A CROSS-VARIETAL CONTINUUM OF UNSTRESSED VOWEL REDUCTION: EVIDENCE FROM BULGARIAN AND TURKISH

Mitko Sabev, Elinor Payne

University of Oxford
mitko.sabev@phon.ox.ac.uk, elinor.payne@phon.ox.ac.uk

ABSTRACT

We compare speech production data from three Bulgarian and two Turkish varieties with respect to spectral and durational reduction of unstressed non-front unrounded vowels, and ensuing height neutralisation. Istanbul Turkish lies at one end of a reduction continuum, with only non-neutralising, gradient F1 frequency undershoot that correlates with duration. Monolingual East Bulgarian lies at the opposite end: unstressed, underlyingly non-high vowels raise considerably and merge with their high counterparts. The Bulgarian speech of bilingual Turkish–Bulgarian speakers from the same region of eastern Bulgaria shows less reduction and neutralisation; perhaps surprisingly, it resembles the reduction pattern of West (Standard) Bulgarian, while at the same time also being gradient, probably under the influence of Turkish. The bilinguals' Turkish speech, on the other hand, exhibits more neutralisation than Istanbul Turkish, but less than their own Bulgarian, which in turn suggests prosodic transfer from these speakers' Bulgarian to their Turkish.

Keywords: vowel reduction, vowel merger, incomplete neutralisation, Bulgarian, Turkish

1. INTRODUCTION

The phonological literature often cites Bulgarian as having a system of six contrastive stressed vowels that neatly contracts into a three-vowel unstressed inventory through reduction and height neutralisation in each of the pairs /i-ɛ, ə-a, u-ɔ/ [1–4]. However, phonetic descriptions make clear that while this may be true of certain eastern dialects, it is not so in Standard (western) Bulgarian, where only /ə-a/ and /u-ɔ/, but not /i-ɛ/, merge in unstressed position. Two claims have traditionally been made about the phonetic realisations of the merged vowels, both of which we challenge in this paper: (1) that unstressed high and non-high vowels merge into intermediate realisations, such as [ɐ] and [o], respectively, and (2) that immediately pretonic vowels are more open than other unstressed vowels. [5–8]

Turkish, too, has a system that can be symmetrically divided into height-contrasting pairs, /i-ɛ, y-œ, ɯ-ɑ, u-ɔ/, but is not known for unstressed vowel reduction (UVR). This is not surprising, for it has been argued that Turkish is a pitch accent language [9] and shows no significant stress-dependent spectral differences [10], and that reduction cannot coexist with vowel harmony (VH) in a language [11].

We consider spectral and durational data from five varieties expected to exhibit different degrees of UVR: East Bulgarian (EB), West Bulgarian (WB), bilingual Bulgarian (BB) and Turkish (BT) as spoken by a long-standing bilingual community from the same region as the EB monolingual group (Tărgovište), and Istanbul Turkish (IT). Based on previous research and assumptions, we expect any spectral UVR in IT to be gradient and correlated with duration, as could result from gestural undershoot under temporal pressure. For WB and EB categorical UVR is expected, with much less or no correlation with duration. EB should show more UVR, and greater merger of high and non-high unstressed vowels. It is difficult to predict where exactly BT and BB will lie on a reduction scale, as these varieties have not been studied before. If we are to believe that VH and UVR are typologically incompatible, there should be a definitive watershed point with BT and BB logically lying on different sides, or one of them having crossed that point as a result of L1–L2 transfer. [11] argues that VH and UVR do not cooccur because VH neutralises backness and roundness while preserving height contrasts, whereas UVR does precisely the opposite. The watershed point should then be defined in terms of both the difference between stressed and unstressed allophones, and – perhaps more importantly – the degree of height contrast neutralisation, i.e. merger.

For space limitations, we focus here on Bulgarian /ə-a/ (*ъ-а*), and Turkish /ɯ-ɑ/ (*ı-a*). As well as being phonologically matched (both pairs are non-front/back and non-round), these vowels are sufficiently phonetically similar to map onto one another in loan adaptation and foreign language learning.

2. METHODS

2.1. Participants, procedure and recording

Native speakers of each variety, aged 18–27, were recorded in return for nominal payment (Table 1). The bilinguals' Bulgarian and Turkish recordings were made on separate days. Participants read out utterances from a computer screen, displayed one at a time, in a quasi-random order. The recordings were made on a MacBook with an iRig Mic HD external microphone, digitised at a sampling rate of 44,100 Hz and a 16-bit resolution, stored as PCM-encoded single-channel WAV files.

Variety	Sp	M	F	Vowel tokens
EB	12	5	7	1216
WB	12	4	8	1344
BB	14	5	9	1064
BT	14	5	9	1232
IT	14	6	8	1176

Table 1: Number of M(ale) and F(emale) Sp(eakers) and vowel tokens per experiment.

2.2. Test items

A set of nonsense words were designed to test the effects of various accentual and segmental contexts on the three variables of interest: F1, F2 and duration. Real words were also included to verify that the nonsense words were reliable indicators for speakers' natural linguistic behaviour. Comparisons of real- and nonsense-word vowels revealed mostly non-significant differences in formant frequencies.

The nonsense words had the shape (C)VCVˈCVC(C)V(C), usually keeping vowel and consonant phonemes constant throughout the word, and conforming to Bulgarian/Turkish phonotactics. As far as possible, the nonsense words had comparable segmental and accentual structure across the two languages. In Bulgarian, stress was indicated with a grave accent on the vowel. Turkish stress is generally word-final, so in order to elicit penultimate stress, the enclitic /lɑ, lɛ/ 'with' (or /diɾ, dyɾ, dɯɾ, duɾ/ 'is') was attached to a trisyllabic nonsense stem. Multiple unstressed syllables per word were necessary to test the repeated claim that pretonic vowels in Bulgarian are more open than other unstressed vowels. Consonants varied for: place of articulation, /p t k/; voicing, /p b/; degree of stricture, /b v/; C ~ ∅ word-initially and finally. In some words stressed and unstressed vowels differed in phonological height. Except for vowels being consistently longer in all varieties when flanked by voiced consonants, segmental context did not yield any conclusive results. The words appeared in the carrier sentences /ˈkazax … ˈpak/ 'I said … again' (Bulgarian) and /ɑhˈmɛt … dɛˈdi/ 'Ahmet said …' (Turkish).

2.3. Analysis

The vowels were manually segmented in Praat [12], on the basis of the synchronised spectrogram, waveform and audio signal. Vowel boundaries were determined by the presence of clear formant structure and sharp changes in intensity. A Praat script was used to extract vowel duration and midpoint F1 and F2 frequencies.

Outliers, defined as values 1.5 times beyond the interquartile range, and devoiced vowels were excluded from analysis. Formant values were normalised using [13]'s adaptation of Neary's vowel-extrinsic, formant-intrinsic log-mean procedure, as implemented in the *vowels* package for R [14,15].

A combination of repeated-measures ANOVA, MANOVA and *t*-tests were performed to measure the effect of stress and syllable ($\sigma_1, \sigma_2, \acute{\sigma}_3, \sigma_4$) on F1, F2 and duration. Pearson product-moment correlation between duration and F1 within $\sigma_{1\text{–}4}$ allophones was used as an indicator for gradient undershoot.

F1/F2/duration distribution overlap was calculated between (1) stressed and unstressed allophones for each vowel, lesser overlap indicating greater acoustic *displacement* in unstressed position; (2) high and low vowels in unstressed position as a measure of *merger*, greater overlap corresponding to a higher degree of merger. Two overlap measurement procedures were used. In the first, allophones are plotted as 2-SD best-fit ellipsoids in the F1/F2/duration space. The fraction of overlap between the two ellipsoids is computed, which can range from 0 (no overlap) to 1 (complete overlap) [16]. The second metric used is Pillai's trace, part of the output of MANOVA performed with F1, F2 and duration as dependent variables. A higher value indicates a greater difference between the two distributions with respect to the dependent variables. For a discussion of these and other vowel overlap metrics, see [17].

3. RESULTS

The ANOVA outputs showed significant main effects in all cases except for F1 in IT /ɯ/. To establish which comparisons are likely to have produced these effects, dependent *t*-tests with Bonferroni correction (α = 0.0125 for Bulgarian, and 0.0167 for Turkish) were performed to compare the three parameters across the three/four syllables (Table 2).

The mean values in Table 3 interpreted in the context of significant differences (Table 2) reveal a number of clear patterns. In all varieties, **duration** is longest in the stressed (third) syllable, and in Bulgarian the second longest syllable is the fourth and final one. This pattern is only broken by /ə/ in BB, where [$ə_4$] is even longer than [$\acute{ə}_3$], and in EB, where [$\acute{ə}_3$] = [$ə_4$]. In all Bulgarian dialects the second and pretonic vowel tends to be the shortest, while in Turkish V_1 and V_2 often have equal durations. **F1** is highest in stressed position across the board for the low vowel, and also for the high vowel in WB and BT, while in EB [$ə_4$] > [$ə_{1,2,3}$], in BB [$ə_4$] > [$ə_1$] > [$ə_{2,3}$], and in IT [$ɯ_1$] = [$ɯ_2$] = [$ɯ_3$]. F1 is lowest in the second (i.e. pretonic) syllable in all varieties but IT. **F2** is higher for the stressed low vowel in Bulgarian and BT, while the high vowel is inconsistent in this respect.

L	σσ	B /ə/ / T /ɯ/ *Dur.*	*F1*	*F2*	B /a/ / T /ɑ/ *Dur.*	*F1*	*F2*
EB	1-2	n.s.	n.s.	0.004	n.s.	n.s.	***
	1-**3**	n.s.	n.s.	n.s.	***	***	0.020
	2-**3**	n.s.	n.s.	n.s.	***	***	***
	4-1	n.s.	***	n.s.	0.006	n.s.	n.s.
	4-2	0.027	n.s.	n.s.	0.002	0.012	***
	4-**3**	n.s.	0.003	n.s.	0.001	***	***
WB	1-2	n.s.	n.s.	***	***	***	***
	1-**3**	***	0.001	n.s.	***	***	0.030
	2-**3**	***	***	***	***	***	***
	4-1	***	***	n.s.	n.s.	n.s.	***
	4-2	0.001	***	***	0.033	0.003	0.022
	4-**3**	***	n.s.	n.s.	***	***	***
BB	1-2	***	***	***	***	***	***
	1-**3**	0.025	0.022	***	***	***	n.s.
	2-**3**	***	n.s.	***	***	***	***
	4-1	***	***	***	***	0.002	***
	4-2	***	***	***	***	***	***
	4-**3**	***	***	n.s.	***	***	***
BT	1-2	n.s.	0.044	***	n.s.	***	***
	1-**3**	***	0.041	0.001	***	***	0.040
	2-**3**	***	***	***	***	***	***
IT	1-2	0.013	n.s.	***	n.s.	***	***
	1-**3**	***	n.s.	***	***	***	***
	2-**3**	***	n.s.	n.s.	***	***	0.006

Table 2: Significance of differences between allophones. L: variety, Dur.: duration, ***: $p < 0.001$

		ə / ɯ σ1	σ2	σ́3	σ4	a / ɑ σ1	σ2	σ́3	σ4
Duration	EB	46.78	47.82	51.44	53.68	44.72	44.06	77.76	58.93
	WB	57.43	59.82	78.36	67.10	62.85	57.96	94.74	61.97
	BB	48.09	43.10	50.62	57.31	53.14	46.37	73.32	60.65
	BT	45.97	45.35	57.83		53.04	54.65	88.67	
	IT	47.31	50.95	57.91		59.23	58.60	87.16	
F1′ freq.	EB	411	429	414	437	426	416	640	454
	WB	392	388	405	409	478	456	639	472
	BB	417	403	409	437	551	482	600	529
	BT	406	395	420		551	494	619	
	IT	389	385	392		533	495	615	
F2′ freq.	EB	972	908	907	923	990	888	1065	940
	WB	1092	1015	1101	1082	1155	1074	1184	1106
	BB	979	876	917	915	1014	915	1003	954
	BT	1015	853	950		1014	889	1044	
	IT	1121	1008	1008		1077	1012	1038	

Table 3: Mean allophone duration in ms, and normalised F1 and F2 frequencies in Hz.

We find consistent linear F1–duration **correlation** in most Turkish cases, as well as in BB (Table 4); we interpret this as an indication of *gradient* undershoot (as opposed to categorical reduction). In the monolingual Bulgarian varieties, correlation is only sporadic, suggesting *categorical* UVR. F2–duration correlations do not present a clear pattern.

		Duration–F1 σ1	σ2	σ́3	σ4	Duration–F2 σ1	σ2	σ́3	σ4
ə / ɯ	EB	n.s.	n.s.	n.s.	n.s.	***	n.s.	0.006	n.s.
	WB	0.009	n.s.	0.001	n.s.	0.001	n.s.	***	n.s.
	BB	n.s.	0.001	0.002	n.s.	***	n.s.	0.001	***
	BT	n.s.	0.026	n.s.		n.s.	n.s.	***	
	IT	n.s.	0.044	0.008		0.030	n.s.	n.s.	
a / ɑ	EB	n.s.	n.s.	***	0.035	0.003	n.s.	***	n.s.
	WB	0.016	***	n.s.	0.001	n.s.	n.s.	0.019	n.s.
	BB	***	0.005	***	***	***	n.s.	***	***
	BT	0.001	0.033	***		n.s.	n.s.	***	
	IT	0.005	***	***		n.s.	n.s.	n.s.	

Table 4: Significance of Pearson's *r* correlations between duration & F1 and duration & F2.

Figure 1 shows degrees of **displacement**, or differences between stressed and unstressed distributions for each vowel, quantified as (**A**) the complement of the fraction of overlap (1 – Overlap), and (**B**) Pillai's trace with stress as independent variable. The low vowel naturally undergoes much greater displacement, with the highest degree in EB and WB, less in BT and IT, and even less in BB. The two statistics produce the same relative results, although **B** is more compatible with a two-way grouping: (EB, WB) > (BB, IT, BT), while **A** yields three apparent categories: (EB, WB) > (BT, IT) > (BB).

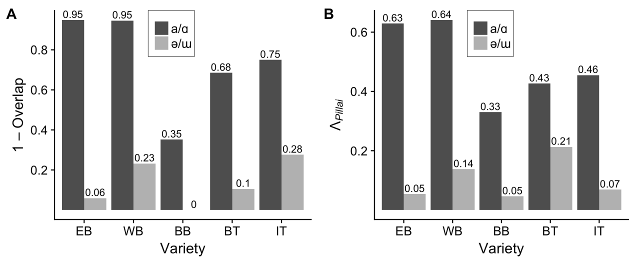

Figure 1: Degrees of unstressed displacement.

To quantify **merger** between high and low unstressed vowels, we take (**A**) their spectral–durational overlap (spectral overlap shown in Fig. 3), and (**B**) the complement of Pillai's trace ($1 - \Lambda_{Pillai}$) with phoneme as independent variable (Fig. 2). Once again, the two metrics show very similar relative rankings but suggest different groupings. While three categories emerge in both – *high* (EB), *medium* (WB, BB) and *low* (IT) – BT is inconclusive, appearing as *medium* in **A**, but closer to *low* in **B**.

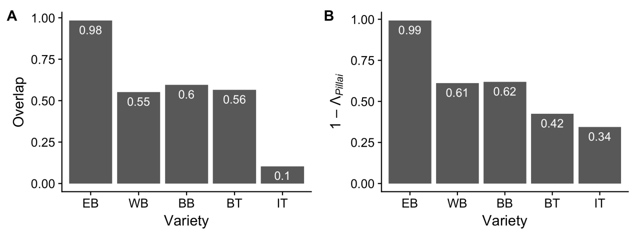

Figure 2: Degrees of merger between high and low unstressed vowels.

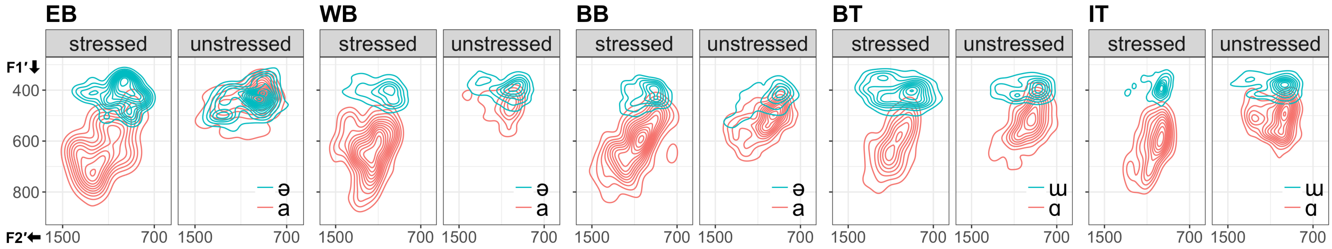

Figure 3: F1/F2 distributions of stressed and unstressed vowels, based on kernel density estimation [17,18].

4. DISCUSSION AND CONCLUSIONS

The results we report indicate that vowel **duration** is a robust correlate for word stress in all five varieties. This is more marked in low than in high vowels, and in WB than in the other varieties. Vowels in final unstressed syllables were analysed for Bulgarian only and proved to be consistently longer than other unstressed vowels, possibly due to boundary effects. In Bulgarian, pretonic vowels are the shortest, while in Turkish durational differences between σ_1 and σ_2 unstressed vowels are mostly non-significant.

F1 frequency is significantly higher in stressed position for low vowels in all varieties. High unstressed vowels remain mostly unchanged in IT, EB and BB, but raise slightly in WB and BT. If there is any L1–L2 transfer in this regard, BT is influenced, surprisingly, by (Standard) WB, not by EB or BB.

The received view that Bulgarian high vowels undergo lowering when unstressed [5–8] has been refuted in previous acoustic studies [19,20] and in recent ultrasound/acoustic work [21]. Here we confirm no lowering for unstressed /ə/, with the slight exception of unstressed final syllables in EB and BB, where we report an F1 increase by up to 28 Hz.

Another established view that our results disprove, to the best of our knowledge for the first time, is that Bulgarian unstressed vowels are more open when pretonic than elsewhere: we find *more* spectral and durational reduction there, not less. The traditional view echoes the two-level UVR system standardly reported for Russian [22–24], where a penultimately stressed four-syllable word has a sonority-sequenced single-peak prominence profile – $_x$ x X $_x$ – while our spectral and durational data reveal a more punctuated, low-level 'trochaic' pattern: x $_x$ X x.

Our **linear correlations** reveal gradient, duration-dependent F1 undershoot in Turkish, while Bulgarian UVR is largely independent of temporal pressure and therefore categorical. BB, however, parallels Turkish in correlating F1 with duration.

Comparisons of F1/F2/duration distributions for stressed and unstressed allophones show much greater **displacement** in low than in high vowels. This is strongest in EB and WB where there is hardly any overlap between stressed and unstressed /a/. As expected, Turkish /ɑ/ shows considerably less displacement. Interestingly, BB has least displacement, but since it also differs from the rest of Bulgarian in correlating F1 with duration, we may in fact be looking at only two degrees of displacement: stronger and categorical (EB, WB) vs weaker and gradient (IT, BT, BB), which also tallies with the Pillai's trace displacement metric (Fig. 1B).

Displacement measurements alone reveal little about the systemic effects of UVR, and we therefore also consider the degree of **merger** of high and low unstressed vowels. In EB we find complete overlap. Although it is at a medium level in WB, we know from previous acoustic and perceptual research [20] that the degree of acoustic overlap between WB unstressed /a-ə/ is sufficient for perceptual neutralisation. BB shows practically the same level of merger as WB, but we should not automatically conclude that this is also perceptually neutralising, because the spectral–durational correlation there may be a factor that enables speakers to perceptually offset displacement. We find least merger in IT, as predicted. BT exhibits more merger than IT, probably due to transfer from Bulgarian. The two merger metrics are at odds here (Fig. 2): the overlap fraction statistic yields a higher output and places BT in the medium group along with WB and BB. Pillai's trace, on the other hand, points to less merger and positions BT closer to IT. This is probably a more reliable metric, as it calculates pooled effect variances, and it is less likely to inflate the output in cases like BT, where a formant frequency correlates with duration.

Our analysis places the five varieties on a UVR continuum as shown in Table 5. BB is influenced by Turkish in its gradient displacement, while BT shows Bulgarian interference in merging more than IT. Our results indicate that UVR is not unequivocally incompatible with VH: both IT and BT exhibit some displacement and merger, with more merger in BT. However, further research is needed to study the perceptual implications of these findings.

	EB	WB	BB	BT	IT
Displace-ment	strong		weak		
	categorical		gradient		
Merger	high	medium		med.-low	low

Table 5: A cross-varietal continuum of UVR.

7. REFERENCES

[1] Trubetzkoy, N.S. 1931/1969. *Principles of Phonology*. Berkeley: University of California Press.

[2] Scatton, E. 1975. *Bulgarian Phonology*. Cambridge, MA: Slavica Publishers.

[3] Crosswhite, K. 2001. *Vowel Reduction in Optimality Theory*. London: Routledge.

[4] Harris, J. 2005. Vowel reduction as information loss. In: Carr, P., Durand, J., Ewen, C.J., (eds), *Headhood, elements, specification and contrastivity*. Amsterdam: Benjamins, 119–132.

[5] Tilkov, D., Bojadžiev, T., Georgieva, E., Penčev, J., Stankov, V. 1982. *Gramatika na săvremennija bălgarski knižoven ezik* (Grammar of contemporary Standard Bulgarian), vol. 1: *Fonetika* (Phonetics). Sofia: Izdatelstvo na BAN.

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

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

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

[9] Levi, S.V. 2005. Acoustic correlates of lexical accent in Turkish. *J. Acoust. Soc. Am*. 35, 73–97.

[10] Konrot, A.K. 1981. Physical correlates of linguistic stress in Turkish. *University of Essex Language Centre Occasional Papers*. 24, 26–53.

[11] Barnes, J. (2006). *Strength and Weakness at the Interface: Positional Neutralization in Phonetics and Phonology*. Berlin: Walter de Gruyter.

[12] Boersma, P., Weenink, D. 1992–2018. Praat, v. 6.0.43. http://www.fon.hum.uva.nl/praat/.

[13] Labov, W., Ash, S., Boberg, C. 2006. *The Atlas of North American English: Phonology, Phonetics, and Sound Change. A Multimedia Reference Tool*. Berlin: Mouton de Gruyter.

[14] Kendall, T., Thomas, E.R. 2010. Vowels: Vowel Manipulation, Normalization, and Plotting in R. R package. Version 1.1. https://cran.r-project.org/web/packages/vowels/index.html.

[15] R Core Team. 2017. R: A Language and Environment for Statistical Computing. Version 3.3.3. https://www.R-project.org/.

[16] Wassink, A.B. (2006). A geometric representation of spectral and temporal vowel features: Quantification of vowel overlap in three linguistic varieties. *J. Acoust. Soc. Am.* 119, 2334–2350.

[17] Nycz, J., Hall-Lew, L. 2013. Best practices in measuring vowel merger. *J. Acoust. Soc. Am.* 134, 4198.

[18] Di Canio, C. 2013. Visualizing vowel spaces in R: from points to contour maps. Accessed 1 Nov 2018. http://christiandicanio.blogspot.com/2013/10/visualizing-vowel-spaces-in-r-from.html

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

[20] Sabev, M. 2015. Reduction of unstressed central and back vowels in Contemporary Standard Bulgarian. In: *Proc. 18th ICPhS*. Glasgow.

[21] Dokovova, M., Sabev, M., Scobbie, J., Lickley, R., Cowen, S. 2019. Bulgarian vowel reduction in unstressed position: an ultrasound and acoustic investigation. In: *Proc. 19th ICPhS*. Melbourne.

[22] Shvedova, N.Ju. 1980. *Russkaia grammatika* (Russian Grammar), vol. 1. Moscow: Nauka.

[23] Timberlake, A. 2004. *A Reference Grammar of Russian*. Cambridge: Cambridge University Press.

[24] Iosad, P. 2012. Vowel reduction in Russian: No phonetics in phonology. *Journal of Linguistics.* 48, 521–571.