speakers misspoke or hesitated or used a noticeably exaggerated pronunciation after recovering from a garden-path ("*oh, I get it,… When the python escaped THE ZOO, it…"*), the experimenter asked them to repeat the sentence in question.

All measurements were performed using the Praat software package (Boersma & Weenik, 2002-2005), based on listening and on visual inspection of the waveform and spectrogram. All durations were measured by a group of three Psychology undergraduate students and one

[1] Strictly speaking, the probability estimates are based on frequencies of syntactic patterns, such as clauses or direct objects, whereas the boundaries in question are prosodic, not syntactic.

Table 1: Verb transitivity biases (from Gahl et al., 2004) and stimulus sentences.

Bias	Verb	Trans.-bias	Sentence
Intransitive	beg	.13	If the tenants beg [,] the landlord [, he] is going to let them stay another month.
	continue	.7	Although the storyteller continued [,] the story [, it] had lost its appeal.
	dance	.7	When the radiant ballerina danced [,] the role [, it] became world-famous.
	escape	.27	When the python escaped [,] the zoo [, it] had to be closed to the public.
	help	.15	Because the teacher helped [,] the children [, they] enjoyed the lessons.
Transitive	accept	.94	Soon after the candidate accepted [,] the money [, it] was found to be illegal.
	attack	.95	When the warriors attacked [,] the city [, it] was nearly destroyed.
	lose	.96	Even though the team lost [,] the match [, it] meant a big success for the new coach.
	pay	.81	When home-owners pay [,] the tax [, it] benefits local public schools.
	perform	.85	Whenever the group performs [,] the show [, it] is sold out.

Linguistics graduate student and checked by the first author. All measurements were condition-blind.

We measured the duration of the verb, the silence following the verb, and the ambiguous noun phrase from its beginning up to the beginning of the next word. The duration of each region was measured from the release of the initial stop for words that started with stops, and from the onset of the initial segment for all other words. When a region began with a stop, the stop's closure was included as part of the preceding silence, since it would have been impossible to identify the beginning of the closure portion of a stop in a period of silence. For analogous reasons, we treated the closure of a region-final stop as the endpoint of that region. This means that the postverbal silence includes the closure portion of any verb-final stops.

All measures were submitted to repeated measures analyses of variance (ANOVA) with speakers (F1) and verbs (F2) as random factors and List (randomly assigned presentation list 1 or 2), Item-group (randomly assigned verb group 1 or 2), Bias (Transitive or Intransitive) and Syntax (Early or Late Closure) as factors. Nonsignificant effects that were of no theoretical interest are not reported here. Effects are reported as significant when at or below the .05 level of significance.

Participants were free to choose which parts of the sentences they wished to accent. Inspection of the data revealed that there were regularities in the location of accent: Speakers tended to accent the subject of the subordinate clause if that subject was a personal name (for example, *When TIM cooks…)*; otherwise, they tended to accent either the verb (in Early-closure sentences) or the following noun phrase (in Late-closure sentences)[2]. Half of the subjects of the subordinate clauses in our materials were personal names. As Schafer and Jun (2001) note, the presence of focal accent on subject noun phrases affects variation in verb durations and might therefore mask the effect under investigation here. We therefore decided to remove sentences with personal names from the analysis. This left a total of 20 sentences, 5 in each combination of verb bias and syntactic structure. This subset of the materials also did not differ significantly in frequency, length, or plausibility of the verbs and their NP arguments across the Transitive Bias and Intransitive Bias items (all $p > .15$).

Predictions

If structural probabilities have similar effects in these sentences as in DO/SC sentences, we should expect words near low-probability boundaries to be lengthened more than those near high-probability boundaries. We therefore make two specific predictions: (1) Intransitive tokens of Transitive Bias verbs (i.e. in Early Closure sentences) will be lengthened more than intransitive tokens of Intransitive Bias verbs; (2) Direct objects of Intransitive Bias verbs (i.e. in Late-closure sentences) will be lengthened more than those of Transitive Bias verbs.

Results

The results of the duration measurements are summarized in Table 2, collapsing across lists and item-groups.

Table 2: Average durations (in milliseconds) of verbs, postverbal silences and postverbal noun phrases in Closure sentences.

Syntax	Bias	Verb	Silence	NP
Early Closure	Intransitive	381	243	414
	Transitive	467	200	381
Late Closure	Intransitive	290	54	704
	Transitive	329	32	626

Consistent with earlier studies, the pronunciation of sentences with Early Closure (*If the tenants beg, the*

[2] Accent placement was judged independently by three coders. We did not attempt to distinguish between focal accent and contrastive accent (Pierrehumbert & Hirschberg, 1990). In cases of coder disagreement, we listened to the recordings again; if no consensus could be reached, that sentence was excluded from the analysis of stress placement.

landlord is going to…) and Late Closure (*If the tenants beg the landlord, he is going to…)* differed, yielding a main effect of Syntax: Intransitive verb tokens and any postverbal pauses were longer than transitive ones (424 vs. 310 ms, $F1(1,19) = 345.3$, $p < .001$, $F2(1,8) = 239.6$, $p < .001$ for the verb; 222 vs. 43 ms, $F1(1,19) = 34.0$, $p < .001$, $F2(1,8) = 149.0$, $p < .001$ for the postverbal pause). Noun phrases representing direct objects (*If the tenants beg the landlord, he…*) were longer than main clause subjects (*If the tenants beg, the landlord might…*) (665 vs. 398 ms, $F1(1,19) = 260.4$, $p < .001$, $F2(1,8) = 215.6$, $p < .001$). These patterns are similar to those reported elsewhere (e.g. Warren, 1985).

The Bias x Syntax interaction was crucial to evaluating our specific predictions. Consistent with our predictions, Transitive Bias verbs in Early Closure sentences were lengthened more, compared to their duration in Late Closure sentences (by 138 ms), than Intransitive Bias verbs in the same contexts (by 91 ms), yielding a significant interaction of Syntax by Bias ($F1(1,19) = 26.3$, $p < .001$, $F2(1,8) = 9.3$, $p < .02$). Figure 1 shows the amount of verb lengthening observed in early closure sentences for intransitive-bias and transitive bias verbs, compared to the durations of the same verbs in late-closure sentences. For the postverbal silence, the Bias x Syntax interaction did not reach significance ($F1(1,19) = 2.3$, n.s.; $F2(1,8) < 1$, n.s.).

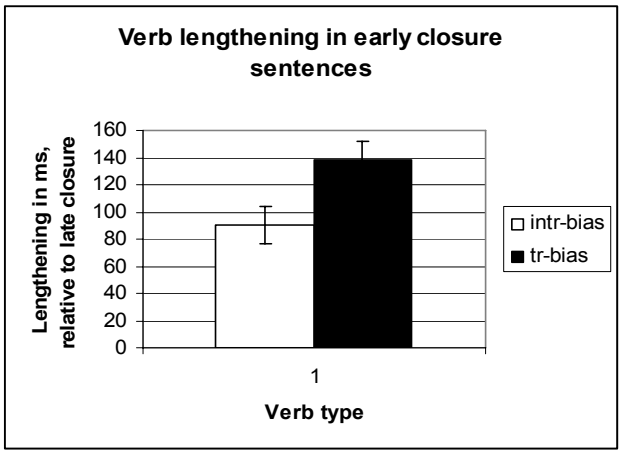

Figure 1: Verb lengthening in Early Closure sentences, relative to verb duration in Late Closure sentences, for intransitive-bias verbs and transitive-bias verbs.

Also consistent with our predictions, noun phrases representing direct objects of Intransitive Bias verbs were lengthened more, compared to their duration as subject noun phrases (by 290 ms), than noun phrases representing direct objects of Transitive Bias verbs (by 245 ms), yielding an interaction of Syntax by Bias that was significant in the analysis by subjects, but not by items ($F1(1,19) = 6.1$, $p < .05$, $F2(1,8) = 2.6$, $p = .14$). Figure 2 shows the amount of NP lengthening observed in late closure sentences with intransitive-bias and transitive bias verbs, compared to the durations of the same NPs in early-closure sentences. The postverbal silences did not yield a significant interaction ($F1(1,19) < 1$, $F2(1,8) < 1$, n.s.). In sum, the pattern of durations was consistent with our hypotheses about the duration of verbs and ambiguous noun phrases.

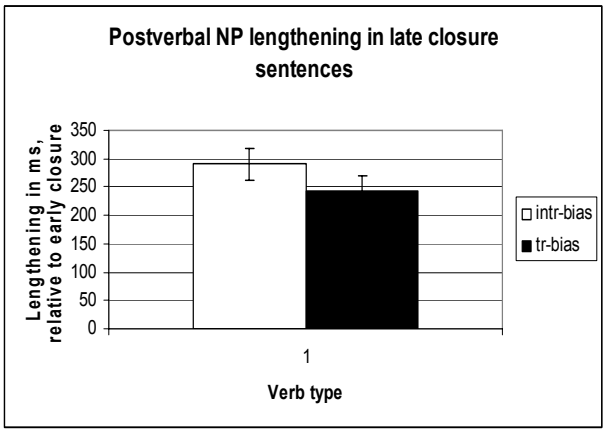

Figure 2: Lengthening of postverbal noun phrases in Late Closure sentences, relative to the duration of those phrases in Early Closure sentences, following intransitive-bias verbs and transitive-bias verbs.

Discussion

The probability of encountering a particular syntactic structure given a verb (or "verb bias"), has long been known to affect language comprehension. In addition, a recent paper (Gahl & Garnsey, 2004) showed that the probability of a verb's taking a direct object vs. a sentential complement also affects language production. In this report, we extended the investigation to a different syntactic pattern, namely sentences with initial subordinate clauses. We found that the pronunciation of high-probability (bias-matching) sentences differed from that of low-probability (bias-violating) sentences, in a pattern similar to that observed by Gahl and Garnsey (2004).

Since accent placement affects duration, could regularities in accent placement have given rise to the observed effects? It is certainly the case that the overall differences between Early Closure and Late Closure sentences in part reflect accent. In our data, intransitive verb tokens tended to be accented, as were noun phrases representing direct objects, a pattern that would reinforce the characteristic lengthening of intransitive verb tokens and direct objects. However, accent placement did not depend on verb bias or on the interaction of bias with syntax: Transitive Bias and Intransitive Bias verbs were accented in 86 and 85 of the 100 intransitive clauses with those verbs; direct objects were accented 69 times (out of a possible 100) following Transitive Bias verbs and 64 times following Intransitive Bias verbs. Thus, accent placement did not vary as a function of bias and therefore cannot explain the observed effects of verb bias.

However, the relationship between stress and verb bias merits further investigation. Some conditions that license intransitive uses of certain heavily Transitive Bias verbs

(Goldberg, 2001) also tend to shift focal accent onto those verbs; hence, accent could reflect verb bias, although it did not do so in our data.

The present findings are subject to limitations: Our analysis was based on only ten verbs. We are currently conducting a follow-up study with full noun phrases (no proper names) and a greater number of verbs. However, the generalizability of our results is supported by the fact that Gahl and Garnsey (2004) found similar effects with a larger number of verbs and sentences of a different type. A second limitation stems from the fact that our findings are based on read speech, whose prosody is known to differ from spontaneous speech (Schafer et al., 2000). It remains to be seen whether the patterns we observed in read speech can also be observed in spontaneous speech.

Our findings have implications for research on pronunciation variation and sentence processing. Research on probabilistic pronunciation variation has until now focused on effects of n-gram probabilities (lexical frequencies and word-to-word transitional probabilities) (Bell et al., 2002; Gregory et al., 1999; Jurafsky et al., 2001). Our findings show that pronunciation also reflects probabilities of syntactic constructs, replicating and extending results reported by Gahl and Garnsey (2004).

Research on sentence comprehension has shown that, generally speaking, bias-violating sentences induce longer processing times and higher error rates than bias-conforming sentences (Garnsey, Pearlmutter, Myers, & Lotocky, 1997; Trueswell et al., 1993). The bulk of this evidence has come from reading. Studies of effects of verb biases on auditory comprehension have shown that prosody and verb biases jointly affect comprehension (Blodgett, 2004). If verb biases themselves affect the production of prosody, it is possible that listeners take this fact into account, such that, for example, strong prosodic cues may facilitate the comprehension of bias-violating sentences. We are currently investigating this possibility. A related possibility is that, if bias-violating sentences are characterized by salient prosodic patterns, prosody may serve as a cue for bias.

A contentious question has been at what point in the comprehension process verb bias comes into play (e.g. Pickering & Traxler, 2003). Given that biases affect production as well as comprehension, it is plausible, in our opinion, to suppose that structural probabilities are part and parcel of speakers' knowledge of their lexicon and syntax, which underlies both comprehension and production and affects both immediately. This point is discussed further in Gahl and Garnsey (2004).

Our work illustrates that studies of language comprehension and production can inform each other, which we consider a promising direction for future research.

Acknowledgments

We are grateful to our colleagues at the Beckman Institute for useful feedback. We would also like to thank Sandra Chavarría-Klahn for help with the stress judgment task and Anthi Diavastou, Yoko Ieuji, Russell Hellin, and Dora Lu for help with the duration measurements. This work was supported by NIH grant HD044458 to the University of Illinois.

References

Beckman, M. E., & Hirschberg, J. (1994). *The ToBI annotation conventions*, [Online ms]. Available: http://www.ling.ohio-state.edu/~tobi/ame_tobi/annotation_conventions.html.

Beckman, M. E., Hirschberg, J., & Shattuck-Hufnagel, S. (2005). The original ToBI system and the evolution of the ToBI framework. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing*. Oxford: Oxford University Press.

Bell, A., Gregory, M. L., Brenier, J., Jurafsky, D., Ikeno, A., & Girand, C. (2002, September 14-15, 2002). Which predictability measures affect content word duration? *Paper presented at the Pronunciation Modeling and Lexicon Adaptation for Spoken Language Technology Workshop*, Estes Park, CO

Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., Gregory, M. L., & Gildea, D. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *Journal of the Acoustical Society of America, 113*, 1001-1024.

Blodgett, A. (2004). *The interaction of prosodic phrasing, verb bias, and plausibility during spoken sentence comprehension.* The Ohio State University.

Boersma, P., & Weenik, D. (2002-2005). Praat: doing phonetics by computer (Version 4.3.04) (Version 4.0.26). Amsterdam.

Gahl, S., & Garnsey, S. M. (2004). Knowledge Of Grammar, Knowledge Of Usage: Syntactic Probabilities Affect Pronunciation Variation. *Language, 80*(4), 748-775.

Gahl, S., Jurafsky, D., & Roland, D. (2004). Verb subcategorization frequencies: American English corpus data, methodological studies, and cross-corpus comparisons. *Behavior Research Methods, Instruments, & Computers, 36*(3), 432-443.

Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory & Language, 37*(1), 58-93.

Goldberg, A. (2001). Patient Arguments of Causative Verbs Can Be Omitted: The Role of Information Structure in Argument Distribution. *Language Sciences, 23*(4-5), 503-524.

Gregory, M. L., Raymond, W. D., Bell, A., Fosler-Lussier, E., & Jurafsky, D. (1999). The effects of collocational strength and contextual probability in lexical production. *Chicago Linguistic Society, 35*, 151-166.

Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. [References], *Bybee,*

Joan (Ed); Hopper, Paul. (2001). Frequency and the emergence of linguistic structure. Typological studies in language, vol. 45. Amsterdam, Netherlands: John Benjamins Publishing Company.

Keller, F., Lapata, M., & Ourioupina, O. (2002). Using the web to overcome data sparseness. *Paper presented at the Paper presented at the Conference on Empirical Methods in Natural Language Processing*, Philadelphia

Kjelgaard, M. M., & Speer, S. R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory & Language. Vol, 40*(2), 153-194.

Levin, B. (1993). *English verb classes and alternations*.Chicago: University of Chicago Press.

Pickering, M. J., & Traxler, M. J. (2003). Evidence against the use of subcategorisation frequency in the processing of unbounded dependencies. *Language & Cognitive Processes. Vol, 18*(4), 469-503.

Pierrehumbert, J., & Hirschberg, J. (1990). The meaning of intonation in the interpretation of discourse. In P. R. Cohen & J. Morgan & M. E. Pollack (Eds.), *Intentions in Communication*. Cambridge, MA: MIT Press.

Schafer, A., & Jun, S.-A. (2001). Effects of focus on prosodic reflections of phrase structure in American English. *Paper presented at the The Prosody in Processing Workshop*, Utrecht University, Utrecht, Netherlands.

Schafer, A. J., Speer, S. R., Warren, P., & White, S. (2000). Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research. Vol, 29*(2), 169-182.

Silverman, K., Beckman, M. E., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J., & Hirschberg, J. (1992). ToBI: A standard for labeling English prosody. *Paper presented at the International Conference on Spoken Language Processing*, Banff, Canada (pp.867-870).

Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, & Cognition. Vol, 19*(3), 528-553.

Warren, P. (1985). *The temporal organization and perception of speech.* University of Cambridge, Cambridge.

Warren, P., Grabe, E., & Nolan, F. (1995). Prosody, phonology and parsing in closure ambiguities. *Language & Cognitive Processes. Vol, 10*(5), 457-486.