

Common and Language Dependent Phonetic Differences between Read and Spontaneous Speech in Russian, Finnish and Dutch

Viola de Silva[†], Antti Iivonen[§], Liya V. Bondarko⁺ and Louis C. W. Pols[#]

[†] Department of Languages, University of Jyväskylä, P.O. Box 35, FIN-40014, Finland

[§] Department of Phonetics, University of Helsinki, P.O. Box 9, FIN-00014, Finland,

⁺ Department of Phonetics, St. Petersburg University, Universitetskaya Naberezhnaya 11, St. Petersburg, Russia,

[#] Institut of Phonetic Sciences, University of Amsterdam, 338 Heregracht, 1016 CG Amsterdam, NL

E-mail: silva@cc.jyu.fi, antti.iivonen@helsinki.fi, lvbon@LB1082.spb.edu, louis.pols@hum.uva.nl

ABSTRACT

This preliminary study aims to reveal both common and language-specific phonetic differences between read and spontaneous speech in three typologically unrelated languages – Russian, Finnish, and Dutch. These languages differ in prosody, sound systems, speech styles, and means for conveying intonational meaning. Spontaneous speech was recorded from 5 to 8 speakers in each language. Transliterated extracts from the spontaneous speech recordings were read aloud by the same speakers. For Dutch, comparable speech material from the labeled IFA corpus and the transcribed Spoken Dutch Corpus was used. The two types of speech in the three languages were studied comparing their F_0 statistics, segment durations, and the number of some consonant elisions.

1. INTRODUCTION

The data presented below are based on the results of the international research project INTAS-915, that aimed at the definition and description of the common and language-dependent phonetic differences between read and spontaneous speech of three typologically unrelated languages — Russian (R), Finnish (F) and Dutch (D). These languages differ in their prosodic and sound systems. Russian is characterized by strong modifications of sounds in the speech flow (co-articulation, reduction, etc), especially where vowels are concerned. Russian stressed vowels are longer than vowels in unstressed syllables, in which they are reduced. In Finnish, the quantity of vowels and consonants is relevant not only in the stressed syllable (“double” vowels) and at the boundary of the first and second syllable (“double” consonants), but also later in words. Therefore, the reduction of Finnish speech sounds follows different rules. Vowel harmony is a typical feature of Finnish. The 18 diphthongs and frequent vowel sequences (e.g. *aioin* ‘I aimed’) increase the vocalic substance of words. According to usual interpretations, intonational

characteristics do not seem to be very prominent in Finnish. Dutch is characterized by a rich intonation system and it exhibits many co-articulation and reduction phenomena.

Changes in speech style result in changes of phonetic characteristics – segmental, suprasegmental, and intonational. Prosodic characteristics are likely to undergo changes in the first place, and under their influence, changes of phonetic characteristics of sounds (co-articulation, reduction, etc.) take place, leading to a rich inventory of speech sound variants.

Comparative phonetic studies are not uncommon. However, the present study is the first effort towards an acoustic-phonetic comparison of spontaneous and read speech styles in different languages.

2. METHODS

2.1 MATERIALS

For Finnish and Russian, 110-minute speech samples representing two speech styles were recorded from 11 native speakers (age 15 to 60 years; five females = F and six males = M).

1. Spontaneous (informal) speech: For Dutch, this style was represented by narrative speech or story telling (for example, about a vacation trip, or kids) face-to-face to an “interviewer” or to another speaker. For Russian and Finnish, the spontaneous speech samples were taken from free dialogue on a loosely defined topic with a close friend.

2. Reading: transliterated extracts from the spontaneous recording read by the same subjects. Hesitations, repetitions, incomplete words, and filled pauses were removed from the text before the recording to ensure fluent reading. The grammar of the texts was not corrected to ensure its comparability with the material recorded in

the first session.

For Dutch, comparable speech samples from the labelled IFA corpus [2] and the transcribed Spoken Dutch Corpus were used.

The results produced for this paper represent only a part of the total amount of recorded speech. 5 (2 female + 3 male) Russian speakers, 6 (3 + 3) Finnish speakers, and 7 (4 + 3) Dutch speakers were included in the analysis. The net duration of spontaneous (read) speech analyzed per speaker was 5 to 11 minutes. The project will continue its work, and the results will be expanded and further discussed later.

2.2 RECORDINGS

All recordings were performed in sound-treated or anechoic rooms, and head-mounted microphones were used.

The recorded samples of spontaneous and read speech were transferred to a computer, segmented into units of various size (sounds, syllables, intonation units) and labelled by trained phoneticians.

2.3 PROCEDURE

To compare the general differences between the pitch of spontaneous and read speech in vowels (in Dutch) or in vowels and non-obstruents (for Russian not in /r/), the average F_0 and standard deviation were measured for each speaker and each speaking style in the three languages. For Russian, F_0 analyses were conducted using the EDS and Praat programs. Before all the relevant values were taken, the pitch curves were manually checked. The other two languages used only Praat for automatic pitch analyses. For Finnish and Russian, the F_0 range was separately pre-set for each speaker, excluding creaky-voiced sections from the analysis as far as possible.

Average sound duration was measured to reflect temporal differences between read and spontaneous speech.

At the segmental (transcribed phoneme or phone) level, the following pronunciation differences between read and spontaneous speech were studied: 1) actual realisation of some very frequent endings in Dutch: 'en' -endings in plurals and verb endings, 2) elisions on final /n/ in Finnish, and 3) /j/-endings in Russian.

3. RESULTS

3.1. F_0 STATISTICS

F_0 mean values for Russian and Finnish vowels and non-obstruents as well as Dutch vowels are presented below in Table 1.

The F_0 means for Russian speakers are close to the

averages reported in some other studies (240 Hz, 256 Hz for female speakers and 120 Hz, 129 Hz for male speakers [3], [4]). The data show considerable inter-speaker differences (see R-M1, in particular) which influenced resulting semitone values of F_0 mean (Sp)- F_0 mean (R) difference, indicating that no significant or stable difference exists between spontaneous and read speech.

The Finnish data (see Table 1) show that there are individual differences of mean F_0 between spontaneous and read speech, but the same tendency can be observed in males and females: F_0 mean is higher in spontaneous speech. In semitones however, the differences are small. An especially great difference can be observed between Finnish and Russian females: Finnish values are considerably lower in both styles.

Language	Speaker(s)	Mean F_0 (Hz)		Difference	
		Spont	Read	Hz	ST
Russian (vowels and non-obstruents)	R-F1	240	247	-7	-0.5
	R-F2	250	244	6	0.4
	R-M1	109	137	-28	-4.0
	R-M2	132	121	11	1.7
	R-M3	101	100	1	0.2
	Male	114	119	5	-0.74
	Female	245	245	0	0
	Total				-0.37
Finnish (vowels and non-obstruents)	F-F1	184	174	10	0.93
	F-F2	176	179	-3	-0.26
	F-F3	194	193	1	0.03
	F-M1	105	110	-5	-0.73
	F-M2	105	103	2	0.55
	F-M3	112	106	6	0.74
	Male	114	105	9	1.14
	Female	187	179	8	0.93
	Total				1.04
Dutch (vowels)	D-F1	196	187	9	0.81
	D-F2	190	209	-19	-1.59
	D-F3	170	192	-22	-2.03
	D-F4	175	195	-20	-1.85
	D-M1	112	112	0	0.1
	D-M2	106	115	-9	-1.23
	D-M3	111	120	-9	-1.47
	Male	110	116	-4	-0.87
	Female	183	195	-12	-1.17
	Total				-1.02

Table 1. Mean F_0 (in Herz and ST=semitones) for read and spontaneous speech, and the differences between the two speaking styles. Ages of speakers Dutch: D-F1 (20), D-F2 (28), D-F3 (40), D-F4 (60); D-M1 (15), D-M2 (40), D-M3 (66), D-M4 (66). Russian: R-F1(18), R-F2 (23), R-M1 (52), R-M2 (43), R-M3 (18). Finnish: F-F1 (22), F-F2 (22), F-F3 (21), F-M1 (28), F-M2 (26), F-M3 (24).

In Dutch there is a tendency that the F_0 values are rising in read speech. The values within the same gender in Finnish and Dutch are alike.

Laryngealisation and some other factors hamper the exact F_0 statistics, and we shall return to this issue in a later paper.

3.2 SPEECH SOUND DURATION

The average duration values of all Russian and Finnish phonemes as well as of Dutch vowels are shown in Table 2.

Speaker	Spont ms	Read ms	Spont/ read, %
Russian, all phonemes			
R-F1	75	72	104
R-F2	71	64	111
R-M1	80	79	101
R-M2	77	69	111
R-M3	65	64	102
mean	73,6	69,6	105,8
Dutch, vowels			
D-F1	82	83	99
D-F2	93	106	88
D-F3	106	102	104
D-F4	103	98	105
D-M1	75	70	107
D-M2	95	90	106
D-M3	108	-	-
mean	96	93	103
Finnish, all phonemes			
F-F1	72	91	79
F-F2	67	67	100
F-F3	71	66	108
F-M1	67	78	86
F-M2	65	71	92
F-M3	68	73	93
mean	68	74	93

Table 2. Average sound duration in milliseconds and the duration ratio between read and spontaneous speech.

Table 2 shows that in Russian and Dutch the segment durations are longer in spontaneous than in read speech, but not much. In Russian, the durations are longer for spontaneous speech (mean of all segments 73,6 ms) compared to reading (69,6 ms) for all speakers regardless of the age and gender. Average shortening (S>R) in percentages = 5.43%.

Dutch vowels show the same tendency (mean in spontaneous speech 96 ms; in read speech 93 ms).

Average shortening (S>R) in percentages = 3.2%. The tendency remains even, if the exceptional speaker D-M2 is removed.

The Finnish data show a different tendency, namely, the duration values in read speech are longer than in spontaneous speech for all sounds and all speakers except in two female speakers. Average lengthening (S<R) in percentages = 6.47%

The average duration of phonologically long vowels in Finnish (85 ms) is only 28,8 ms longer (20%) than that of the short vowels (56 ms) in spontaneous speech. In read speech the average difference is greater: 43 ms. The average of the long vowels is 102 ms and that of short vowels 59 ms, i.e. the long ones are about 74% longer than the short ones. That means that in read speech the quantity distinction is clearer. Earlier investigations with read isolated words and carrier sentences have showed even more distinct differences [2], [7].

It is necessary to regard Finnish diphthongs separately, because their segmentation into two vowel components was not possible. The measurements showed that diphthongs are comparable more to short vowels in duration than to long ones and that they do not differ much in the two styles (mean differences 59 ms/S and 62 ms/R, that is only 3 ms in favor of read speech).

3.3 OMISSION OF SPEECH SOUNDS

In connected speech, both consonants and vowels, or even whole syllables may be elided, changing the phonetic make-up of words. This phenomenon is common in the Dutch word final plurality suffix /@n/ and its derivatives. The underlying final /n/ is redundant and only surfaces in some phonological contexts as a connection sound. However, it appears in the orthography and it resurfaces in read speech. The data in Table 3 shows that different speech styles differ in this aspect.

Speaking style	/@n/ plural		/@n/ non-plural	
	/@n/	/@/	/@n/	/@/
Read	34.98	65.02	34.5	65.53
Spont.	2.53	97.47	9.73	90.27
Total	24.08	75.92	27.51	72.49

Table 3. Omission of word final /n/ in /@n/ suffixes in Dutch (%). The non-plural suffixes are either derived from plurals or behave as such 'by similarity'. @ = schwa vowel.

As Table 3 proves, the resurfacing of /n/ in word-final /@n/ suffixes is strongly driven by speaking style. In Spontaneous speech it is all but absent for plurals (2.5%),

but in read speech the word-final /n/ is pronounced in over a third of all occurrences. Conceptually, this can, at least partially, be interpreted as an attempt to speak "correct" Dutch based on the feeling that orthography represents the language 'standard'.

In Russian, /j/ is particularly prone to elision [5]. Figure 1 illustrates the percentage of words where /j/ was preserved in spontaneous and read speech. /j/-dropping appears to be more common in spontaneous speech — regardless of position. Elision after front vowels took place more often than in other vowel contexts. Word-final /j/-s disappear less frequently than non-final /j/-s. The first post-stressed syllable with /j/ is affected more often than the second syllable after stress.

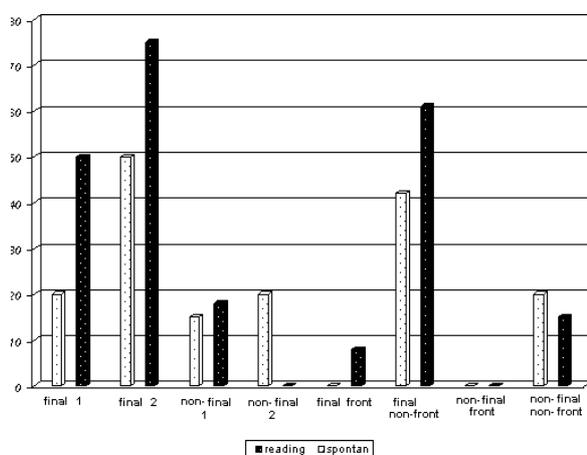


Figure 1: Percentage of /j/ preservation in read and spontaneous speech in Russian.

One of the typical features of the Finnish language is that the spoken language, even the standard Finnish, differs clearly from the written language. The dropping of final /n/ (e.g. in *sitte(n)* 'then') is possible even in cases in which it is a part of grammatical suffix (e.g. *me tultii(n)* 'we came'). The data on spontaneous speech showed that 51,2 % of word final /n/ which is very common in literary form of Finnish got elision. In this group also the cases were included in which /-n/ was assimilated according to the initial (non-nasal) consonant of the following word. In the rest of the cases the nasal /n/ was 1) intact or 2) occurred as a labial or velar nasal as a result of an assimilation according to the initial consonant of the following word, or 3) the remaining final vowel was nasalized.

4. CONCLUSIONS

The fundamental characteristics that distinguish spontaneous speech from read speech are partly due to the amount of planning required to either produce speech on the spot or to read a text prepared beforehand (even if it is a transcript of one's own spontaneous speech).

Comparison over styles and languages shows that

1) The material investigated so far allow us to conclude that in Russian, the F_0 values in the two speech styles do not vary much. At the same time, spontaneous speech is characterised by higher average sound duration and a greater number of cases with /j/ elision compared to read speech.

2) Spontaneous-read differences for Dutch subjects do not show a consistent difference in mean F_0 . There seems to be a tendency for longer sound durations in spontaneous speech. Speakers mark Read speech clearly by adding redundant underlying /n/ sounds to their word-endings.

3) In Finnish, the F_0 statistical parameters do not indicate much difference between the two speaking styles either for female speakers or for the male speakers. The duration of speech sounds is longer in read than in spontaneous speech. Elision or assimilation with the initial consonant of the next word were common processes for final /n/ in Finnish spontaneous speech.

At this point of research the data does not allow any definitive conclusions about the phonetic phenomena associated with speech styles. Besides that the number of speakers studied so far is small, some linguistic differences between the three languages apparently give rise to methodological problems. However, interesting tendencies in similarities and differences could be discovered which earn further elucidation.

ACKNOWLEDGMENTS

This study was supported by INTAS (project number 00-915) and the Academy of Finland (project numbers 78014 and 79394).

REFERENCES

- [1] B. Blaauw, "On the perceptual classification of spontaneous and read speech", *Proceedings of the ICPhS95*, Stockholm, Sweden, V.3: pp.254-257, 1995.
- [2] J. Lehtonen, *Aspects of Quantity in Standard Finnish*, Gummerus, Jyväskylä, 1970.
- [3] V.S.Martynov *Statistical parameters of F0 in speech*, PhD thesis, Leningrad, 1965 (in Russian).
- [4] Yu.V.Mikheyev, "Statistical law of F_0 periods distribution in Russian speech", *Journal of Acoustics*, vol.16, 4, pp.558-562, 1970 (in Russian).
- [5] *Phonetics of spontaneous speech*, ed. N.Svetozarova, Leningrad: Leningrad State University, 1988 (in Russian).
- [6] R. J. J. H. van Son et al., "The IFA CORPUS: a Phonemically Segmented Dutch 'Open Source' Database". *Eurospeech 2001*.
- [7] K. Wiik, *Finnish and English vowels*, Turku: University of Turku, 1965.