Available online at www.sciencedirect.com
ScienceDirect
Journal of Phonetics 35 (2007) 162–179
Journal of Phonetics
www.elsevier.com/locate/phonetics

Prosodic planning: Effects of phrasal length and complexity on pause duration

Jelena Krivokapić*

Department of Linguistics, University of Southern California, 3601 Watt Way, GFS 301, Los Angeles, CA 90089-1693, USA

Received 27 April 2004; received in revised form 20 February 2006; accepted 25 April 2006

Abstract

Research on pause duration has mainly focused on the impact of syntactic structure on the duration of pauses within an utterance and on the impact of syntax, discourse, and prosodic structure on the likelihood of pause occurrence. Relatively little is known about what factors play a role in determining the *duration* of pauses *between* utterances or phrases. Two experiments examining the effect of prosodic structure and phrase length on pause duration are reported. Subjects read sentences varying along the following parameters: (a) the length in syllables of the intonation phrase (IP) preceding and following the pause, and (b) the prosodic structure of the IP preceding and following the pause, specifically whether or not the IP branches into smaller phrases. In order to minimize variability due to speech rate and individual differences, speakers read sentences synchronously in dyads. The results showed a significant postboundary effect of prosodic branching and significant pre and postboundary phrase length effects. The results are discussed in terms of production units.

1. Introduction

Starting with work done in the 1960s (see the overview in Goldman Eisler, 1968), researchers have been investigating the variables determining the occurrence and duration of pauses and have explored the implications of their findings for models of speech production. For example, based on studies of silent intervals, it seems likely that syntactically complex phrases are more demanding for the production system than syntactically simple phrases because pauses are longer between syntactically complex phrases (e.g., Ferreira, 1991; Strangert, 1997). Through the study of pause duration and sentence initiation times we are beginning to understand the more intricate mechanisms of speech planning and production.

The present study examines the impact of prosodic structure and length on pause duration. The results are analyzed from the point of view of modeling speech production, addressing in particular issues relating to the domain of planning units in production and the possible causes of pre and postboundary effects on the duration of pauses.

*Tel.: +1 213 740 2986; fax: +1 213 740 9306.
E-mail address: krivokap@usc.edu.

0095-4470/$ - see front matter
doi:10.1016/j.wocn.2006.04.001

2. Prosodic structure

A brief introduction to the levels of the prosodic hierarchy assumed in this experiment is in order, as different studies argue for the existence of different prosodic constituents (see Shattuck-Hufnagel & Turk, 1996 for an overview). In this study the prosodic categories assumed in the intonational framework of Pierrehumbert and Beckman (Beckman & Pierrehumbert, 1986) are adopted. Above the word, the two levels of prosodic phrasing in English are the intonation phrase (IP) and the intermediate phrase (ip). The IP is the largest unit and is defined as the domain of a coherent intonational contour that has a nuclear pitch accent, a phrase accent, and a boundary tone. IPs further branch into ip's. The ip includes at least a nuclear pitch accent and a phrase accent. It further branches into words, which are in turn composed of syllables. Fig. 1 schematizes those aspects of the prosodic hierarchy in English which are of immediate relevance here.

While the intonational model adopted here is widely used, it should be noted that the issue of prosodic structure is not settled. Apart from the number and nature of prosodic constituents, a question that is disputed is whether prosodic structure is recursive, meaning whether one prosodic constituent can be embedded within a prosodic constituent of the same type (for example, an IP embedded within another IP), as argued for by Ladd (1988), or whether the hierarchical structure is nonrecursive, as assumed by the Strict Layer Hypothesis (Nespor & Vogel, 1986; Selkirk, 1984; see Ladd, 1996, p. 235ff for a discussion of the opposing views). The nonrecursive view is manifest in the Tones and Break Indices (ToBI) prosodic annotation system (Beckman & Elam, 1997). This system describes prosodic boundaries and tonal events of the prosodic structure. The basis of ToBI is the work of Pierrehumbert and Beckman (Beckman & Pierrehumbert, 1986; Pierrehumbert, 1980) for intonation and the work of Price, Ostendorf, Shattuck-Hufnagel, and Fong (1991) and Wightman, Shattuck-Hufnagel, Ostendorf, and Price (1992) for the break indices. The distinct prosodic break indices in ToBI correspond to the three distinct prosodic categories: word, ip and IP. For this study, the constituents ip and IP of Beckman and Pierrehumbert's (1986) model are adopted, and generally the ToBI guidelines (Beckman & Elam, 1997) are used to assign prosodic structure to the production of sentences. Contrary to this framework, here the possibility is left open that embedding of prosodic constituents of the same category, in particular the embedding of IPs within IPs, is possible (this issue will be briefly addressed again in Section 4.3.2).

3. Pauses

3.1. Pause occurrence

Pause occurrence has been examined in terms of prosodic structure, syntactic structure, speaker variability and speaking task.[1] Goldman Eisler (1968) shows that pause occurrence depends on whether the utterance is spoken spontaneously or read, in that in read speech pauses occur at grammatical junctures only, whereas in spontaneous speech they are present elsewhere in the utterance as well. In a number of studies it has been shown that the slower the speech, the more frequent the pauses (e.g., Goldman Eisler, 1968; Lane & Grosjean, 1973; Fletcher, 1987 for French, and for German Butcher, 1981 and Trouvain & Grice, 1999), although not all speakers decrease pause frequency when increasing their speech rate (Trouvain & Grice, 1999). Lastly, in both read and spontaneous speech there are individual differences between speakers and fewer pauses in well-learned material (Goldman Eisler, 1968).

Most of the other factors have only been investigated in read speech, since it is only in read speech that these factors can easily be controlled. In terms of prosodic structure, it is generally accepted that pauses can, but do not have to, occur between IPs and do not occur elsewhere (Nespor & Vogel, 1986; Selkirk, 1981). The length of a phrase also plays a role in boundary placement. Grosjean, Grosjean, and Lane (1979) show that speakers tend to pause in the middle of sentences and in the middle of syntactic constituents, even if syntactic complexity would predict the pause to occur at another location (see the next paragraph on the relation between syntactic complexity and pause occurrence). Further, Watson and Gibson (2004) show that taking into account the number of phonological phrases within a preceding and a following syntactic phrase, and

[1]Unless otherwise indicated, the studies cited refer to English.

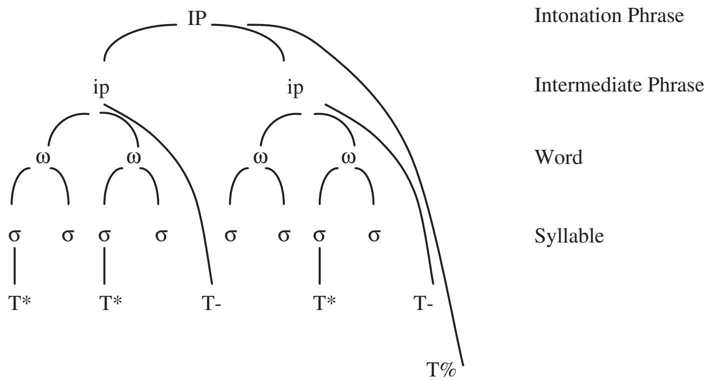


Fig. 1. Prosodic constituents (adapted from Beckman, 1989). "T*" stands for different types of pitch accents, 'T-' for phrase accents and 'T%' for boundary tones. Only the immediately relevant aspects of prosodic structure are shown.

thus taking the length of syntactic phrases into account, is a good predictor of how likely an IP boundary is to occur. Since in read speech the probability of an IP boundary is highly correlated to the probability of a pause being produced, this means that the length of syntactic phrases is an important determinant of pause occurrence.

Research on the effects of syntactic structure on pause occurrence has shown that pauses often occur at major syntactic boundaries and that the more complex the syntactic constituent, the more likely a pause is to occur (see for example Cooper & Paccia-Cooper, 1980; Grosjean et al., 1979; Strangert, 1991 for Swedish). Certain syntactic structures are known to almost always coincide with an IP boundary (and possibly a pause at the end of the IP). Namely, vocatives, appositives, preposed clauses, nonrestrictive relative clauses, and parenthetical expressions obligatorily constitute IPs (Selkirk, 1995, p. 567). Furthermore, there are phrasings that are disallowed due to syntactic constraints.

While syntax constrains the prosodic realization of sentences, it does not fully determine it. This can be shown in cases where one sentence has more than one prosodic phrasing, as given below:

(1) (a) (Elizabeth agreed reluctantly)$_{\mathrm{IP}}$,
(b) (Elizabeth) $_{\mathrm{IP}}$ (agreed reluctantly)$_{\mathrm{IP}}$,
(c) (Elizabeth agreed)$_{\mathrm{IP}}$ (reluctantly)$_{\mathrm{IP}}$.

In sum, the occurrence of a pause is constrained by a number of factors: prosodic, syntactic, length, speaker and speaking task, and rate.

3.2. *Pause duration*

As with pause occurrence, pause duration is determined by a number of factors: speaker, speech rate, discourse, prosodic structure, phrase length, and syntax.

Goldman Eisler (1968) notes that pause duration varies with individual speakers (see also Butcher, 1981). This is more manifest sentence-internally, while between complete sentences, pauses show a more uniform pattern across subjects (as shown by Fant, Kruckenberg, & Ferreira, 2003 for Swedish). Pause duration is also dependent on speaking rate. It often increases in slower speech (see e.g., Goldman Eisler, 1968; Fletcher, 1987 for French; and Trouvain & Grice, 1999 for German) although not always (Butcher, 1981 reports no increase in pause duration in German, and only some speakers show such a pattern in Fletcher's study of French; Fletcher, 1987). Taken together, these findings indicate speaker dependency rather than a speech rate dependency alone.

A structural factor affecting pause duration is discourse organization. Speakers have been shown to pause longer at topic shift than at other discourse boundaries (e.g., Smith, 2004 for English; Bannert, Botinis, Gawronska, Katsika, & Sandblom, 2003 for Greek; Swerts & Geluykens, 1994, for Dutch spontaneous speech). Pause duration also depends on the type of text being read. Fant et al. (2003) examine Swedish news readings over the radio as compared to readings of a novel, and find that pauses between sentences are shorter in news readings than in readings of novels.

The main body of research has addressed whether and how syntactic structure plays a role in pause duration. This research is directly relevant to the experiment reported below. Cooper and Paccia-Cooper (1980) report that boundaries inside a major constituent are weaker than those between major constituents, and similarly, pauses within sentences have been found to be shorter than pauses at sentence boundaries (Butcher, 1981; Grosjean et al.,1979; Sanderman & Collier, 1995 for Dutch). Grosjean et al. (1979) look at the relation between pause duration and syntactic complexity. Their main finding is that there is a strong correlation between syntactic complexity and pause duration. Even though there are mismatches, it generally holds that the more complex the preceding and following syntactic constituent, the longer the pause duration (Grosjean et al., 1979; see also Ferreira, 1991; Strangert, 1991, 1997 for Swedish). The weighting of the preceding and upcoming syntactic phrase is not well understood. Cooper and Paccia-Cooper (1980) report that preboundary phrases determine boundary strength to a larger degree than postboundary phrases. However, studies examining pause duration between a subject NP and the following VP have found varying degrees of effect of length and complexity of the subject NP and of the VP (see the different effects reported in Ferreira, 1991; Sanderman & Collier, 1995, for Dutch; Strangert, 1997 for Swedish). In sum, studies examining syntactic effects on pause length show that pause duration is influenced by syntactic structure in that it increases with length and complexity of surrounding syntactic constituents. The strength of effect of the preceding and following syntactic constituents differs across studies.

A number of studies show the effect of prosodic structure and length. Gee and Grosjean (1983) look at the sentences examined by Grosjean et al. (1979) and achieve stronger predictions for pause duration by including both syntactic and prosodic structure into the calculation of pause duration, rather than just syntactic structure (as in Grosjean et al., 1979). In a seminal study, Ferreira (1993) varies prosodic and syntactic structure affecting the pause duration after a constituent and shows that it is the prosodic structure that induces pause duration effects, not the syntactic structure. Similarly Martin (1970) reports that in the perception of pauses in spontaneous speech, prosodic structure (final lengthening) overrides syntactic structure. A study by Horne, Strangert, and Heldner (1995) shows further effects of prosody on pause duration. In their study of Swedish, they show that pause duration increases with prosodic boundary strength. A corpus-based study by Choi (2003) finds similarly that pause duration increases from nonboundary to intermediate to intonation boundary, and also that pause length depends on boundary tones, in that pause duration increases in the order of: L–H%, H–L%, H–H%, L–L%.

A study that specifically investigates the effects of adjacent phrase length of the prosodic constituents is that of Zvonik and Cummins (2003). They show that the length of the IPs preceding and following the boundary has an effect on pause duration. Short pauses are associated with the length of the preceding and following IP. Pauses of less than 300 ms almost only occur when the preceding or following IP is less than ten syllables. For pauses of longer than 300 ms there is no predictability. The probability of a pause being less than 300 ms rises when both the IP preceding and following the pause are less than ten syllables. An earlier study, Zvonik and Cummins (2002), also indicates that pause length might be influenced by the length of the preceding phrase (see Fig. 1, in Zvonik & Cummins, 2002). Ferreira (1991) also shows that sentence initiation time increases with the number of phonological words in a sentence.

The main conclusion to be drawn from these studies is that length of prosodic utterance plays a role in pause placement and pause duration and that syntactic structure has a large impact on pause placement and duration, but does not fully determine it. The varying syntactic effects and the studies by Ferreira (1993) and Gee and Grosjean (1983) suggest that syntactic effects might be better described if related to prosodic structure. The studies further show that there is large variability in both pause placement and in pause duration due to speaker variability and speaking rate.

The present study is motivated by the studies of Ferreira (1991) and Zvonik and Cummins (2003) showing phrase length effects and by the hypothesis that prosody and not syntactic structure might be the crucial factor in determining pause duration (Ferreira, 1993). As a first step in assessing the impact of prosodic structure on pause duration, the first experiment described here is designed to test the effect of prosodic phrase length on pause duration. The second experiment examines the effect of prosodic branching on pause duration.

Zvonik and Cummins (2003) showed a relation between syllable count and short pauses (i.e., between phrases less than ten syllables long and pauses less than 300 ms long). The experiment reported here tests the relation between two very different syllable counts (two syllables and 38 syllables) for preceding and following phrases and pause duration.

The present experiment was conducted using the synchronous reading paradigm introduced by Cummins (2002, 2003) and Zvonik and Cummins (2002, 2003) in which dyads read a text simultaneously. Cummins (2002) compares the pause placement for speakers in solo reading with pause placement in synchronous reading by dyads. He shows that variability in pause placement is significantly reduced in synchronous speech. Zvonik and Cummins (2002) find that synchronous speech also reduces variability in pause duration. The distinct advantages of the synchronous speech paradigm are that the reduction of variability in both pause placement and pause duration facilitates comparison across speakers. Furthermore, Cummins (2002, 2003) and Zvonik and Cummins (2003) suggest that this method captures in important ways the speakers' shared knowledge of *linguistic* timing, minimizing individual or stylistic differences.

Zvonik and Cummins's (2002) study also showed that while pause duration was less variable in synchronous speech, the length of pauses was comparable—speakers paused both synchronously and in solo readings longer after a longer phrase—again indicating that the synchronous reading does not invoke its own dynamics.

4. Experiments

Two experiments are presented. The first experiment tests for an effect of the length of adjacent phrases on pause duration. The second tests the effects of prosodic branching. The stimuli for the two experiments were combined, randomized and recorded in one session.

4.1. Method: Experiment I (length effects)

4.1.1. Stimuli

Twelve sentences were constructed (see Appendix A) to measure the effect of phrase length on pause duration. The independent variables were: (a) the length of the preboundary phrase (whether it is long or short) and (b) the length of the postboundary phrase (whether it is long or short). As these factors were crossed, this yields four conditions (long/long, long/short, short/long, short/short). For each condition, three sentences were constructed. The short sentences had two syllables and the long sentences 38 syllables. To ensure that the pause duration did not vary due to acoustic properties of the phonemes preceding and following it and to ensure that the pause interval could be consistently measured, the last phoneme preceding the pause was a voiceless stop, and the first phoneme following the pause was a vowel in all 12 sentences.

The 12 sentences were put into four blocks (one block corresponding to each condition). The materials for this experiment were combined with 12 sentences for Experiment II (see Section 4.3.1), which were also put together into four blocks of three sentences. The resulting eight blocks were randomized for each dyad. The sentences within each block were also randomized for each dyad. The randomized list of sentences was read three times by each dyad. Data from eight dyads were collected, and the results of seven dyads are reported, as one dyad was unable to complete the experiments. This yields a total of 252 sentences for each experiment, and 504 sentences for both experiments together ((12 sentences for Experiment I + 12 sentences for Experiment II) $\times$ 3 repetitions of each sentence $\times$ 7 dyads).

4.1.2. Subjects

The subjects are native speakers of American English, all of them students at the University of Southern California. They were paid for their participation and were naive as to the purpose of the experiment.

The two subjects were seated facing each other. Prior to the start of the recording, they were given the sentences and asked to familiarize themselves with them, by reading them aloud. Already during this familiarization period subjects started reading the sentences synchronously, even though they were not required to do so. Once familiar with the sentences, the subjects were asked to read them aloud, together with their co-speaker, in their normal reading styles, at the prompt of the experimenter. In cases of errors, they were asked to read the sentence again.

4.1.3. Recording

Subjects were recorded on a DAT recorder, using two Shure head mounted unidirectional microphones. The recordings were made at a 44,100 Hz sampling rate. These recordings were transferred to a PC onto the right and left channels of a stereo file at a sampling rate of 22,050 Hz.

4.1.4. Measurements

Due to nonfluent productions and mispronunciations, 5.2% of the total number of sentences in Experiment I were discarded. The data were measured using Audacity.[2] The end of each preboundary phrase was a voiceless stop, and each postboundary phrase started with a vowel. Pause duration was measured from the end of periodic voicing (beginning of the voiceless stop closure) to the beginning of voicing for the following phrase initial vowel. The end and the onset of voicing were straightforward to detect. At the end of the silent interval occasionally there was a single glottal pulse preceding the periodic voicing for the vowel. The glottal pulse was taken to be evidence of the laryngealized onset of voicing often correlating with a glottal onset (see Dilley, Shattuck-Hufnagel, & Ostendorf, 1996). In such cases, the end of the pause was taken to be at the beginning of the glottal pulse. In all other cases, the end and the beginning of periodic voicing were taken to be the beginning and end of the silent interval, respectively. Three sentences (sentences 1, 11 and 12 in Appendix A) ended with two voiceless stops (with the words 'wept' and 'slept'). The beginning of the voiceless stop closure was used, as before, as the onset of the pause. Finally, in a few cases there was in the recording of one speaker interference from the other speaker, in the form of aperiodic noise at the end of the silent interval. Again, the beginning of periodic voicing was taken in these cases to mark the end of the silent interval. The beginning and end of the pause were also used to determine synchrony between the speakers. For each sentence, the temporal difference in the beginning of the pause between the speakers of a dyad was calculated (the average pause beginning difference was 38 ms, $SD = 48$ ms), and the difference in the end of the pause was calculated (the average pause end difference was 79 ms, $SD = 64$ ms).

The author ascertained the prosodic structure of the produced sentences using the ToBI prosodic transcription criteria (Beckman & Elam, 1997). The IP boundaries were realized by final lengthening and a boundary tone. The speakers produced nonbranching structures (an IP preboundary and an IP postboundary). Subjects of a dyad used the same prosody. This was always the case.

Silent intervals of less than 200 ms were excluded from consideration (6.3% of the total number of sentences in Experiment I), since the intervals included the voiceless stops in each sentence and could not with certainty be asserted always to be pauses.

4.1.5. Statistical analysis

For each sentence token, the duration of the pause was averaged between the two speakers of each dyad. A two-factor ANOVA was performed on these data for each dyad separately, testing the effect of the factors: (1) preboundary length, with the levels long and short and (2) postboundary length, with the levels long and short, and the interaction between the two factors.

One of the dyads (dyad R) had very small (<200 ms) pause durations for all the short/short tokens, so this condition was therefore excluded for this dyad, yielding a one-factor ANOVA with three levels for the length factor, with the levels coded as: long/long, long/short and short/long. A Scheffé test was performed on the results of this one-factor ANOVA. Significance for all tests was defined as $p<.05$.

4.2. Results for Experiment I

Significant length effects were found both pre and postboundary for all six dyads (H, J, S, C, L, M), as shown in Table 1. Longer phrases had significantly longer (preceding and following) pauses. Two dyads (S and C),

[2]Free software available at http://audacity.sourceforge.net/.

Table 1
Length effect ANOVA results and means (ms)

	H	J	S	C	L	M
Preboundary	$F(1,29)=27.159$, $p<.001$ *Means* Short: 490 Long: 650	$F(1,32)=36.686$, $p<.001$ *Means* Short: 421 Long: 565	$F(1,24)=39.467$, $p<.001$ *Means* Short: 486 Long: 672	$F(1,30)=82.070$, $p<.001$ *Means* Short: 414 Long: 709	$F(1,28)=102.438$, $p<.001$ *Means* Short: 378 Long: 560	$F(1,29)=6.100$, $p<.05$ *Mean:* Short: 464 Long: 526
Postboundary	$F(1,29)=33.783$, $p<.001$ *Means* Short: 486 Long: 663	$F(1,32)=41.119$, $p<.001$ *Means* Short: 417 Long: 569	$F(1,24)=122.331$, $p<.001$ *Means* Short: 415 Long: 743	$F(1,30)=102.235$, $p<.001$ *Means* Short: 423 Long: 735	$F(1,28)=36.057$, $p<.001$ *Means* Short: 419 Long: 518	$F(1,29)=34.679$, $p<.001$ *Means* Short: 427 Long: 567

The columns refer to individual dyads.

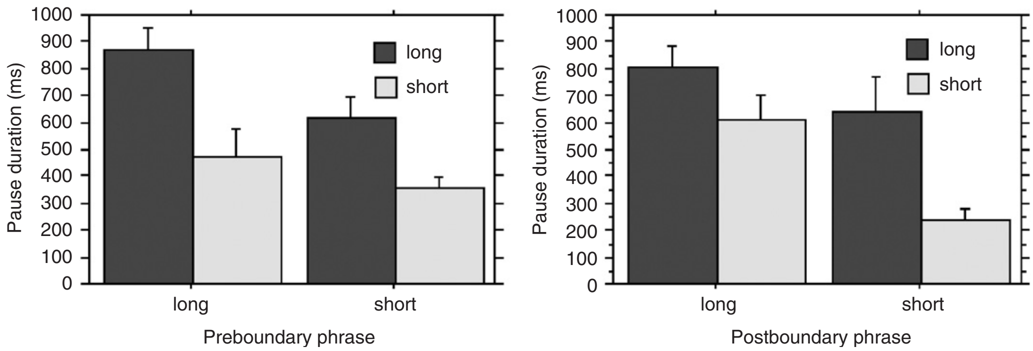


Fig. 2. Interaction of pre and postboundary length effects (pause duration and standard deviations) for dyad S (left) and dyad C (right).

Table 2
Average pause duration (ms) with standard deviations (in brackets) for the length conditions

	H	J	S	R	C	L	M
Long, long	730 (102)	657 (61)	871 (83)	463 (32)	808 (73)	590 (40)	592 (37)
Long, short	569 (56)	473 (67)	474 (1 0 0)	395 (55)	610 (92)	525 (44)	460 (79)
Short, long	588 (56)	481 (81)	616 (80)	463 (53)	641 (128)	445 (71)	541 (90)
Short, short	378 (73)	361 (74)	357 (37)	—	236 (42)	277 (55)	395 (52)

Each row refers to one condition. The columns refer to individual dyads.

had a significant interaction effect: for S $F(1,24)=5.423$, $p<.05$, for C $F(1,30)=12.058$, $p<.05$, shown in Fig. 2.

Table 2 shows the mean pause duration across the length conditions, for the seven dyads. As can be seen from this table, for the six dyads that do have a pause in the short/short condition, this pause is the shortest. Dyad R follows this pattern (by not having a pause at the short/short condition), indicating that the nonexistence of the pause in the short/short condition is likely a floor effect.

For dyad R the results of the one-factor ANOVA show that there is a difference between the long/long, long/short and short/long condition ($F(2,24)=6.036$, $p<.01$. A post hoc test was conducted, and the Scheffé test was chosen as it does not require an equal number of cells across conditions. The results showed a significant difference between the long/long and the long/short condition ($p<.05$), such that the pause for the long/long condition was longer, and between the long/short and the short/long condition ($p<.05$), such that the pause for the short/long condition was longer. Since the conditions are significantly different when

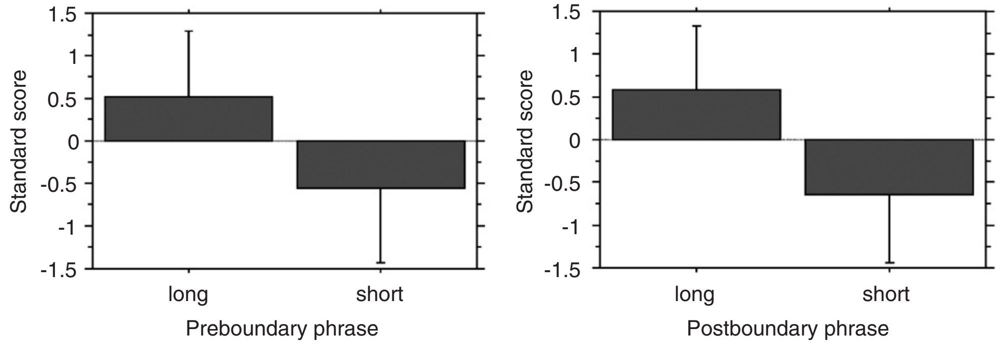


Fig. 3. Preboundary (left) and postboundary (right) length effects (z-scores and standard deviations), all dyads pooled.

comparing the postboundary long and postboundary short factor (in long/long vs. long/short), but not when comparing long/long vs. short/long, it can be concluded that it is the length of the postboundary phrase that has an effect for dyad R.

As the dyads patterned similarly with regard to phrase length, the standard score was taken of the average duration of the pauses for each of the six dyads, and a two-factor ANOVA performed on these data. The use of the standard score allowed the pooling of the data of the six dyads. The results show a significant preboundary effect ($F(1, 192) = 189.862$, $p < .001$) and postboundary effect ($F(1, 192) = 251.604$, $p < .001$), and no interaction effect. The effects are shown in Fig. 3, where a score above zero indicates lengthening (relative to the mean) and a score below zero indicates shortening (relative to the mean).

To summarize, there is a pre and postboundary effect for all six dyads; longer phrases, both before and after the phrasal boundary, yield longer pauses. For the seventh dyad (dyad R), the three length conditions differ, and there is a post hoc significant difference as a function of following phrase length.

4.3. Method: Experiment II (prosodic complexity)

4.3.1. Stimuli

In Experiment II, 12 sentences were constructed (see Appendix B) to measure the effect of prosodic structure on pause duration. The independent variables were: (a) whether or not the IP preceding the pause branches into ip's and (b) whether or not the IP after the pause branches into ip's. As these factors were crossed, this yields four conditions (branching/branching, branching/nonbranching, nonbranching/branching, nonbranching/nonbranching). For each condition, three sentences were constructed. In order to elicit the intended prosodic structure, the syntactic structure and the length of syntactic constituents were varied. Overall the length of the preboundary and postboundary phrases and the point of branching were kept constant across all sentences. The sentences consisted of 28 syllables, and the branching was targeted at the 14th syllable (see Appendix B for details on the stimuli construction). To ensure that the pause duration did not vary due to acoustic properties of the phonemes preceding and following it and to ensure that the pause interval could be consistently measured, the last phoneme preceding the pause was a voiceless stop, and the first phoneme following the pause was a vowel in all 12 sentences.

As already mentioned, the 12 sentences for this experiment were combined with the materials for Experiment I (see Section 4.1.1 for the randomization procedures). Each sentence was read three times by each dyad, yielding 36 sentences, and a total of 252 sentences for all dyads together (12 sentences for Experiment II $\times$ 3 repetitions of each sentence $\times$ 7 dyads).

4.3.2. Subjects, recording and measurements

As the data for Experiments I and II were randomized and collected in one session, the subjects, recording procedures, and pause duration measurement procedures are the same as in Experiment I. The results of the seven dyads that were able to complete the experiment are presented. All pauses were longer than 200 ms. Due

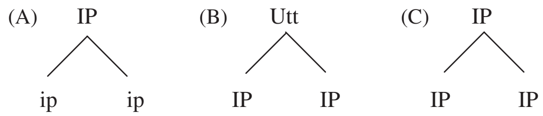


Fig. 4. Prosodic branching. (A) Target structure; (B) and (C) possible structures of realized prosody: (B) according to the SLH; (C) allowing recursion.

to mispronunciations or nonfluent productions, 17% of the tokens were excluded from the analysis. As a measure of synchrony, for each sentence the temporal difference in the beginning of the pause between the speakers of a dyad was calculated (the average pause beginning difference was 33 ms, SD = 27 ms), and the difference in the end of the pause was calculated (the average pause end difference was 71 ms, SD = 58 ms).

The author screened the recordings to determine whether the subjects used the intended prosody (i.e., whether they produced the prosodic structure that the stimuli were designed to elicit). The ToBI criteria for prosodic transcription (Beckman & Elam, 1997) were used, and pre and postboundary phrases were examined for whether they contained ip boundaries, which for the purposes of this experiment were used to indicate prosodic branching. ip boundaries were identified by a phrase accent. The sentences were also checked for IP boundaries, which were identified by phrase accent, final lengthening and a boundary tone. In this way, the prosodic structures used in the analyses below were verified after data collection.

The one deviation speakers made from the target prosodic structure was that they produced multiple ip's in all cases. Significantly however, subjects did produce branching and nonbranching structures in the intended sentences, though the branches were IP rather than ip. So the intended prosodic conditions were achieved in the sense that there were branching and nonbranching prosodic phrases, though at the IP rather than ip level. Theoretically, a phrase branching into two IPs can have two different structures. If the Strict Layer Hypothesis is accepted, the structure is as in Fig. 4B, and if the SLH is not assumed, (4C) is another possible structure. For the purposes of this study, the only relevant property is that there is prosodic branching, the exact structure not of immediate concern. The target prosodic structure for prosodic branching and the possible structures of the realized prosody are schematized in Fig. 4.

Only the tokens where both subjects of a dyad had the same prosodic realization were included in the analysis (across dyads, 7.9% of the total number of tokens were discarded because of mismatches in prosodic realization). It is interesting to note that this is half the rate that general disfluencies occurred at, indicating a high level of within dyad agreement on prosody. Any asynchrony within dyads was caused in most cases by one of the subjects having a stronger prosodic boundary than the other, rarely by one subject having a prosodic boundary and the other not. The number of sentences discarded due to prosodic mismatches was similar across dyads, with dyad J having the smallest number (2.8%) and dyad S having the largest number (13.9%).

4.3.3. Statistical analysis

For each sentence token, the duration of the pause was averaged between the two speakers of each dyad. A two-factor ANOVA was performed on these data for each dyad separately testing the effect of the factors: (1) preboundary prosodic complexity factor (with the two levels branching and nonbranching) and (2) postboundary prosodic complexity factor (with the two levels branching and nonbranching) and their interaction. Significance for all tests was defined as $p < .05$.

4.3.4. Results for Experiment II

The average pause duration times for the prosodic complexity conditions are given in Table 3.

There were no effects of preboundary branching. Two dyads showed a postboundary effect: dyad J ($F(1,24) = 4.271$, $p < .05$) and dyad S ($F(1,14) = 7.031$, $p < .05$). Shorter pauses occur in the postboundary branching condition for these two dyads. One dyad (M) has a significant interaction effect, ($F(1,24) = 6.501$, $p < .05$), shown in Fig. 5, such that in the preboundary nonbranching condition there was a longer pause when the postboundary phrase was nonbranching than when it was branching; whereas a preboundary branching

Table 3
Average pause duration (ms) with standard deviations (in brackets) for the prosodic complexity conditions

	H	J	S	R	C	L	M
Branching, branching	582 (59)	559 (69)	799 (84)	438 (34)	789 (126)	539 (66)	501 (82)
Branching, nonbranching	649 (43)	624 (77)	856 (84)	464 (42)	823 (65)	571 (48)	474 (75)
Nonbranching, branching	617 (109)	580 (75)	709 (51)	422 (50)	726 (41)	529 (105)	468 (80)
Nonbranching, nonbranching	639 (58)	622 (40)	838 (24)	445 (54)	821 (113)	613 (67)	586 (47)

Each row refers to one condition. The columns refer to individual dyads.

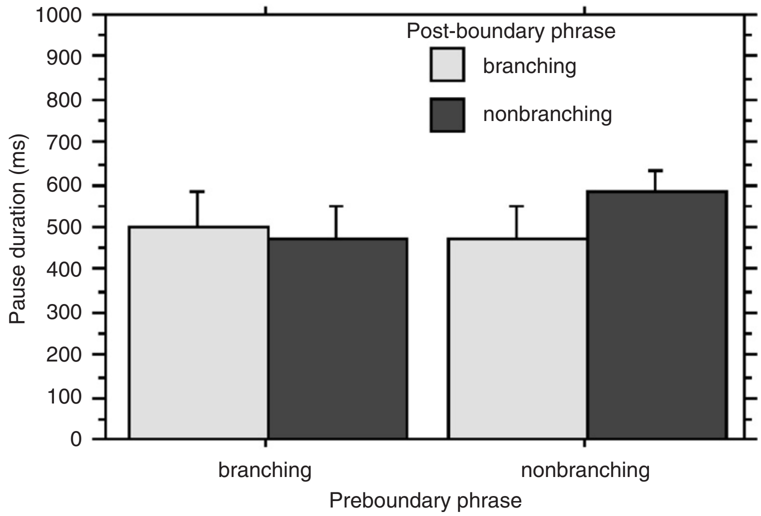


Fig. 5. Interaction of pre and postboundary prosodic complexity effects (pause duration and standard deviations). Dyad M.

Table 4
Postboundary average pause durations (ms) with standard deviations (in brackets)

	H	J	S	R	C	L	M
Branching	601 (88)	568 (70)	772 (85)	429 (43)	762 (102)	535 (86)	486 (80)
Nonbranching	645 (48)	622 (52)	847 (58)	454 (48)	823 (91)	594 (61)	526 (84)

Postboundary 'branching' means that the phrase following the pause was branching, postboundary 'nonbranching' means that the phrase following the boundary was nonbranching. The columns refer to individual dyads.

phrase was associated with a longer pause when the postboundary phrase was branching than when it was nonbranching. No other main effects or interactions are significant.

As can be seen in Table 4, all dyads patterned similarly with regard to postboundary prosodic complexity, in that for each dyad, a postboundary branching phrase had a shorter average pause duration than a postboundary nonbranching phrase. The standard score was taken of the average duration of the pauses for each dyad, and a two-factor ANOVA performed on these data for the seven dyads. The use of the standard score allowed the pooling of the data of the seven dyads. There was a significant postboundary effect ($F(1, 185) = 23.076$, $p < .001$), and no interaction effect. The effects are shown in Fig. 6, where a score above zero indicates lengthening (relative to the dyad's mean) and a score below zero indicates shortening (relative to the dyad's mean). The dyads did not all pattern in the same direction with regard to preboundary prosodic complexity, so the preboundary results will not be discussed in the pooled analysis.

To conclude, there is a prosodic complexity effect, such that postboundary branching induces shorter pauses.

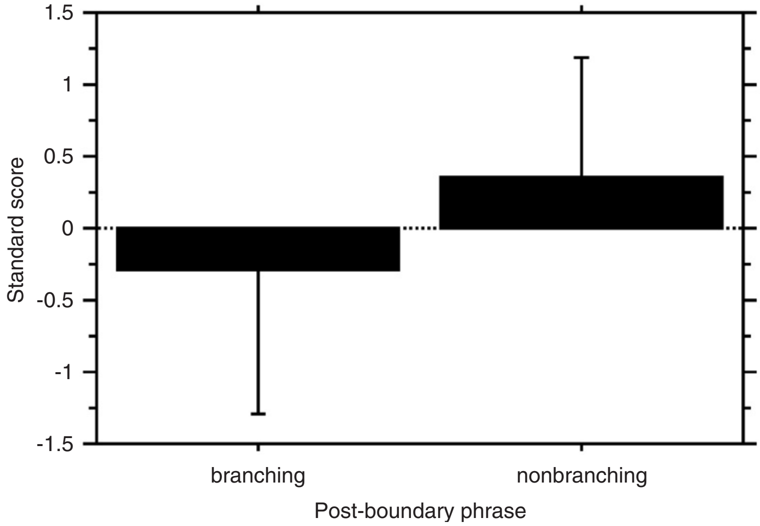


Fig. 6. Postboundary branching effect (z-scores and standard deviations), all dyads pooled.

5. Discussion

The experiments were designed to test the effects of prosodic structure and length on the duration of pauses, in order to further our understanding of the mechanisms of planning in speech production. It was found that the length of prosodic phrases has a significant effect both pre and postboundary. Shorter phrases occur with shorter pauses, and longer phrases with longer pauses. Prosodic branching (i.e., complexity) was also shown to have a significant effect. Postboundary, branching phrases have shorter preceding pauses. In what follows, these results are discussed in terms of speech production units.

5.1. Sources of pause duration effect

Three possible sources are now considered for the effects of phrase length and prosodic complexity on pause duration. The pause duration effects could be due to structural linguistic factors. These factors could be syntactic, prosodic, or related to phonological structure. Syntactic structure has been investigated thoroughly (see the literature review in Section 3.2.), and it has been shown to have an effect on pause duration (e.g., Ferreira, 1991; Grosjean et al., 1979). Prosodic structure is one of the influential factors in the present experiment and has also been shown to have an effect on pause duration (e.g., Choi, 2003; Ferreira, 1993; Gee & Grosjean, 1983). Finally, constituent length can be conceived of in terms of phonological structure (for example, number of feet or, as in the present experiment, number of syllables) and it has been shown to have an effect on pause duration (e.g., Sternberg, Monsell, Knoll, & Wright, 1978; Zvonik & Cummins, 2003).

The second possibility is that information load is influencing pause duration, where information load can be processing demands of some sort. However, the postboundary effects are unlikely to be related to information load, as a study by Ferreira (1991) has shown that semantic implausibility does not have an effect on initiation times for sentences (initiation times are, like the postboundary effects, related to the planning of upcoming material). For the sentences in the present experiment, the information load is likely to be minimal, since the sentences are read not spontaneous, and the subjects are familiar with them.

Finally, pause duration could be due to cognitive factors. For example it could be due to cognitive factors relating to the generation of the phonological structure and the phonetic plan for an intended utterance. A buffer is proposed by Sternberg et al. (1978) for storing the motor program (i.e., the phonetic plan) for a list of words before it is articulated and a similar buffer is proposed by Levelt (1989). Based on their experiments, Sternberg et al. (1978) suggested that when producing a list of words, speakers store the motor program for the words in a short-term memory buffer, and when producing the list, they have to retrieve the motor programs for the words. The time required for the retrieval depends on the number of words stored in the buffer. The larger the number of the words in the list, the longer the retrieval time and the longer the initiation time for the list. Similarly, in explaining syntactic complexity effects, Ferreira (1991) suggests that the capacity of this buffer limits the size of a structural unit that can be encoded. For the experiments presented here, such an explanation can be related to the postboundary effects only, as it is related to speech planning by the speaker (since the speaker is buffering the encoded material prior to articulation).

Next the structural, information and cognitive factors are evaluated, first as they relate to the observed preboundary effects, and then as they relate to the postboundary effects.

5.1.1. Preboundary effects

Preboundary effects could be driven by speaker or listener processing needs. If preboundary effects are listener driven, presumably a pause after a phrase provides time for the processing of that phrase. A listener would presumably require a shorter pause for a just uttered short phrase and a longer pause for a just uttered long phrase. Of the options mentioned, both the structural and the information load factors seem likely. Phonological structural properties of the preceding phrase in terms of syllable (or foot) count could allow the speaker to gauge the listener's processing needs. This would mean that in the course of phrase production a 'counter' is keeping track of the number of syllables, and pause duration increases as the number on the counter grows. While this is a grossly oversimplified description, some measure of quantifying elapsed structure or elapsed time seems a plausible explanation of length effect. A second possibility is that information load influences pause duration, in that a longer phrase is likely to have a larger information load. The longer pause would reflect the time made available to the listener for processing the preceding information. The same explanation could be available for a speaker-driven account for the preboundary length effects. Namely, the speaker may be consolidating or de-activating processed information. At this point, it is an open question as to whether the pauses are speaker or listener driven, and the experiment results do not allow for any distinction to be made between the possible causes of this effect.

5.1.2. Postboundary effects

Postboundary effects are most sensibly understood as speaker-driven since the listener has not yet been presented with this material for processing. One thing to note is that the results for the prosodic complexity experiment show that speakers can have a very large lookahead, namely one IP. This is why an effect of prosodic complexity is observed.[3]

Three factors affecting pause duration were discussed: information load, structural linguistic factors, and cognitive factors. As mentioned, information load is not a postboundary factor in pause duration. This leaves the option of the planning buffer and of linguistic structure for understanding the postboundary effects. In fact, it is suggested below that the two interact in an interesting way.

Looking at the results for the prosodic complexity condition, the question arises as to why a more complex structure (i.e., the branching structure) is correlated with a *shorter* pause. Studies on the effects of *syntactic* structure on pause duration show that more complex structures lead to longer planning pauses (e.g., Ferreira, 1991; Strangert, 1997). Ferreira (1991) finds that the pause preceding a VP is significantly longer depending on the complexity of the VP, but crucially that there is an interaction effect, and that the complexity of the VP has an effect only when the subject is complex too. Ferreira concludes that the pause between the subject NP and the VP is the time when the speaker can pause to plan the following unit, and that when the subject NP is not complex, the planning of the VP has been done during the executing of the NP, so there is no need for a longer pause, but when both the subject and the VP are complex, the pause duration increases, as the planning of the

[3]Such a large lookahead is interesting for the question of incrementality in speech production. Incrementality requires that as soon as a chunk of information is available at one level of processing, this will trigger the processing at the next level (e.g., Levelt, 1989; Levelt, Roelofs, & Meyer, 1999; Ferreira & Swets, 2002; Keating & Shattuck-Hufnagel, 2002). In terms of articulation for example, incrementality means that the speaker can start generating the articulatory plan before knowing the syntactic structure of the entire utterance. While incrementality is assumed as a general principle in speech production, the important question remains as to how big a chunk of information triggers processing at the next level. The results of the prosodic complexity experiment suggest that speakers, at the time of initiation of the postboundary phrase, were aware of the postboundary branching, therefore, at least on some level, had a lookahead of one IP. So the increment triggering articulation was at least one IP.

VP is still in progress at that time. Ferreira concludes that phonological encoding is done in chunks—structurally defined units—and that the size of these units is limited by the capacity of a short-term memory buffer.

A similar account is possible for the results of the present study, making use of Ferreira's (1991) idea that speech is encoded over structurally defined units. In the case of postboundary branching, speakers may be planning up to the branching node only, and then start to speak. In the case of nonbranching, they have planned the complete upcoming phrase. The effect of the branching is that the unit to be planned is smaller—just the first following IP. The second IP is then prepared during the production of the first IP. The prosodic structure determines what amount of information will be planned at one time, and depending on the size of that information (i.e., the whole postboundary phrase or just one phrasal component of that larger phrase), the pause will be longer or shorter.[4]

The postboundary effect seen for phrase length can also be explained by the speaker planning her speech—again, the longer the unit to be processed, the longer the pause. In this case, the relevant structure appears to be the relative number of syllables within a phrase (or possibly some other indicator of phonological metrical structure). Critically, all phrases in Experiment I were nonbranching IPs, so there could not have been an effect of prosodic branching. If this phonological structure were to have the same effect as the prosodic structure, there should not be a length effect, as the processing unit in all the length conditions would always be the same—a syllable. An alternative is that the postboundary length effect is due to each syllable adding pause time for phonological and phonetic encoding. Pause duration then depends on the number of structural units to be processed. So even though this is a structural effect, it is a very different effect than the prosodic complexity effect. Prosodic structure determines the size of the chunk to be processed, whereas phonological structure (length) determines how long the encoding time for that chunk will be. This difference explains why the prosodic and phonological effects on preceding pause duration are different—namely more phonological structure leads to longer pauses whereas more complex prosodic structure occurs with shorter pauses.

Summarizing, this section argued that preboundary effects are driven by information processing purposes or are due to linguistic structure. It was suggested that postboundary effects are driven by the speaker's planning of the upcoming phrase. The postboundary prosodic complexity effects were analyzed as resulting from cognitive factors (e.g., the possibility of buffering) interacting with prosodic structure, and the postboundary length effects as resulting from cognitive factors interacting with phonological structure.

6. Conclusions

Two experiments testing the effects of prosodic complexity and IP length before and after pauses were conducted. The results showed that both length and prosodic structure influence pause duration. The length effects are symmetrical, while the effects of prosodic structure are asymmetrical (postboundary), with a lookahead of one IP. Possible explanations for pause length effects were examined. It was argued that preboundary effects are due either to linguistic structure or to information load. The postboundary effects on pause length were explained by linguistic structure and cognitive factors. It was further suggested that phonological and prosodic structure have different effects, with upcoming phonological structure driving the time needed for encoding and with upcoming prosodic structure determining the constituent unit for which that encoding is taking place.

Acknowledgments

This research was supported by NIH DC 03172 (PI Dani Byrd) and the USC Phonetics Laboratory. I am very grateful to Sun-Ah Jun, the editor and an anonymous reviewer for very helpful comments and to Sungbok Lee for helping with the recording. I am especially grateful to Dani Byrd for helping with all parts of

[4]Note that the idea that speech planning and execution operate in parallel is commonly assumed as a principle of the speech production process (e.g., Levelt, 1989; Ferreira, 1991; Ferreira & Swets, 2002; Griffin, 2003; Howell, 2004). It is a further instance of incrementality in speech production.

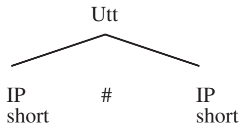
the project. Parts of this work have been presented in ''Prosodic Complexity and Phrase Length as Factors in Pause Duration'', at the Acoustical Society of America meeting, New York, New York, May 2004.

Appendix A

Materials for Experiment I (Length)

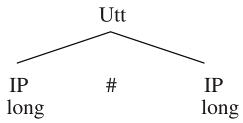
(# indicates the measured pause)

I short # short



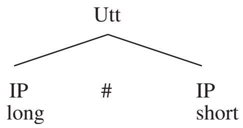
(1) John wept; Ann talked.
(2) John sat; Ann stood.
(3) John went; Ann came.

II long # long



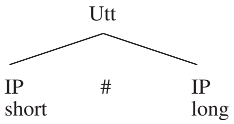
(4) Jonathan was slowly reading an unbelievably boring book about beautifully and strangely colored carpets in various parts of the state; Elizabeth-Ann Whittiker was meticulously drawing a colorful picture of incredibly fast and rather weird looking animals.
(5) Jonathan was desperately trying to finish the java program for playing an incredibly fascinating fantasy game with a hobbit; Elizabeth Whittiker was slowly reading a very strange and unbelievably funny book about auto-racing and extraterrestrials.
(6) The older students were barely paying attention to the professor's lamenting about the museum's tragic loss of the strangely colored carpet; all the younger students were meticulously and studiously taking notes for the final exam on carpet fashion in the nineteenth century.

III long # short



(7) The room above Ann was unfortunately immediately occupied by an extremely boisterous and untalented heavy-metal rock group; Ann moved.
(8) Nobody in the neighborhood seemed to really appreciate their incredibly loud and unbelievably persistent practicing in the attic; Abe moved.
(9) She was known for her ability to unintentionally upset even the most loyal hardworking people in that group; Abe left.

IV short # long



(10) John ate; Elizabeth-Ann Whittiker was visibly impressed and delighted with his strange and exotically delicious chocolate fritter appetizers.
(11) John slept; Elizabeth-Mary Whittiker was energetically and successfully ruining his masterfully sculpted symmetrical beach sand castles.
(12) Marc slept; Elizabeth was absorbed in drawing a surprisingly accurate picture of fantasy-creatures from her favorite Scandinavian fairy tale.

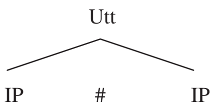
Appendix B

Materials for Experiment II (prosodic complexity)

In the sentences below, in order to elicit prosodic branching and nonbranching, the syntactic structure and the length of syntactic constituents were manipulated (while the length of the pre and postboundary phrases and the point of branching were kept constant across all sentences). The criteria used are based on results from the literature showing that speakers tend to place a prosodic boundary after preposed clauses and in the middle of sentences, and indications that pauses are likely after long subjects (Cooper & Paccia-Cooper, 1980; Ferreira, 1991; Grosjean et al., 1979; Selkirk, 1995; Strangert, 1997). In addition to these elements, inverted subject phrases were also used. Each sentence had 28 syllables, and the branching was targeted at the 14th syllable. In this way, the tendency to break was enhanced by targeting the branching in the middle of the pre or postboundary phrase. Finally, punctuation marks were not used, in order not to influence the strength of the produced boundary. Below, after each utterance, if there was branching, the elements inducing the prosodic branching are mentioned: 'long subject' means that the subject phrase is long, and that the following syntactic constituent is of the same length; 'inverted subject' means that the subject is inverted and 'preposed adjunct clause' means that the adjunct clause is preposed. For the nonbranching stimuli the sentences consisted of a number of small syntactic constituents, so that according to the above mentioned criteria nonbranching IPs phrases would be elicited. The actual production of the intended prosodic structures was verified using the ToBI criteria for prosodic transcription (Beckman & Elam, 1997).

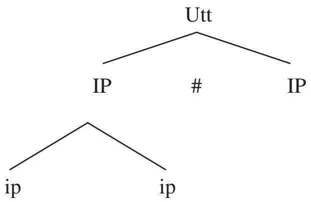
(# indicates the measured pause)

I nonbranching # nonbranching



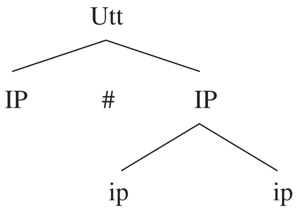
(1) The room was filled with exotically shaped and wonderfully colored large cushions for sitting on the pale-green mat; all the large windows were covered with very heavy and meticulously embroidered decorative curtains.
(2) The water pool was surrounded by a large number of tall growing and long living sugarpines from New Brunswick; all the children were often enchantedly looking at the symmetrically shaped and beautifully smelling cones.
(3) The tall electrician did not really appreciate the extremely loudly barking and aggressive looking dogs in the back lot; and the Dobermans did not really seem to mind the noisy and light generating electrician all that much.

II branching # nonbranching



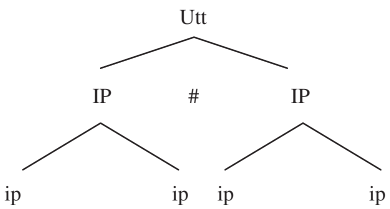
(4) The friendly smiling girl from the top floor of the building was listening to John's discussion with the assistant; all the other people were excitedly talking about green peppers and fancy spices from distant countries [preboundary branching: long subject].
(5) Standing almost invisibly in the furthest corner was a very tall woman holding a yellowish sack; all the other people were attentively listening to the rather vehement discussion about peppers [preboundary branching: inverted subject].
(6) In the middle of the beautifully and meticulously carved table were various tasty dishes from all parts of the state; a small girl was hungrily staring at all the exotically smelling and wonderfully looking fancy food [preboundary branching: inverted subject].

III nonbranching # branching



(7) The stylish woman was continuously and excitedly talking about a lovely orange-red carpet; accompanying her very energetic story were various examples of fabulous new carpets [postboundary branching: inverted subject].
(8) The smaller children were enthusiastically exploring the cavelike room in a wildly disorganized route; a huge amount of wonderfully tasting Dutch candy was hidden in various dark corners in the large room [postboundary branching: long subject].
(9) The greenhouse garden was completely covered with radiantly colored flowers from various parts of the state; an exotically smelling vanilla bean orchid was dividing this extravagant garden into two parts [postboundary branching: long subject].

IV branching # branching



(10) Looking at the funny shaped and oddly colored furniture Johnny wondered about the strange taste of the tenant; admiring the beautiful and comfortable chairs Ann thought of her own pale and uncomfortable sofa [preboundary branching: preposed adjunct clause # postboundary branching: preposed adjunct clause].

(11) The curious and polite people from the large pink room had never seen the maple leaf-shaped chairs in the attic; all the pleasantly smiling people from the orange room had always liked the maple leaf-shaped chairs in the attic [preboundary branching: long subject # postboundary branching: long subject].

(12) In the middle of the oddly decorated bedroom was a somewhat elaborately carved brown cabinet; on the right side of the green painted and very clean desk was an unusually tightly packed wooden bookshelf [preboundary branching: inverted subject # postboundary branching: inverted subject].

References

Bannert, R., Botinis, A., Gawronska, B., Katsika, A., & Sandblom, E. (2003). Discourse structure and prosodic correlates. In: *Proceedings of the XVth international congress of phonetic sciences*, Barcelona, Spain, August 3–9, 2003 (pp. 1229–1232). Adelaide: Causal Productions.

Beckman, M. E. (1989). Notes on intonation. In M. E. Beckman (Ed.), *Course reader of Ling 694B (Intonation), Winter 1989*. Department of Linguistics, Ohio State University.

Beckman, M., & Elam, G. (1997). Guidelines for ToBI labeling. (version 3.0). Unpublished ms (available online at: http://www.ling.ohio-state.edu/~tobi/ame_tobi/labelling_guide_v3.pdf).

Beckman, M. E., & Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, *3*, 255–309.

Butcher, A. (1981). *Aspects of the speech pause: Phonetic correlates and communicative functions*. Kiel: Arbeitsberichte, 15. Institut fur Phonetik.

Choi, J.-Y. (2003). Pause length and speech rate as durational cues for prosody markers. *Journal of the Acoustical Society of America*, *114*(4), 2395.

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

Cummins, F. (2002). On synchronous speech. *Acoustic Research Letter Online*, *3*, 7–11.

Cummins, F. (2003). Practice and performance in speech produced synchronously. *Journal of Phonetics*, *31*, 139–148.

Dilley, L., Shattuck-Hufnagel, S., & Ostendorf, M. (1996). Glottalization of vowel-initial syllables as a function of prosodic structure. *Journal of Phonetics*, *24*, 423–444.

Fant, G., Kruckenberg, A., & Ferreira, J. B. (2003). Individual variations in pausing. A study of read speech. In Proceedings of Fonetik 2003, Reports in Phonetics, Umeå University, PHONUM, Vol. 9 (pp. 193–196).

Ferreira, F. (1991). Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language*, *30*, 210–233.

Ferreira, F. (1993). Creation of prosody during sentence production. *Psychological Review*, *100*(2), 233–253.

Ferreira, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, *46*, 57–84.

Fletcher, J. (1987). Some micro and macro effects of tempo change on timing in French. *Linguistics*, *25*, 951–967.

Gee, J. P., & Grosjean, F. (1983). Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology*, *15*, 411–458.

Goldman Eisler, F. (1968). *Psycholinguistics. Experiments in spontaneous speech*. London and New York: Academic Press.

Griffin, Z. (2003). A reversed word length effect in coordinating the preparation and articulation of words in speaking. *Psychonomic Bulletin and Review*, *10*, 603–609.

Grosjean, F., Grosjean, L., & Lane, H. (1979). The patterns of silence: Performance structures in sentence production. *Cognitive Psychology*, *11*, 58–81.

Horne, M., Strangert, E., & Heldner, M. (1995). Prosodic boundary strength in Swedish: Final lengthening and silent interval duration. In *Proceedings of the XIIIth international congress of phonetic sciences*, Stockholm, Sweden, August 13–19, 1995 (Vol. 1, pp. 170–173). Stockholm: KTH and Stockholm University.

Howell, P. (2004). Assessment of some contemporary theories of stuttering that apply to spontaneous speech. *Contemporary Issues in Communication Science and Disorders: CICSD*, *31*, 122–139.

Keating, P., & Shattuck-Hufnagel, S. (2002). A prosodic view of word form encoding for speech production. *UCLA Working Papers in Phonetics*, *101*, 112–156.

Ladd, R. (1988). Declination "reset" and the hierarchical organization of utterances. *Journal of the Acoustical Society of America*, *84*(2), 530–544.

Ladd, R. (1996). *Intonational phonology, Cambridge Studies in Linguistics, Vol. 79*. Cambridge: Cambridge University Press.

Lane, H., & Grosjean, F. (1973). Perception of reading rate by listeners and speakers. *Journal of Experimental Psychology*, *97*(2), 141–147.

Levelt, W. J. M. (1989). *Speaking. From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Brain and Behavioral Sciences*, *22*(1), 1–38.

Martin, J. (1970). On judging pauses in spontaneous speech. *Journal of Verbal Learning and Verbal Behavior*, *9*, 75–78.

Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Dordrecht, Holland/Riverton, USA: Foris Publications.

Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation*. Ph.D. Dissertation, MIT, published 1987 by Indiana University Linguistic Circle.

Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America*, *90*(6), 2956–2970.

Sanderman, A. A., & Collier, R. (1995). Prosodic phrasing at the sentence level. In F. Bell-Berti, & L. J. Raphael (Eds.), *Producing speech: Contemporary issues. For Katherine Safford Harris* (pp. 321–332). New York: American Institute of Physics.

Selkirk, E. (1981). On prosodic structure and its relation to syntactic structure. In T. Fretheim (Ed.), *Nordic Prosody II* (pp. 111–140). Trondheim: Tapir.

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

Selkirk, E. (1995). Sentence prosody: Intonation, stress, and phrasing. In J. A. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 550–569). Cambridge, MA, and Oxford, UK: Blackwell.

Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, *25*(2), 193–247.

Smith, C. (2004). Topic transitions and durational prosody in reading aloud: Production and modeling. *Speech Communication*, *42*, 247–270.

Sternberg, S., Monsell, S., Knoll, R. L., & Wright, C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G. E. Stelmach (Ed.), *Information processing in motor control and learning* (pp. 117–152). New York: Academic Press.

Strangert, E. (1991). Pausing in texts read aloud. In , *Proceedings of the XIIth international congress of phonetic sciences, Aix-en-Provence, France, August 19–24, 1991*, Vol. 4 (pp. 238–241). Université de Provence, Service des Publications.

Strangert, E. (1997). Relating prosody to syntax: Boundary signaling in Swedish. In: Kokkinakis, G. Fakotakis, N., & Dermatas, E. (Eds.), *Proceedings of the fifth European conference on speech communication and technology*, Rhodes, Greece, Vol. 1 (pp. 239–242).

Swerts, M., & Geluykens, R. (1994). Prosody as a marker of information flow in spoken discourse. *Language and Speech*, *37*, 21–43.

Trouvain, J., & Grice, M. (1999). The effect of tempo on prosodic structure. In *Proceedings of the XIVth international congress of phonetic sciences*, San Francisco, CA, August 1–7, 1999 (Vol. 2, pp. 1067–1070). Martinez: East Bay Institute for Research and Education, Inc.

Watson, D., & Gibson, E. (2004). The relationship between intonational phrasing and syntactic structure in language production. *Language and Cognitive Processes*, *19*(6), 713–755.

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal Acoustical Society of America*, *91*(3), 1707–1717.

Zvonik, E., & Cummins, F. (2002). Pause duration and variability in read texts. In *Proceedings of the 2002 international conference on spoken language processing (ICSLP '02),* Denver, CO (pp. 1109–1112).

Zvonik, E., & Cummins, F. (2003). The effect of surrounding phrase lengths on pause duration. In *Proceedings of Eurospeech 2003*, Geneva, Switzerland (pp. 777–780).