

Virtual and Real Pauses at Clause and Sentence Boundaries

Nina B.Volskaya

St.Petersburg State University, Department of Phonetics,
11, Universitetskaya emb., St.Petersburg, Russia, 199034
e-mail: nina@PS1089.spb.edu

ABSTRACT

This study investigates pause duration at the boundaries between syntactic units in read aloud speech in Russian. The results show predictable relationship between the type of the syntactic phrase and pause duration. At the same time, data obtained in the study indicate that among the acoustic determinants of and cues to the perception of a phrase boundary pause duration is not the leading factor, since for a considerable number of syntactic units pauses have no correlates in acoustic silence. The results of the investigation can provide reliable data for accurate modelling of pause duration at prosodic boundaries of phrases, clauses and sentences in the read text.

1. INTRODUCTION

Pause placement and duration show the speaker's strategy in structuring the text, helping the listener to interpret it accordingly. Pausing strategy reveals the syntactic structure of the text. Many studies confirm a relationship between prosody and syntax, at the same time claiming that prosodic and syntactic units do not necessarily coincide. "Intonation-groups usually align with relatively large syntactic structures, the most typical correspondence being the co-extensiveness between intonation-groups and clauses (the term clause is defined as a grammatically unified group of words which contains a subject and a finite verb)"[1:98]. This normal alignment Halliday calls "neutral tonality" [2:32]. In the data reported by Quirk et al., half of the clauses were co-extensive with intonation groups [3].

A division of the text into intonation units and assigning appropriate pause duration is one of the first and most important steps in generating intelligible and natural sounding synthetic speech. A reliable quantitative model for pause duration still remains the aim of intensive research.

A pause is an important member of prosodic cues used among other components — pitch changes, pre-boundary lengthening, declination reset — for boundary marking and thus for structuring the text into intonation units. Speakers use variable combination of prosodic cues to signal the end of one unit and the beginning of the other. Which of them are of primary importance, how the speaker marks boundaries of different strength and

highlights the structural make up of the sentences in the text — these issues have been the subject of intensive research for quite a long time. Studies have shown that prosodic unit boundaries are characterised by internal and external cues. Each intonation unit usually contains one pitch prominence, the nucleus. The occurrence of the nucleus is the internal criterion for the demarcation of the intonation units. Among external cues — pauses of various length [4], [5], [6], [7:36-41], [8:171,205-206]; final syllable lengthening [5], [6], [9], [10], [11]; pitch changes and declination reset — changes in the relative height of F0 level at the boundaries of two units [12], [13], [14].

The development of a quantitative model for pause duration is very complicated due to its great variability caused by many factors, such as speech rate, speaking style, syntactic unit length, etc.

Some studies which attempted to relate prosody to syntax, found that pause duration depends on the complexity of the syntactic phrases and the length of the words, which precede the boundary. The longer the syntactically complex phrases or sentences, the longer the pauses [15], [16]. At the same time, empirical studies show that the vast majority of intonation-groups have lengths which range between one and seven words [8:256].

The present study investigated pause duration as one of the prosodic cues used in marking syntactic boundaries in the Russian read text.

2. SUBJECTS

5 female and 3 male native speakers of Russian participated in this study, all students of the Department of Humanities of Saint-Petersburg State University. Their age ranged from 17 to 20 years old at the time of recording.

3. MATERIAL AND PROCEDURE

A phonetically balanced text [17] was recorded individually with a Sony DAT recorder in the sound-treated room. Total recording time is about 30 minutes. The text contained syntactically simple, compound and complex sentences. It was read by each speaker in the most appropriate and natural for him or her manner. Syntactic analysis of the recorded material and intonation

transcription, in which prosodic boundaries were marked, were performed by the experimenter. Since the subjects read a prepared text, the freedom of interpretation was limited by the syntactic structure of the text. The number of syntactic units realised by each speaker, is presented in Table 1. Average number of syntactic units is 130, standard deviation 5,59. In a great majority of cases prosodic boundaries coincided with the boundaries between syntactic units. The recordings were then digitised. The beginning and the end of each intonation unit was determined at the on-set and off-set of the acoustic signal on the oscillogram. The oscillograms and the waveforms were obtained using the EDS program. In order to investigate the duration, number and position of pauses, the speech data was processed automatically. Analysis of pause duration at syntactic boundaries was carried out in several steps. First, silent intervals longer than or equal to 250 ms duration, defined as acoustic correlate for pausing, were automatically indicated on the oscillogram of the analysed signal by a mark of a selected colour. Second, the results of the automatic marking were compared with the results of manual intonation transcription. Pauses which were shorter than 250 ms, were marked on the oscillogram by hand. At the third step, the segmentation files were exported into a specifically structured text file which was treated using the EXCEL for MS Office 97. Unanimously defining pauses as silent intervals, researchers express different views as to their duration, required for a pause to be perceived as such. Thus, Fletcher, suggests 200ms [18] as an acoustic correlate for pausing; other studies select an interval of 100ms [19]; according to [20:216], a pause is noticeable if its duration is *at least* 0.3 second. In the present study we selected a silent interval equal to (or longer than) 250 ms for a pause which reflects a speaker's deliberate choice of one of the prosodic markers of a syntactic boundary.

Speaker	Number of units	Pauses <250 ms, %	Average pause duration, ms
S. 1	127	48	157
S. 2	134	48	159
S. 3	138	51	177
S. 4	127	34	188
S. 5	128	40	180
S. 6	136	39	174
S. 7	121	50	174
S. 8	130	19	179
Means	130	41	173,5

Table 1. Number of units, number of pauses < 250 ms and their average duration values (in ms)

4. ANALYSIS OF RESULTS

Table 1 presents the number of intonation units obtained as a result of intonation transcription of the recorded material, percent and average duration of pauses marked

automatically. The analysis of the data showed that pauses with duration below 250 ms constitute, on average, 41% of all pauses, observed at clause and sentence boundaries (total number of units analysed is 1041). The average duration value of all pauses below 250 ms is 173,5 ms. To get a clearer picture of pause duration variability, a more detailed study of pause duration was performed with regards to the type of the syntactic unit. At this step of analysis pause duration values were categorised. The results obtained are presented in figures 1, 2.

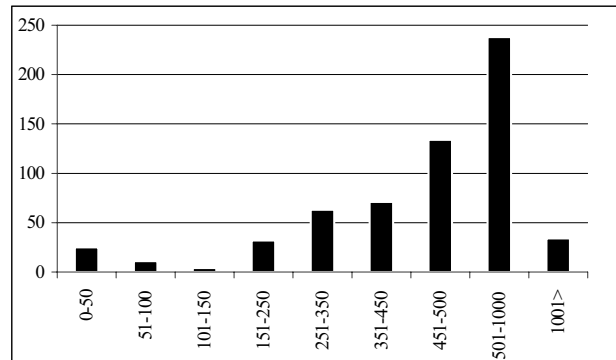


Figure 1. Number of observations for 9 pause categories (in ms) for sentence boundaries in the read text.

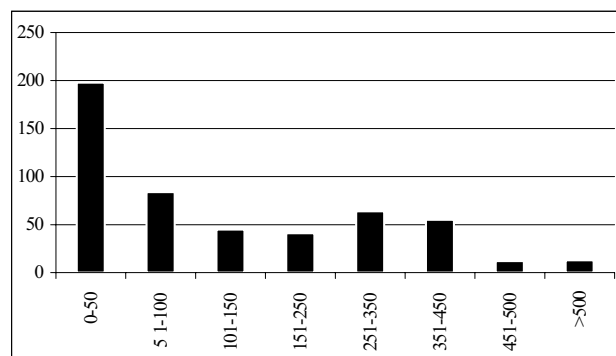


Figure 2. Number of observations for 9 pause categories (in ms) for clause and phrase boundaries in the read text

5. SENTENCE BOUNDARIES

As it follows from Fig.1, though pause duration in sentence final positions varies, the differences have a very narrow range. 88% of all pauses are over 250 ms, most of them, between 450 and 1000 ms. Pauses with duration over 1000 ms are those which mark the end of a paragraph.

6. CLAUSE AND PHRASE BOUNDARIES

71,9 % (368 pauses out of the total of 512) of all pauses found in non-sentence final position have duration values below 250 ms. Pauses with duration under 100 ms constitute 55 % of the whole number, most of them are

below 50 ms (see Fig. 2). Percentage of very short pauses, occurring at the end of non-final units (grammatically-dependent phrases and clauses) per each speaker (Table 2) is very high, with the exception for speaker 8. This data leads to the conclusion that clause boundaries in the read text are marked by a set of prosodic parameters of which a pause length is not of primary importance. Speakers leave it to the listeners and researchers to evaluate the relevance of other prosodic cues.

Speaker	S. 1	S. 2	S. 3	S. 4	S. 5	S. 6	S. 7	S. 8	Average
Percentage	94,9	80,6	84,7	58,4	63,5	73,5	87,5	34,3	71,9

Table 2. Percentage of pauses below 250 ms at clause and phrase boundaries for 8 speakers

Empirical data, obtained in some other studies [21] show, that "zero" pauses and very short pauses with duration below 200 ms at the boundaries of syntactic units in fluent reading and speech is a very common phenomenon, but their distribution and number can be "style-specific" (Table 3): the number of zero and very short pauses in the experiment described was greater for spontaneous speech (S) than for reading (R) for all the speakers, who participated in the experiment. Most of these pauses can be regarded as virtual from the point of view of the listener, since there was no discontinuity in the speech signal at the perceived boundary of the intonation unit.

Speaker	Number of units with 0 pause, %		Number of pauses <200ms, %		Total of virtual pauses, %	
	S	R	S	R	S	R
S.1	66	49	14	19	80	68
S.2	43	20	19	26	62	46
S.3	33	15	15	22	48	37
S.4	31	26	11	17	42	43
S.5	58	62	11	22	69	84
Mean	46,2	34,4	14	21,2	60,2	55,6

Table 3. Percentage of units with "zero" and very short pauses for 5 speakers (S-spontaneous, R-read speech)

Studies on pause duration demonstrate high variability of pause behaviour due to many factors which influence it. At the same time, some experimenters claim that duration of pauses may be language-specific as well [22].

7. PAUSE DURATION VS SYLLABLE DURATION

Traditionally, pause duration is described either in absolute values, milliseconds, reflecting their physical duration, or in some relative terms, like "long", "medium", "short" (see, for example, [22], where pauses are categorised as brief (<200 ms), medium (200-1000ms) and long (>1000ms)], depending on their possible duration

values and functional load in the utterance. In intonation modelling abstract descriptions are of little value, as they merely imply that pauses of the first type are longer (or shorter) than others. At the same time, analysing pause duration values along with other data, which reflect the temporal pattern of the intonation unit or a sentence, we arrive to the conclusion that some element can be selected to help us predict pause duration. In a recent study performed by Skrelin [23], average syllable duration was proposed as a measuring unit for calculating relative pause duration in the text. Table 4 presents data on average duration of syllables, intonation units and pauses, number of syllables in the intonation unit as well as their number that would potentially "fit in the pause" in the recorded material. The author comes to the conclusion that mean duration of pauses can be equal to the average syllable duration multiplied by 2 (3 or 4). Compared to the average number of syllables in the intonation unit, this data varies to a greater extent (standard deviation 0,31 vs 0,54), and would probably be more convincing if pause duration at sentence final and non-final boundaries were calculated separately.

Speaker	Syllable duration (ms)	Pause duration (ms)	Pause / syllable duration ratio	Intonation unit / pause duration ratio	Average number of syllables in the intonation unit
S.1	146	264	1,8	4,3	7,7
S.2	160	319	2,0	3,7	7,3
S.3	112	351	3,1	2,3	7,2
S.4	143	389	2,7	2,8	7,7
S.5	134	371	2,8	2,8	7,7
S.6	150	319	2,1	3,4	7,2
S.7	152	295	1,9	4,2	8,1
S.8	152	464	3,1	2,3	7,6

Table 4. Averages of duration values and their ratio in the material recorded from 8 speakers

Anyway, the average pause duration / average syllable duration ratio seems to reflect certain rhythmical tendencies and reveals the speaker's strategy in planning the temporal organisation of the text as a whole and producing pauses of required duration in particular.

8. CONCLUSION

The majority of "noticeable" silent intervals are observed at sentence boundaries of the read text. Clause and phrase boundaries are marked by other prosodic parameters. The force of a pause as a prosodic marker in demarcating syntactic units is debatable. Pauses for 41% of sentence boundaries and 71,2 % of non-final syntactic units do not have acoustic correlates and should be regarded as virtual. This result leads to the conclusion, that acoustic and perceived pauses may not coincide.

The data obtained challenges the results reported in some other studies, according to which a speaker primarily adapts the duration of pauses and secondarily chooses a particular type of melodic boundary marker. The results are more in agreement with those studies which state the priority of pitch changes in this respect. This is particularly true for non-final intonation units in Russian, since observations show that pre-boundary lengthening is not characteristic for them either, due to very fast and sharp pitch changes at the intonation centre. But experimental evidence for this claim is the aim of our future studies.

ACKNOWLEDGEMENT

The study was supported by the INTAS grant 00-00915 "Spontaneous Speech of Typologically Unrelated Languages (Russian, Finnish and Dutch): Comparison of Phonetic Properties".

REFERENCES

- [1] J.Toivanen, "Perspectives on Intonation:English, Finnish and English Spoken by Finns," in *Forum Linguisticum*, Frankfurt am Mein; Berlin; Bern; Bruxelles; New York; Oxford; Wien: Peter Lang, Band 37, p.98, 2001.
- [2] M.A.K.Halliday, "Intonation and Grammar in British English," The Hague: Mouton, p.32,1976.
- [3] R.Quirk, A.P.Duchworth, J.Swartwik, J.P.L.Rusiecki, A.J.T.Colin, "Studies in the correspondence of prosodic to grammatical features in English," in *Proceedings of the IXth Congress of Linguistics*, pp.679-691, Boston, the Hague:Mouton, p.20, 1964.
- [4] F. Goldman-Eisler, "Pauses, clauses, sentences", *Language and Speech* , 15, pp. 103-113, 1972.
- [5] L.A. Streeter, "Acoustic determinants of phrase boundary perception", *JASA*, 64 (6), pp. 1582-1592, 1978.
- [6] I. Lehiste, J.P. Olive, L.A. Streeter, "Role of duration in disambiguating syntactically ambiguous sentences," *JASA*, 60(5), pp.1192-1202, 1976.
- [7] A. Cruttenden, *Intonation*, Cambridge: Cambridge University Press, 1986.
- [8] D. Crystal, *Prosodic Systems and Intonation in English*, Cambridge, Cambridge University Press, 1976.
- [9] D.H. Klatt, "Vowel lengthening is syntactically determined in a connected discourse", *Journal of Phonetics*, 3, pp. 129-140, 1975.
- [10] I. Lehiste, "The role of temporal factors in the establishment of linguistic units and boundaries", *JASA*, 60 (5), pp. 1199-1202, 1975.
- [11] D.R. Scott, "Duration as a cue to the perception of a phrase boundary", *JASA*, 71(4), pp. 996-1007, 1982.
- [12] P. Hansson, "Perceived boundary strength," in *Proceedings of the ICSLP'02*, pp. 2277-2280, 2002.
- [13] W.E. Cooper, J.M. Sorensen, "Fundamental frequency contours at syntactic boundaries," *JASA*, 62, pp. 683-693, 1977.
- [14] J.R. de Pijper and A.A. Sanderman, "On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues," *JASA*, 96 (4), pp. 2037-2047, 1994.
- [15] E. Strangert, "Relating prosody to Syntax: Boundary Signalling in Swedish," in *Proceedings of the 5th European Conference on Speech Communication and Technology*, vol.1, pp. 239-241, 1997
- [16] E. Zvonik and F. Cummins, "Pause Duration and Variability in Read Texts," in *Proceedings of ICSLP'02*, pp.1109-1112. 2002.
- [17] Bulletin of the Phonetic Fund of the Russian Language, Supplement N3, SPb-Bohum, pp. 69-70, 1993.
- [18] J. Fletcher, "Some micro and macro effects of tempo change on timing in French," *Linguistics*, vol. 25, pp. 951-967, 1987.
- [19] B. Megyesi and S. Gustafson-Čapková, "Production and perception of pauses and their linguistic context in read and spontaneous speech in Swedish," in *Proceedings of ICSLP'02*, pp.2153-2156, 2002.
- [20] S. Schuetzhe-Coburn and E. Weber, "Units of Intonation in Discourse: a Comparison of Acoustic and Auditory Analyses," *Language and Speech*, 34, pp. 207-234, 1991.
- [21] L.Bondarko, N.Volskaya, S.Tananayko, L.Vasilyeva, "Phonetic Properties of Russian spontaneous speech," in *Proceedings of the ICPHS'03*, forthcoming.
- [22] E.Campione and J.Véronis, A large-scale multilingual study of silent pause duration, in *Proceedings of Prosody 2002*, Aix-en-Provence, 2002.
- [23] P. Skrelin, "Pause duration: in search of measuring unit," in *Problems and methods of the experimental phonetic studies*, SPb, pp. 224-229, 2002.