# Length, Ordering Preference and Intonational Phrasing: Evidence from Pauses

*Gerrit Kentner*

Institut für Linguistik, Universität Potsdam, Germany
gerrit@ling.uni-potsdam.de

## Abstract

This paper reports a speech production experiment in which the effects of surrounding phrase lengths and head-argument distance on intra-sentential pause duration were tested. While the results confirm an effect of phrase length on pausing, this effect is found to be distinctly stronger for long phrases preceding the pause than for long upcoming phrases. The results are discussed with respect to intonational phrasing tendencies and ordering preferences for unequal-sized constituents.

**Index Terms:** intonational phrasing, pausing, phrase length, phrasal ordering preferences, speech production.

## 1. Introduction

This study is concerned with intra-sentential pauses in read speech. The purpose of this study is to ascertain the effects of the length of preceding and upcoming material and its syntactic structure on pause duration at a given position within a sentence.

Pauses are an important feature of intonational phrase (IP) boundaries in speech [1, 2]. They are, however, a highly variable phenomenon governed by numerous factors such as speaking style, speaking rate, and the speaker's emphasis. One factor determining pause distribution and pause duration that has attracted the interest of several researchers [3, 4, 5] is the length of material preceding or following a pause. As Watson and Gibson [4] note, intonational phrase boundaries, and correspondingly pauses, can serve as refractory periods which are needed after the production of particularly long constituents. Similarly, a long upcoming constituent requires planning time which in turn might induce a pause. Watson and Gibson [4] formulate the LRB algorithm to predict IP boundaries at a given position within a sentence. It is mainly based on the number of words preceding and following that point such that the probability of an IP boundary rises when the preceding and/ or upcoming constituent is long. It is further constrained by syntactic factors. It has been shown that the likelihood of a pause and pause duration at a given position in a sentence increase with the number of syntactic brackets associated with that position [6]. That is, speakers preferably place phrase boundaries between rather than within syntactic constituents, i.e. speakers intonationally wrap the constituents [7].

A further constraint on IP boundary placement is again related to prosody: If possible, speakers concatenate IPs of approxiamtely equal size [8]. If the syntactic constituents to be concatenated in production have greatly differing lengths, the above constraints on intonational boundary placement conflict. In such a situation, speakers of English and German tend use the order short before long [9, 10]. The preference to place long IPs at the end of utterances can be seen as a prosodic constraint that might override syntactic requirements on constituent order, as evidenced by the phenomenon of Heavy NP Shift [11]. However, speakers do not always have the choice to order the constituents according to the needs of such prosodic constraints and might be forced to utter long constituents first.

Two recent studies on pause duration, namely [3] and [5] have scrutinised inter-sentential pause duration as an effect of preceding and following phrase lengths. In their study, Zvonik and Cummins [5] used synchronously read speech. The authors report that inter-sentential pauses shorter than 300ms almost exclusively occur when the preceding and following phrase consists of 10 syllables or less. The probability of a pause being short was shown to rise greatly if both the preceding and following phrase contained only 10 or fewer syllables, suggesting that the two predictors act superadditively.

Krivokapic [3] also used the method of synchronous reading. In her study on inter-sentential pauses, she compared pause length in four conditions, namely short/short, short/long, long/short, and long/long. Krivokapic found significant length effects for both preboundary and postboundary phrases, indicating that, irrespective of the order of the phrases, long phrases induce increased pause duration relative to short ones.

In a recent article, Watson and Gibson [12] have tested the hypothesis that the likelihood for an IP boundary increases with the integration distance between heads and their arguments. In their production experiment, however, they could not find a significant effect.

The present experiment is designed to ascertain the effect of the respective ordering of a long and a short constituent on pause duration between the constituents within German sentences. The syntactic structure of the experimental sentences is systematically varied in order to disentangle the different effects of syntax (i.e. head-argument distance) and constituent length on pause duration.

## 2. Methods

### 2.1. Experimental design

A speech production experiment is designed to test the influence of three factors on pause duration in speech production. These are 1) the respective ordering of a short and a long constituent, 2) the direction of the lexical head of the long constituent and 3) the position of the main verb within the sentence. With factors 2) and 3), the integration distance between syntactic heads and their arguments is systematically varied. All factors are two sided and crossed for this experiment.

Reading material is constructed according to the above factors resulting in 8 conditions. Sentences including coordinations are chosen as a test bed for this experiment. These constructions contain a proper name as subject in first position followed by a transitive verb frame with two coordinated objects. One of the object NPs is relatively long (10-15 syllables) and the other short (2-3 syllables). The

lexical head of the long object is either preceded by a modifying adjective phrase (head right) or followed by a modifying PP (head left). The transitive verb frame either features with the main verb in second position followed by the objects or the main verb appears sentence finally and a modal verb occupies the second position. The respective structures are exemplified in (1).

(1)

a. [Subj] [Verb] [Obj1] [&] [Obj2]
b. [Subj] [Mod] [Obj1] [&] [Obj2] [Verb]

The stimuli for the production experiment consist of 24 sets of 8 sentences each. The experimental sentences were allocated to the subjects in a latin square design such that each subject would see only one sentence from each of the 24 sets. That way, each subject was presented 24 sentences, three from each condition. The 24 experimental sentences were embedded in 54 filler sentences. This set of 78 items was fed into a DMDX presentation [13] and pseudo-randomised for each subject such that sentences of the same condition did not appear adjacently.

In (2), the eight conditions are shown. The range specified in the parentheses refers to the relative position of the conjunction in that condition. It is calculated dividing the number of syllables preceding the conjunction by the total number of syllables in that sentence. As can be seen, the range for short-long sentences is distinctly below 0.5 while the range for long-short versions is higher than 0.5. The total length of the sentences ranges from 18-24 syllables.

(2)

a. short-long, V2, Head left (.25-.41)
b. short-long, V2, Head right (.25-.41)
c. short-long, V-end, Head left (.23-.33)
d. short-long, V-end, Head right (.22-.31)
e. long-short, V2, Head left (.75-.85)
f. long-short, V2, Head right (.75-.86)
g. long-short, V-end, Head left (.61-.77)
h. long-short, V-end, Head right (.63-.77)

In (3) the first 4 conditions are exemplified. Conditions e-h are made up of the same material but the order of the underlined objects is reversed.

(3)

a.
Paul malt den Fluss und das Haus von Melanies Großtante.
*Paul paints the river and the house of Melanie's grand aunt*

b.
Paul malt den Fluss und das winzige geklinkerte Gartenhaus.
*Paul paints the river and the tiny clinker-bricked summer house*

c.
Paul will den Fluss und das Haus von Melanies Großtante malen.
*Paul wants to paint the river and the house of Melanie's grand aunt*

d.
Paul will den Fluss und das winzige geklinkerte Gartenhaus malen.
*Paul wants to paint the river and the tiny clinker-bricked summer house*

### 2.2. Subjects

16 undergraduate students (7 male, 9 female) from the University of Potsdam took part in the experiment. All are native speakers of German and naïve to the purpose of the experiment. They either received course credit or were paid.

### 2.3. Recordings

Recordings took place in an acoustically shielded room with an AT4033a audio-technica studio microphone. Each subject was seated in front of a 15'' computer screen with the microphone placed approximately 30cm from the subject's mouth. A keyboard was placed on a table within close reach of the subject. Recordings were made on a computer using the RecordVocal function of DMDX and a C-Media Wave soundcard at a sampling rate of 44.1 kHz with 16 bit resolution.

### 2.4. Procedure

After a short instruction and three practice items (not part of the experimental set) the first sentence was presented on the screen. In order to enhance reading fluency, subjects were asked to familiarise themselves with the sentence and to press the space bar key afterwards. On pressing the space bar, the screen blanked for 200ms until the sentence reappeared on the screen. At this point, the subject's task was to read the sentence aloud. After that, a new sentence appeared and the procedure was repeated. For each sentence, there was only one realisation by subject. No corrections were recorded in the case of hesitations or slips of the tongue.

### 2.5. Data analysis

The data of the 16 subjects contains numerous slips of the tongue or hesitations due to self corrections (8.6%); the affected sentences were discarded. Overall, 332 sentences were manually annotated using the TextGrid device of Praat acoustic speech analysis software [14]. Duration analyses of the sentences were carried out automatically with Praat scripts. The durations of the two conjuncts and the duration of the silent interval before the conjunction were measured.

## 3. Results

Figure 1 shows the duration of the silent interval between the conjuncts as a function of the ordering of the long and the short conjunct.

The mean pause duration is 87ms for long-short sentences and 51ms for short-long sentences. A linear mixed effects model [15] with the crossed fixed factors "ordering" (short-long vs. long-short), "headedness" and "position of main verb" was employed; "subjects" and "sentence" were included as random effects. The logarithm of pause duration was chosen as the dependent variable. This model yields a significant main effect for "ordering" on logarithmised pause duration ($t=-2.11$, $df=324$, $p=0.036$). All other main effects and interactions remain non-significant.

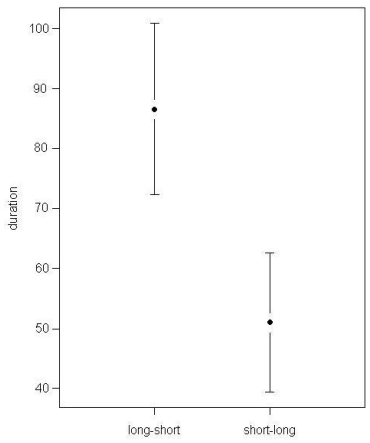


Figure 1: *Mean pause duration (in ms, with 95% confidence interval) at the conjunction as a function of the ordering of a long and a short conjunct.*

## 4. Discussion

These data confirm that surrounding phrase lengths have an effect on pause duration. Speakers pause longer at the conjunction of two unequal-sized conjuncts when the longer constituent precedes the shorter one. This result suggests that, in these asymmetric coordinations, the duration of the pause is positively correlated with the size of the preceding phrase but not with the size of the upcoming one.

Pauses are among the defining features of IP boundaries. The likelihood and strength of an intonational boundary grows with increasing pause length [1, 2]. The present evaluation of the experiment remains agnostic as to whether the silent intervals coincide with other IP boundary cues such as phrase-final lengthening or boundary tones. A closer examination of these cues would certainly be adequate to verify the results. However, given that items involving hesitation pauses were discarded, it is unlikely that the sentences under scrutiny contain pauses that interrupt intonational phrases. Since the intervals were measured at major constituent boundaries, namely at the conjunction, the dependent variable can be considered a good measure for intonational boundary strength. Therefore, it can be inferred from the result that the boundary strength and thus the likelihood of an IP boundary at the conjunction is higher in sentences with long-short ordering of the constituents than in short-long versions.

The findings of the present experiment complement and qualify Watson and Gibson's [4] LRB algorithm on IP boundary placement since only an effect of the length of the preceding phrase but not of the upcoming one can be confirmed. This is not to contest the results of Ferreira [16] and others who find that the size of an upcoming constituent is a predictor for pause length. However, it follows from the results here that the size of the preceding phrase is a stronger predictor for IP boundary placement. Watson and Gibson [4] themselves hypothesise that their LRB algorithm might be more successful when the relative influence of the upcoming phrase on boundary placement would be more restricted. This corresponds well with the notion of incrementality in the speech production process: A speaker does not always complete the planning of a constituent before he starts uttering it. Therefore, its ultimate size cannot be determined in advance and thus its influence on pause duration is limited.

The outcome of the experiment is especially interesting against the background of the short-long preference for constituent ordering in German and English [9, 10]. It seems that the violation of this preference has an effect on intonational phrasing. That is, while sentences which obey the preferred constituent order do not show a strong prosodic break, the long-short order tends to result in a more complex prosodic structure with an IP boundary between the unequal-sized constituents (as substantiated by the relatively longer pauses in this condition). A possible interpretation of this is that, when forced to utter the unpreferred order, speakers avoid a violation of the short-long preference on IP level by inserting an IP boundary after the long constituent. Thus, the IP can be seen as a domain for the short-long ordering preference.

Given the preference for long-short ordering of constituents in Japanese [17], it would be interesting to set up a similar experiment in that language to compare the effects of pause duration on preferred vs. dispreferred constituent order.

## 5. Acknowledgements

This research was financed by a grant from the German Research Foundation (DFG priority programme no. 1234) to the project "Prosody in Parsing" at the University of Potsdam. I would like to thank Caroline Féry, Frank Kügler, Anja Mietz, Umesh Patil, Shravan Vasishth, and three anonymous reviewers for helpful comments and suggestions.

## 6. References

[1] Cruttenden, A., Intonation, Cambridge University Press, Cambridge, 1997.

[2] Price, P.J., Ostendorf, S., Shattuck-Hufnagel, S., Fong, C., "The use of prosody in syntactic disambiguation", J. Acoust. Soc. Am. 90 (6), 1991, 2956-2970.

[3] Krivokapic, J., "Prosodic planning: Effects of phrasal length and complexity on pause duration", Journal of Phonetics, article in press.

[4] Watson, D., Gibson, E. "The relationship between intonational phrasing and syntactic structure in language production", Language and cognitive processes, 19, 2004, 713-755.

[5] Zvonik, E., Cummins, F., "The effect of surrounding phrase lengths on pause duration", Paper presented at Eurospeech 2003

[6] Gee, J.P., Grosjean, F., "Performance structures: A psycholinguistic and linguistic appraisal", Cognitive psychology, 15, 1983, 411-458.

[7] Truckenbrodt, H., "On the relation between syntactic phrases and phonological phrases", Linguistic Inquiry, 30, 1999, 219-255.

[8] Ferreira. F., Prosody, In: Nadel, L. (Ed.). Encyclopedia of cognitive science, Macmillan Reference, London, 2002.

[9] Wasow, T., "End-weight from the speaker's perspective", Journal of Psycholinguistic Research, 26, 1997, 347-361.

[10] Behaghel, O., "Von deutscher Wortstellung", Zeitschrift für Deutschkunde, 44, 1930, 81-89.

[11] Jackendoff, R.., "Foundations of Language". Oxford University Press, 2002.

[12] Watson, D., Gibson, E., “Intonational phrasing and constituency in language production and comprehension”, Studia linguistica, 59, 2005, 279-300
[13] Forster, K., Forster, J., “DMDX: A windows display program with millisecond accuracy”, Behavior Research Methods, Instruments and Computers 35, 2003, 116-124
[14] Boersma, P., Weenink, D. 2005. Praat: doing phonetics by computer. http://www.praat.org/ visited 5-Feb-07.
[15] Pinheiro, J., Bates, D., Mixed-Effects Model in S and S-PLUS. Springer, New York, 2000.
[16] Ferreira, F., “Effects of length and syntactic complexity on initiation times for prepared utterances”, Journal of Memory and Language, 30, 1991, 210-233.
[17] Yamashita, H., Chang, F., ”Long before short preference in the production of a head-final language”, Cognition, 2001, B45-B55.