

The role of grammatical and semantic context in spoken word recognition

Olga Raeva

[St. Petersburg State University](#)

olgaspaces@rambler.ru

Abstract

In order to recognize reduced word forms, listeners use phonetic cues (Riekhakaynen 2010) and contextual information (Van de Ven et al. 2012). Some highly reduced word forms in spontaneous Russian (for instance, [zœz] for seychas 'now') do not contain enough phonetic cues for reconstruction of the canonical variant. However, according to the results of the pilot experiment reported in this paper, most of these reduced forms are recognized without any difficulties. The paper suggests that the role of contexts for the identification of reduced word forms should be considered for future studies of spoken word recognition.

Keywords: *Spontaneous speech; spoken word recognition; reduction; context.*

1. Introduction

One of the main characteristics of spontaneous speech is considered to be the elements' reduction in speech signals. This has been studied and documented for a number of languages; examples include (Kohler 1990) for German, (Ernestus 2000) for Dutch, (Shockey 2003) for British English, (Johnson 2004) for American English, (Heegård 2012; Schachtenhaufen 2012) for Danish, (Bondarko L.V. et al. 1988) for Russian, etc.

The information about characteristics of casual speech allows us to make a functional model for spoken word recognition, namely, to explain those procedures that take place in utterance processing.

1.1 Aim of the article

The main purpose of the research project from which this paper has evolved is to verify the fact that native speakers are able to identify strongly reduced word forms in natural speech recognition without any difficulties. In order to address this issue, the pilot study of perception of reduced Russian speech is described in this article.

1.2 Previous research on Russian

Reduction and variability of spontaneous Russian were investigated in many studies (e.g., Zemskaya 1973 and Bogdanova 2009). Previous research has shown that up to 50% of all Russian word forms are reduced in the spoken language (Apushkina et al. 2011). These results are based on the data from the [Corpus of Spontaneous Russian](#).

1.3 Data

The Corpus of Spontaneous Russian consists of TV and radio programs (overall duration is about 90 minutes). The corpus has been transcribed in the following way: first, the auditory texts were converted into the orthographic form. During this procedure, the expert group of phoneticians had no difficulties identifying lexical units. In order to reach as an exact phonetic transcription as possible, five professional phoneticians have identified word forms unanimously, while listening to the utterances involving a broad context.

The next step of processing speech materials was to transcribe them carefully. The phonetic transcriptions confirmed the fact that spontaneous Russian contains a large amount of variation, ranging from canonical, full-distinct pronunciation to strongly reduced realizations.

With the goal of finding out whether listeners have difficulties with recognition of strongly reduced word forms (namely, whether all of respondents can identify such kind of realizations accurately), an auditory experiment on perception of short speech fragments utilizing dictation task was run.

2. Method

Strongly reduced word forms have been chosen for creating two experimental sequences.

For instance, the canonical realization [pata'muʂta] for the word form *potomu_çto*¹ 'because' consists of 9 elements: 4 vowels and 5 consonants. Reduced realization [dvaʂ] consists of 4 elements: 1 vowel and 3 consonants. These forms have only 2 components in common: the vowel [a] and the fricative consonant [ʂ].

The differences between canonical and reduced pronunciation variants can be illustrated with their dynamic spectrograms (see Figure 1).

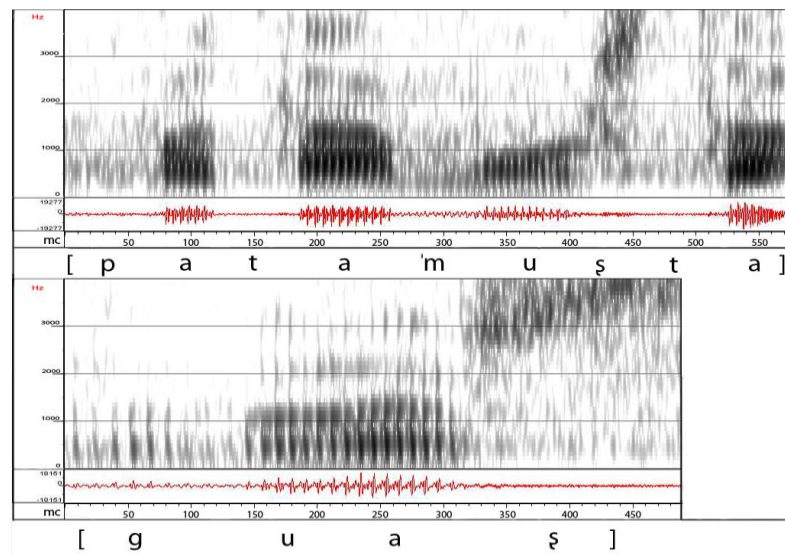


Figure 1. The spectrograms of pronunciation variants for the word form *potomu_çto* 'because'

Twenty-two strongly reduced realizations have been extracted from the Corpus of Spontaneous Russian, with right and left context (7 words on average). The main condition for including forms in the experiment is that these pronunciation variants are very different from canonical ones (see Table 1).

¹ The connection between two graphic words with an underscore indicates a phonological word.

Table 1. The strongly reduced realization included to the experiment

N	Word form	Canonical variant	Reduced variant
1	<i>vstrechayetsya</i> '(is) met'	[fstr'i'te aitsa]	[stij'te eets]
2	<i>poluchayut</i> '(they) get'	[palu'te ajut]	[bu'te et]
3	<i>govorim</i> '(we) say'	[gava'r'im]	[veəm]
4	<i>situatsiyu</i> (SG.ACC) 'situation'	[s'itu'atsiju]	[stots]
5	<i>lyudey</i> (PL.ACC) 'people'	[l'u'd'ej]	[v'i'd'e:]
6	<i>chisle</i> — the part of the conjunction that means 'including'	[tsei'sl'e]	[ete]
7	<i>ochen</i> 'very'	['ote in ^j]	[lutε]
8	<i>seychas</i> 'now'	[s'i'teas]	[zœz]
9	<i>tut</i> 'here'	[tut]	[də]
10	<i>mne</i> (1SG.DAT) 'I'	[mn ^j e]	[n'œ]
11	<i>u_nego</i> 'he (has)'	[un ^j i'vo]	[n'uu ⁱ]
12	<i>moikh</i> (PL.ACC) 'my'	[ma'ix]	[ba'œ]
13	<i>yego</i> (SG.ACC) 'his'	[ji'vo]	[u]
14	<i>kazhetsya</i> '(it) seems'	['kazits a]	['kasu]
15	<i>izvinite</i> 'excuse'	[izv ^j i'n ^j it ^j i]	[i'idʒ]
16			[dvaʃ]
17	<i>potomu_cho</i> 'because'	[pata'muʃta]	[guaʃ]
18			['toʃə]
19	<i>kogda</i> 'when'	[ka'gda]	[kaw ⁱ]
20	<i>skazat_odnu_ochen</i> 'to say one very'	[ska'zat ^j a'dnu 'otein ^j]	[ske'zadnœte]
21	<i>cho_eto</i> 'what is this'	['ʃtoeta]	['zida]
22	<i>sotsialnoy</i> (SG.GEN) 'social'	[sats i'al'naj]	[s:a'ne]

Two experimental sequences, each including 11 stimuli and one training stimulus, were prepared. Due to the fact that any fragment of spontaneous speech is an alternation of full and deformed variants, the listeners could not guess that they were being presented with strongly reduced realizations. That is why only one training stimulus was included in each sequence, and why it was presented at the beginning to help participants in adaptation. There was a 25 second pause after each stimulus to allow listeners to write

down what they had just heard. Every fragment was presented only once, preceded by a special, short warning signal.

The overall duration is 6.05 minutes for the first sequence, and 6.08 minutes for the second one (including time for instruction). The list of stimuli is presented in the Appendix.

Forty participants took part in the experiment (27 females and 13 males). Each experimental sequence was presented to 20 listeners.

3. Results

Any word form that was recognized by listeners as one word and did not contradict the semantics of surrounding context was considered to be a correctly identified unit. During the calculation of respondents' reactions, different grammatical word forms were considered as one and the same word.

A response that did not correspond to the semantics of the context was considered to be an incorrectly recognized word form. Only 12 identifications out of 480 reactions (2.5%) were wrong. This small number of incorrect responses allows us to conclude that participants in general transformed what they had heard into semantically adequate phrases. Listeners probably built the semantic structure using grammatical and semantic information available from the context, and identified the reduced word form according to it. We believe that this indicates that if several variants are possible in a context, the decision is made in compliance with acoustic cues of a reduced word form and/or with the word frequency.

In terms of analysis of identification processes, the most interesting realizations are those that were recognized ambiguously and, on the contrary, unambiguously within presented context. Thus, the fact of ambiguity or unambiguity was the reason for distributing reduced word forms into two groups.

Eleven realizations in the first group were recognized as one word form (for instance, *mne* (1SG.DAT) 'I', *seychas* 'now', etc.); the second group consists of realizations recognized as several different word forms (for instance, *tut* 'here', *sotsialnoy* (SG.GEN) 'social', etc., altogether 11 units) (see Table 2). The analysis of respondents' reactions allows us to hypothesize that in some cases, activation of only one word form is possible (probably in case of strong context influence). However, this assumption requires further research.

Table 2. Groups of realizations created during data processing

1. Reduced realizations recognized as one word form	2. Reduced realizations recognized as several word forms
<i>mne</i> (1SG.DAT) 'I'	1. <i>potomu_cho</i> ['toʂə] 'because'
<i>ochen</i> 'very'	2. <i>moikh</i> (1PL.GEN) 'my'
<i>kazhetsya</i> '(it) seems'	3. <i>chisle</i> — the part of the conjunction that means 'including'
<i>situatsiyu</i> (SG.ACC) 'situation'	4. <i>u_nego</i> 'he (has)'
<i>vstrechayetsya</i> '(is) met'	5. <i>kogda</i> 'when'
<i>potomu_cho</i> [guaʂ] 'because'	6. <i>sotsialnoy</i> (SG.GEN) 'social'
<i>poluchayut</i> '(they) get'	7. <i>potomu_cho</i> [dvaʂ] 'because'
<i>lyudey</i> (PL.ACC) 'people'	8. <i>tut</i> 'here'
<i>seychas</i> 'now'	9. <i>izvinite</i> 'excuse'
<i>govorim</i> '(we) say'	10. <i>yego</i> (3SG.ACC) 'his'
<i>skazat_odnu_ochen</i> 'to say one very'	11. <i>cho_eto</i> 'what is that'

The reduced variants of the second group were recognized as several different word forms (see Table 3). Thus, it is necessary to think about a likely explanation of this fact. The realization [s:a'ne] was chosen for more detailed analysis.

Table 3. Respondents' reactions received for stimulus [s:a'ne].

Respondents' answers	<i>sotsialnoy</i> 'social' (Gen.Sg.)	The first part of the compound <i>samo(reabelitatsii)</i> 'self-rehabilitation'	<i>v zone</i> 'in the area of'
Canonical realization	[saʈs i'al'naj] ²	[sama-]	['vzon'i] ³
Percentage	60%	35%	5%

The important fact is that all of these options do not contradict the context. Furthermore, their canonical forms have common elements with the strong-

² Common sounds are in bold.

³ Realizations ['vzon'i] and [s:a'ne] have no common components, but such elements as the palatalized consonant [nʲ] and the voiced fricative [z] are very close to the consonants [n] and [s].

ly reduced variant [s:a'ne]. We interpret these data in accordance with the idea that available contextual information and acoustic features are important to listeners during spoken word recognition.

4. Future perspectives

In the next stage, the necessary amount of context for variability removal will be determined. We are going to understand how the context should be expanded for definite identification of strongly reduced forms. Another goal is to expose to what extent acoustic cues, preserved in the phonetic shape of the reduced form, will be useful. We are going to conduct an additional experiment to conduct a deeper analysis of the role of semantic structure in a phrase. The same stimuli with missing target words will be presented to participants in the written form. It will help us to obtain those lexical units that are acceptable in terms of semantic and grammatical structure, and to explain why these words are not chosen during the processing of auditory speech fragments.

5. Conclusion

In conclusion, the described experiment was a trial. The results confirmed the assumption that in most cases, native Russian speakers can identify strongly reduced word forms presented once in a short context just as well as the group of phoneticians. It can probably be explained by the critical role of the semantic and grammatical information as well as by the remaining acoustic features used for spoken word recognition.

Furthermore, we plan to test this hypothesis, changing the amount of context for the stimuli and estimating which acoustic information is used for the interpretation of a reduced form.

Acknowledgements

This work was supported by research grant number MK-3646.2013.6 from the President of the Russian Federation.

References

Apushkina, Irina, Riekhakaynen, Elena, Slepokurova, Natalia, Ventsov, Anatolii. 2011. Frequency word list of spontaneous Russian. Poster presented at Production and Comprehension of Conversational Speech

- Workshop 2011, Max Planck Institute for Psycholinguistics, Nijmegen, 12–13 December.
- Bogdanova, Natalia V. 2010. Redutsirovannye formy — porcha yazyka ili fakt yego evolyutsii? [Reduced word forms: are they the damage of language or the fact of its evolution?] *Fonetika segodnya. Materialy dokladov i soobshcheniy VI mezhdunarodnoy nauchnoy konferentsii 8–10 oktyabrya 2010 goda*. P. 17–20.
- Bondarko, Lia V. et al. 1988. Fonetika spontannoy rechi [Phonetics of spontaneous speech].
- Ernestus, Mirjam. 2000. Voice assimilation and segment reduction in casual Dutch. A corpus-based study of the phonology-phonetics interface.
- Heegård, Jan. 2012. Functional indistinctiveness: Danish verbal and adjectival –te and –ede endings. *Speech in Action. Proceedings of the 1st SJUSK Conference on Contemporary Speech Habits*. P. 29–51.
- Kohler, Klaus J. 1990. Segmental reduction in connected speech in German: Phonological facts and phonetic explanations. *Speech Production and Speech Modelling*. P. 69–92.
- Johnson, Keith. 2004. Massive reduction in conversational American English. *Spontaneous speech: Data and analysis*. P. 29–54.
- Schachtenhaufen, Ruben. Reduction of word final schwa in Danish — Rhythmic and syntactic factors. *Speech in Action. Proceedings of the 1st SJUSK Conference on Contemporary Speech Habits*. P. 53–71.
- Shockey, Linda. 2003. *Sound patterns of spoken English*. Blackwell.
- Van de Ven, Marco, Ernestus, Mirjam, Schreuder, Robert. 2012. Predicting acoustically reduced words in spontaneous speech: The role of semantic/syntactic and acoustic cues in context. *Laboratory Phonology*, 3. P. 455–481.
- Riekhakaynen, Elena I. 2010. Vzaimodeystviye kontekstnoy predskazuyemosti i chastotnosti v protsesse vospriyatiya spontannoy rechi (na materiale russkogo yazyka) [The interaction of context predictability and frequency in processing of spontaneous speech: Evidence from Russian]. Russia: St. Petersburg State University dissertation.
- [The Corpus of Spontaneous Russian](#). 2009.
- Zemskaya, Elena A. (ed.). 1973. *Russkaya razgovornaya rech* [Russian spoken speech]. Moscow.

Appendix

List of stimuli №1

1. *k* *deystvitelno* *seryoznoy* *literature,* *mne* *kazhetsya*
 [k dʲi'stʲitn sʲirʲi'æzni lʲitʲira'tuʲrʲi 'mnʲe 'kazsə]
 to really serious literature I seem
 'to really serious literature, it seems to me' (duration: 0:00:03:00)
2. *mne* *bylo* *semnadtsat* *ili* *vosemnadtsat* *let*
 [nʲæ 'bəle sʲə'mnats̃ əlʲ vasə'mnats̃ 'lʲet]
 I be seventeen or eighteen year
 'I was seventeen or eighteen years old' (duration: 0:00:02:00)
3. *ya* *seychas* *budu* *dumat* *ob* *udarenii*
 [ʲe zœz vivə 'duvətʲ eb ədaʲ'rʲinʲi]
 I now will think about stress
 'now I am going to think about the stress' (duration: 0:00:01:11)
4. *u* *nego* *net* *vremeni* *na* *eto* *on* *dolzhen* *bezhat*
 [ø* nʲu nʲet 'vrʲemʲinʲi na 'et un 'dozən bʲe'zətʲ]
 PREP he no time for this he must run
 'he has no time for this, he must be fast' (duration: 0:00:02:11)
5. *kogda* *poluchayut* *ochen* *malenkuyu* *zarplatu*
 [ka'gda bu'tʲet 'oʲtʲe inʲ 'malʲenʲku zar'platu]
 when get very small salary
 'when they get very small salary' (duration: 0:00:01:29)
6. *u* *nas* *mnogo* *chego* *tut* *sdelano*
 [ʲu 'nas 'mogəʲ tʲe æ də 'zdʲelənə]
 PREP we much what here do
 'we have a lot that is done here' (duration: 0:00:01:08)
7. *ya* *sidela* *snachala* *mne* *ochen* *hotelos* *skazat*
 [je sʲi'dʲe sna'tʲeʲa 'vnʲe lu'tʲe xa'tʲæ ska'za'tʲ]
 I sit at first I very want say
 'at first I sat, I really wanted to say' (duration: 0:00:01:21)
8. *mne* *kazhetsya* *chto* *zdes* *dolzno* *byt* *splavleno*
 [nʲe 'kasu ʂto 'zdʲinʲ do'zno bitʲ splavlʲuna]

I seem that here must be fuse
'it seems to me that it must be fused here' (duration: 0:00:02:03)

9. *desyat tysyach, dvenadtsat tysyach eto izmenit situatciyu*
['dʲesʲi 'tisʲitɕ dvʲe'nats 'tʲɕ 'etʲɪz,mʲenʲ sʲtɔts]
ten thousand twelve thousand this change situation
'ten thousand, twelve thousand, will this change the situation?' (duration: 0:00:02:16)

10. *sprashivat moikh znakomykh tam ya ne znayu*
['spraʂoʲtʲ baʲæ zna'komix tam ʲe nʲe 'znae]
ask my friend there I no know
'to ask my friends, well, I don't know' (duration: 0:00:02:19)

11. *rugat shkolu bessmyslenno potomu_chno mozhno*
[ru'gaʲtʲ 'ʂkoli bəs'mislʲnə dvaʂ moʒ]
criticize school meaninglessly because can

tochno tak zhe rugat vsyo ostalnoye
['toʲtɕ n tɕ zɕ lgaʲtʲ fsʲɕɛsta'ʲnoʲe]
exactly so PARTICLE criticize everything rest
'to criticize the school is meaninglessly, because one can criticize everything else in the same way' (duration: 0:00:03:16)

12. *togda u nas chto-to izvinite poluchitsya*
[ta'gda u 'nas 'ʂtota ʲi'ɪdʒ pa'lutɕ]
then PREP we something excuse result
'then, excuse me, we will be successful in something' (duration: 0:00:03:08)

List of stimuli №2

1. *hotya vse eti elementy zdes prisutstvuyut*
[xaʲ'tʲa: 'fsʲi ə'tʲilʲi'mʲenti z'dʲisʲ pri'sustuvʲɪt]
though all these element here attend
'though all these elements are here' (duration: 0:00:03:01)

2. *esli yego babushka ne byla shkolnoy uchitelnitsey*
[esʲu u 'babʲʂkə 'nʲebəle 'ʂkoʲʲnə ə'tɕ itnʲitɕe]
if his grandmother no be school teacher
'if his grandmother was not a school teacher' (duration: 0:00:02:09)

3. *ya hochu skazat_odnu_ochen veshch prostuyu*
[ʲa xetɕ ske'zadnətɕ vʲɕɕ pra'stue]
I want say one very thing simple

'I want to say one very simple thing' (duration: 0:00:01:23)

4. <i>opat</i>	<i>zhe</i>	<i>imeyushchikh</i>	<i>problemy</i>	<i>sotsialnoy</i>	<i>reabilitatsii</i>
[a'pʲelʲ]		zʲi'mʲiɛx	pro'bʲlɛm	s:a'ne	rʲebʲəlʲe'tatse]
again	PARTICLE	have	problem	social	rehabilitation

'again having problems of a social rehabilitation' (duration: 0:00:02:13)

5. <i>my</i>	<i>vsò</i>	<i>vremya</i>	<i>govorim</i>	<i>vot</i>	<i>uchitelya</i>	<i>takie</i>
[mʲi]	vsʲɔ	re'mə	veəm	vot	əstə'lʲe	ta'kii
we	all	time	say	PARTICLE	teacher	such

uchitelya *syakie*
 u'tɛ itʲe'lʲe sʲi'kʲii]
 teacher like that

'we always say: «teachers are so-and-so»' (duration: 0:00:02:25)

6. <i>ona</i>	<i>vstreichayetsya</i>	<i>znayete</i>	<i>ona</i>	<i>i</i>	<i>v</i>
[əna	stijʲtɛɛts		zni'tʲani	ə	ø
she	meet	know	she	and	in

VUZ'e *inogda* *vstreichayetsya*
 'vuʲzʲena'gda stri'itɛ aets]
 university sometimes meet

'it is met, you know, sometimes it is met even in the University' (duration: 0:00:03:10)

7. <i>kogda</i>	<i>chelovek</i>	<i>prishol</i>	<i>v</i>	<i>parikmakherskuyu</i>
[kaɐʲ	tɛɛ	pə'ʂol	ø	prʲik'ma'xʲərskuə]
when	man	come	in	hairdressing salon

'when a man has come to hairdressing salon' (duration: 0:00:01:27)

8. <i>segodnya</i>	<i>tam</i>	<i>prisut-...</i>	<i>nu</i>	<i>chto_eto</i>	<i>v</i>
[sʲi'voʲdʲe	tam	prʲi'sut	nu	'zʲɪda	f
today	there	attend	PARTICLE	what	this in

pervom *klasse*
 'pʲervom 'kla'sʲe]
 first class

'today they are there..., and what is this, in the first class' (duration: 0:00:02:48)

9. <i>i</i>	<i>v</i>	<i>tom</i>	<i>chisle</i>	<i>tam</i>	<i>i</i>	<i>togo</i>	<i>mesta</i>
[i:	f	tom	ɛte	'tam	ə	ta'vo	'mʲestə]
and	PREP	that	number	there	and	that	place

'and including there and that place' (duration: 0:00:01:14)

10. *ne day Bog vospityvat potomu_chno segodnyashnyaya shkola*
 [nʲe 'dae 'bok va'spʲid gʷaʂ sʲi'vodnʲiʂnʲe 'ʂkole]
 no give God educate because today school
 'God forbid to educate because today's school' (duration: 0:00:02:16)

11. *potomu_chno vezde rabotayut lyudi*
 ['toʂə vʲi'zdʲe ra'bitət 'lʲu'dʲi]
 because everywhere work people
 'because people work everywhere' (duration: 0:00:01:17)

12. *bolshinstvo nashikh molodykh lyudey okazhetsya potom*
 [bəlʲʂʲi'stvo 'naʂex maa'dix vʲi'dʲe: a'kastsə pa'tom
 majority our young people find then
 v toy kolonyi
 f 'toʲj ka'lo'nʲie]
 PREP that colony
 'then the majority of our young men will be found in that colony' (duration: 0:00:05:23)

* The element was not presented acoustically, but it could be reconstructed without any difficulties in accordance with the semantic and grammatical context.