REDUCTION OF UNSTRESSED CENTRAL AND BACK VOWELS IN CONTEMPORARY STANDARD BULGARIAN

Mitko Sabev

University of Oxford
mitko.sabev@phon.ox.ac.uk

ABSTRACT

This paper reports two experiments that examine unstressed vowel reduction and neutralisation in Contemporary Standard Bulgarian. Stress-dependent F1 and F2 frequency variation was measured for the central and back vowel pairs in careful speech. The results reveal a less unequivocal picture of the pairwise convergence of formant frequencies than has previously been reported: the data confirm a high degree of acoustic overlap only for the unstressed back vowels, /ɔ – u/, whereas the differences in the unstressed central vowels, /a – ə/, remain statistically significant. A perception experiment consisting of an identification and a discrimination task was designed to test whether the existing differences in unstressed vowels are perceptually significant. The correct-response rates in both tasks were not significantly higher than chance. This partial mismatch between acoustic and perceptual data points to incomplete, rather than categorical neutralisation.

Keywords: vowel, reduction, neutralisation, Bulgarian

1. INTRODUCTION

Contemporary Standard Bulgarian (CSB) has six contrastive vowels in stressed position: front /ɛ, i/, central /a, ə/, back /ɔ, u/. /ə/ is typically close-mid and somewhat retracted from central, therefore sometimes transcribed /ɤ/ [6, 9]. [7] establishes what may be called the 'traditional view' of vowel reduction (VR) in CSB, which has since been replicated and reaffirmed in numerous accounts, such as [3, 6, 9]. This view can be summarised as follows. (i) Of the three contrastive-height pairs, /ɛ – i, a – ə, ɔ – u/, the central and the back neutralise when unstressed in CSB. [7] also suggests that this is more pronounced with /a – ə/ than with /ɔ – u/. The front pair is neutralised in Eastern dialects only. (ii) The neutralised height value for each pair is intermediate between the values of the two stressed vowels. (iii) Two 'degrees of reduction' are recognised: the 'first' appears in the first pretonic only and is a more open realisation, while the 'second' occurs elsewhere and approaches the height of the higher vowel. [5, 8] report that the back vowels neutralise into the quality of the higher vowel, [u], while also demonstrating that unstressed /a – ə/ may remain acoustically distinct. [1] is the first corpus-based study of VR in CSB, and it further challenges claim (ii) by concluding that the neutralised realisations for both pairs largely overlap with the stressed realisations of the higher vowel in each pair.

This paper reports the results from two experiments carried out to assess the extent of acoustic and perceptual neutralisation of unstressed central and back vowels in CSB, as well as to identify their reduced realisations in careful speech.

2. EXPERIMENT 1: PRODUCTION

This experiment aimed to determine the effects of stress on the F1 and F2 frequencies and duration of central and back vowels, as well as to measure the acoustic overlap of the unstressed realisations of /a – ə/ and /ɔ – u/.

2.1. Subjects and procedure

8 female native speakers of CSB, all undergraduates at the University of Sofia, aged 18–24, volunteered to take part in this production experiment in return for nominal payment. Six of the subjects also participated in the second experiment.

The subjects were asked to read out a list of 40 target items and 40 distractors, in quasi-random order. The target items consisted of 5 repetitions of 8 Bulgarian words containing stressed and unstressed /a ə ɔ u/ – one word for each allophone (Table 1). The pairs with a stressed and an unstressed allophone of the same vowel were derivationally or inflectionally related; each word was also in a potential minimal or homophone pair with one of the other target words. The distractors mirrored the morphemic structure of the target items. The words were displayed to the participants on a computer screen, one at a time.

320 vowel tokens were analysed in total: 40 instances of each of the eight allophones (stressed and

Table 1: Target words used in the production experiment ([+*str*] = stressed; [−*str*] = unstressed)

	/a/	/ə/
[+*str*]	/ɔˈm**a**ʒa/ ‘splatter.1.sg.pres’	/ɔˈm**ə**ʒa/ ‘marry.1.sg.pres’
[−*str*]	/ɔm**a**ˈʒɛtɛ/ ‘splatter.imper.pl’	/ɔm**ə**ˈʒɛtɛ/ ‘marry.imper.pl’
	/ɔ/	/u/
[+*str*]	/ɔbraˈz**ɔ**vam/ ‘educate.1.sg.pres’	/ɔbraˈz**u**vam/ ‘form.1.sg.pres’
[−*str*]	/ɔbraz**ɔ**ˈvaniɛ/ ‘education’	/ɔbraz**u**ˈvaniɛ/ ‘formation’

unstressed /a ə ɔ u/). The speech was digitised at a sampling rate of 44,100 Hz. The vowels were manually segmented in Praat [2], on the basis of the synchronised spectrogram, waveform and audio signal. Vowel boundaries were determined by the presence of clear formant structure and a sharp change in intensity. F1 and F2 frequencies and duration were measured using a Praat script; formant frequencies were measured at the vowel mid-points.

2.2. Results

Outliers, defined as values beyond the interquartile range by 1.5 times the interquartile range, were excluded from further analysis (4 for F1, 3 for F2, 4 for duration). Table 2 shows the mean F1 and F2 frequencies and duration times for each allophone. This reveals a general pattern of decrease in F1 frequency for the unstressed lower vowels, /a, ɔ/, and a considerable decrease in duration for all vowels. The F2 differences point to a tendency towards fronting for all unstressed vowels except /a/.

Table 2: Mean F1 and F2 frequencies and duration (female) for stressed and unstressed /a ə ɔ u/ (*M* = mean; *SD* = standard deviation; $[V^{+}]$ = stressed; $[V^{-}]$ = unstressed)

Vowel	F1 (Hz)		F2 (Hz)		Duration (ms)	
	M	*SD*	*M*	*SD*	*M*	*SD*
$[a^{+}]$	792	120	1416	86	163.66	19.28
$[a^{-}]$	590	72	1407	111	81.04	24.53
$[ə^{+}]$	522	29	1327	122	114.55	17.53
$[ə^{-}]$	549	52	1451	124	77.12	23.47
$[ɔ^{+}]$	559	35	1157	95	130.67	17.92
$[ɔ^{-}]$	442	33	1244	135	58.96	11.84
$[u^{+}]$	430	25	1080	127	100.55	16.11
$[u^{-}]$	435	26	1245	119	56.68	14.22

A MANOVA was performed for each of the four vowel phonemes with stress as an independent variable, and F1 and F2 frequencies and duration as dependent variables. There was a significant overall effect of stress on the dependent variables for each of the four phonemes. However, univariate tests on the separate dependent variables revealed no significant effect of stress on F1 frequency for /u/, and on F2 frequency for /a/. Although significant, the effect of stress on F2 frequency was small ($\eta_p^2 < .3$) for /ɔ/ and /ə/ and medium-sized ($\eta_p^2 < .5$) for /u/.

Paired *t*-tests were performed to assess the significance of differences in F1 and F2 frequency across allophones (Table 3). The following differences are *non*-significant. For F1: ($[u^{+}]$, $[u^{-}]$) and ($[ɔ^{-}]$, $[u^{-}]$); ($[ɔ^{-}]$, $[u^{+}]$) is only marginally significant ($p = .0494$). For F2: ($[a^{+}]$, $[a^{-}]$), ($[a^{+}]$, $[ə^{-}]$) and ($[ɔ^{-}]$, $[u^{-}]$). The effect size is large ($r \geq .50$) for all significant results except for three cases where it is medium ($r = [.30, .50)$): ($[ə^{+}]$, $[ə^{-}]$) and ($[ɔ^{-}]$, $[u^{+}]$) for F1, and ($[a^{-}]$, $[ə^{-}]$) for F2.

Table 3: Paired *t*-test results for F1 and F2 freq.

Vowel pair	F1 frequency		
$[a^{+}]$, $[a^{-}]$	$t(39) = 12.56$	$p < .001$	$r = .9$
$[ə^{+}]$, $[ə^{-}]$	$t(36) = -3.4$	$p < .003$	$r = .49$
$[a^{+}]$, $[ə^{+}]$	$t(36) = 13.97$	$p < .001$	$r = .92$
$[a^{-}]$, $[ə^{-}]$	$t(39) = 3.92$	$p < .001$	$r = .53$
$[a^{+}]$, $[ə^{-}]$	$t(39) = 12.67$	$p < .001$	$r = .9$
$[a^{-}]$, $[ə^{+}]$	$t(36) = 5.51$	$p < .001$	$r = .68$
$[ɔ^{+}]$, $[ɔ^{-}]$	$t(39) = 13.06$	$p < .001$	$r = .9$
$[u^{+}]$, $[u^{-}]$	$t(38) = -0.86$	$p > .3$	$r = .14$
$[ɔ^{+}]$, $[u^{+}]$	$t(39) = 16.3$	$p < .001$	$r = .93$
$[ɔ^{-}]$, $[u^{-}]$	$t(38) = 1.83$	$p > .07$	$r = .28$
$[ɔ^{+}]$, $[u^{-}]$	$t(38) = 14.9$	$p < .001$	$r = .92$
$[ɔ^{-}]$, $[u^{+}]$	$t(39) = 2.03$	$p = .0494$	$r = .31$
Vowel pair	**F2 frequency**		
$[a^{+}]$, $[a^{-}]$	$t(37) = 0.75$	$p > .4$	$r = .12$
$[ə^{+}]$, $[ə^{-}]$	$t(39) = -6.79$	$p < .001$	$r = .74$
$[a^{+}]$, $[ə^{+}]$	$t(37) = 5.64$	$p < .001$	$r = .68$
$[a^{-}]$, $[ə^{-}]$	$t(39) = -2.33$	$p < 0.03$	$r = .35$
$[a^{+}]$, $[ə^{-}]$	$t(37) = -1.49$	$p > .1$	$r = .24$
$[a^{-}]$, $[ə^{+}]$	$t(39) = 3.91$	$p < .001$	$r = .53$
$[ɔ^{+}]$, $[ɔ^{-}]$	$t(39) = -3.86$	$p < .001$	$r = .53$
$[u^{+}]$, $[u^{-}]$	$t(38) = -7.57$	$p < .001$	$r = .78$
$[ɔ^{+}]$, $[u^{+}]$	$t(39) = 4.16$	$p < .001$	$r = .55$
$[ɔ^{-}]$, $[u^{-}]$	$t(38) = 0.25$	$p > .8$	$r = .04$
$[ɔ^{+}]$, $[u^{-}]$	$t(38) = -3.97$	$p < .001$	$r = .54$
$[ɔ^{-}]$, $[u^{+}]$	$t(39) = 7.16$	$p < .001$	$r = .75$

A significant effect of stress on duration was found for all phonemes, as shown in Table 4. An unfortunate confound must be acknowledged here: the target words for unstressed vowels are longer than those for stressed vowels, which may have somewhat inflated the reported effects.

Table 4: Effect of stress on duration (univariate analysis)

/a/	$F(1,76) = 265.03$	$p < .001$	$\eta_p^2 = 0.78$
/ə/	$F(1,74) = 57$	$p < .001$	$\eta_p^2 = 0.44$
/ɔ/	$F(1,75) = 422.03$	$p < .001$	$\eta_p^2 = 0.85$
/u/	$F(1,76) = 157.67$	$p < .001$	$\eta_p^2 = 0.67$

Pearson's r computed to assess the correlation between vowel duration and F1 frequency *within allophones* reveals *no* significant correlation for any allophones except [a$^-$] ($r = 0.353$, $n = 40$, $p = .025$) and [ə$^+$] ($r = -0.347$, $n = 36$, $p = .038$).

3. EXPERIMENT 2: PERCEPTION

This experiment consisted of two tasks, identification and discrimination; its purpose was to test whether the acoustic differences between unstressed vowels established in the production study are perceptually significant.

3.1. Subjects and procedure

13 native speakers of CSB, 10 female and 3 male, all undergraduates at the University of Sofia, aged 18–24, volunteered to take part in this production experiment in return for nominal payment.

Recordings of the eight target words from Experiment 1 (Table 1) were used as stimuli for the *identification* task. The words with a stressed target vowel were used as a control sample. Each stimulus was played 10 times (80 tokens per subject), in quasi-random order. The stimuli were produced by an 18-year-old female subject in Experiment 1, who did not take part in either part of the perception experiment. The stimuli were chosen so that there would be at least a 30 Hz difference in F1 frequency between the target vowels in each neutralisation pair ([a$^-$] 585 Hz, [ə$^-$] 547 Hz; [ɔ$^-$] 472 Hz, [u$^-$] 424 Hz). The recordings were played to the subjects through headphones, in a quiet room. After listening to each stimulus, participants were given a forced-choice task to choose between two images, each representing one of the concepts in the minimal or homophone pair.

The stimuli for the *discrimination* part consisted of paired tokens of the same recordings as those played for the identification task: 16 pairs of either 2 identical words, or each of the words in a minimal/homophone pair. Each stimulus was played 5 times (80 tokens per subject), in quasi-random order. The stimuli were played to the subjects under the same conditions. After listening to each stimulus, participants were given a forced choice to identify the two words as 'same' or 'different'.

3.2. Results

Identification 99.42% of the stressed vowels (control) and 59.42% of the unstressed vowels were identified correctly. Table 5 shows the results for each vowel.

Table 5: Correct identification rates for stressed and unstressed /a ə ɔ u/

	/a/	/ə/	/ɔ/	/u/
$[+str]$	99%	100%	99%	99%
$[-str]$	59%	57%	59%	62%

A paired t-test was performed to assess the significance of the difference in correct responses to stressed vs unstressed target vowel stimuli and showed a significant strong effect of stress on correct identification ($t(12) = 14.5$, $p < .001$, $r = .97$). A second paired t-test was used to assess the difference in the correct identification of [a$^-$ – ə$^-$] (58%) vs [ɔ$^-$ – u$^-$] (61%). No significant difference was found ($t(12) = -0.68$, $p > .5$, $r = .19$).

Discrimination The overall results from this task match quite closely the results obtained in the first part: the correct discrimination rate is 99.62% for both stressed pairs, 57.31% for [a$^-$ – ə$^-$] and 59.62% for [ɔ$^-$ – u$^-$]. The vast majority (82%) of incorrect responses to unstressed target vowel stimuli are different words identified as 'same'.

A paired t-test was performed to assess the significance of the difference in correct responses to stimulus pairs containing stressed vs unstressed vowels, and showed a significant strong effect of stress on correct discrimination ($t(12) = 25.46$, $p < .001$, $r = .99$). Another paired t-test was applied to assess the difference in the discrimination between [a$^-$ – ə$^-$] (57% correct) vs [ɔ$^-$ – u$^-$] (60% correct). No significant difference was found ($t(12) = -0.66$, $p > .5$, $r = .19$). *D*-prime was calculated to quantify mean discrimination performance by subtracting the z-transformed 'false alarms' (proportion of same stimuli identified as different) from the z-transformed 'hits' (proportion of different stimuli identified as different). The result was mean $d' = 0.568$, and does not significantly differ from 0, or chance ($p > .07$).

4. DISCUSSION AND CONCLUSIONS

The production experiment described here reveals a less orderly spectral pattern of VR in CSB than has previously been reported or assumed. First, while the data confirm a high degree of acoustic overlap for the unstressed *back* vowels /ɔ – u/, the F1 difference for the unstressed *central* vowels remains statistically significant ($p < .001$), which is consistent with [8]. Second, /ə/ appears to undergo an increase in F1 and F2 frequency in unstressed contexts (i.e., mid-centralisation from a somewhat closer and more retracted stressed articulation), rather than undershoot. Third, while /ɔ/ is clearly raised in unstressed syllables, both /ɔ/ and /u/ are also fronted when unstressed. Table 6 summarises the statistically significant F1 and F2 frequency differences I have found across allophones.

Table 6: Significant F1 and F2 frequency differences across allophones of /a ə, ɔ u/

Vowel pair	Diff.	Vowel pair	Diff.
[a⁺], [a⁻]	F1, —	[ɔ⁺], [ɔ⁻]	F1, F2
[ə⁺], [ə⁻]	F1, F2	[u⁺], [u⁻]	—, F2
[a⁺], [ə⁺]	F1, F2	[ɔ⁺], [u⁺]	F1, F2
[a⁻], [ə⁻]	F1, F2	[ɔ⁻], [u⁻]	—, —
[a⁺], [ə⁻]	F1, —	[ɔ⁺], [u⁻]	F1, F2
[a⁻], [ə⁺]	F1, F2	[ɔ⁻], [u⁺]	(F1), F2

This is at odds with previous accounts, which maintain that height contrasts are neutralised in both back and central unstressed vowels. In terms of unstressed *realisations*, for /ɔ – u/ my measurements point to [u̟], which is more in line with [5, 8, 1] than with the 'traditional view', which argues for an intermediate unstressed value [7, 3, 6, 9]. As regards /a – ə/, however, the unstressed realisations measured here are indeed closer than the values for [a⁺] but more open that those for [ə⁺], which is at variance with what [1] report. (It should be noted, however, that my sample was not controlled for coarticulatory effects, whereas such effects will have largely cancelled each other out in the considerably larger and more diverse sample analysed in [1].)

All unstressed vowels analysed here are in the first pretonic syllable. It remains to be verified that first pretonic allophones are significantly different to unstressed realisations elsewhere.

There is a significant effect of stress on both duration and (except for /u/) F1 frequency. Duration therefore appears to be a rather salient, perhaps categorical cue for stress in CSB. At the same time the correlation between duration and F1 frequency within allophones is either non-significant or rather weak. Since significant F1 frequency differences obtain only between categorically different conditions (stressed/unstressed), and F1 frequency is not a linear function of duration *within* conditions, one may at least tentatively conclude that unstressed realisations implement separate, phonologised targets. While these unstressed targets are merged for the two back vowels, which are therefore categorically neutralised, the neutralisation of the central vowels is incomplete as their unstressed realisations remain acoustically (though not perceptually) distinct.

F1 frequency differences of 38 to 48 Hz did not prove to be perceptually significant in unstressed vowels: neither the identification nor the discrimination task yielded a correct response rate that is significantly higher than chance. Thus, while acoustic neutralisation has only been established for the pair [ɔ⁻ – u⁻], the perceptual data show that both [ɔ⁻ – u⁻] and [a⁻ – ə⁻] are non-distinct. Table 7 summarises the neutralisation patterns for the three pairs of CSB vowels (including the front vowels, which are not examined in this paper).

Table 7: Acoustic and perceptual neutralisation of CSB unstressed front, central and back vowels

Vowel pair		Neutralisation	
		Acoustic	Perceptual
Front	/ɛ – i/	−	−
Central	/a – ə/	−	+
Back	/ɔ – u/	+	+

This uneven pattern of neutralisation appears to mirror the asymmetrical shape of the vowel space, which is in turn a function of the asymmetrical properties of the vocal tract. Unstressed vowels are fully merged at the back, where there is least space, whereas the somewhat more spacious central area can still host two separate targets, though these are perceptually indistinguishable; the front, on the other hand, allows enough mouth opening and acoustic distance for the two vowels to remain perceptually distinct in all contexts. This conforms to the model of VR proposed in [4].

It is not at all unlikely that the formant frequency differences between [a⁻] and [ə⁻] found in careful speech are not as a rule maintained in more natural or casual speech, as [1] may be taken to demonstrate. However, the very fact that reduction may be dependent on style or register, and thus be a matter of conscious choice, as it were, points once again to incomplete neutralisation rather than a categorical merger for this pair.

Acknowledgements I am grateful to Elinor Payne, John Coleman and Matt Husband for many helpful comments. This paper has also benefited from feedback received at the 2014 LEL PG Conference (University of Edinburgh). Any remaining errors or omissions are mine alone.

5. REFERENCES

[1] Andreeva, B., Barry, W., Koreman, J. 2013. The Bulgarian Stressed and Unstressed Vowel System. A Corpus Study. *Proc. 14th Interspeech.* Lyon, France. 345–348.

[2] Boersma, P., Weenink, D. 1992–2012. Praat, v. 5.3.16. http://www.fon.hum.uva.nl/praat/.

[3] Boyadžiev, T., Tilkov, D. 1997. *Fonetica na bălgarskija knižoven ezik* (The phonetics of Standard Bulgarian). Veliko Tărnovo: Abagar.

[4] Flemming, E. 2005. *A Phonetically-Based Model of Phonological Vowel Reduction.* Unpublished manuscript. MIT, Cambridge, MA. web.mit.edu/∼flemming/www/paper/vowelred.pdf (Retrieved 7 Nov 2013).

[5] Pettersson, T., Wood, S. 1987. Vowel reduction in Bulgarian and its implications for theories of vowel reduction: a review of the problem. *Folia linguistica* 21, 261–280.

[6] Ternes, E., Vladimirova-Buhtz, T. 1999. Bulgarian. In: *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press 55–57.

[7] Tilkov, D., (ed) 1982. *Gramatika na săvremennija bălgarski knižoven ezik* (Grammar of contemporary Standard Bulgarian) volume 1: *Fonetika* (Phonetics). Sofia: Bulgarian Academy of Sciences Press.

[8] Wood, S., Pettersson, T. 1988. Vowel reduction in Bulgarian: the phonetic data and model experiments. *Folia linguistica* 22, 239–262.

[9] Žobov, V. 2004. *Zvukovete v bălgarskija ezik* (The sounds of Bulgarian). Sofia: Sema RŠ.