INFORMATION DENSITY AND VOWEL DISPERSION IN THE PRODUCTIONS OF BULGARIAN L2 SPEAKERS OF GERMAN

Erika Brandt[1], Bistra Andreeva[2], Bernd Möbius[2]

[1]Friedrich Schiller University Jena, [2]Saarland University
erika.brandt@uni-jena.de, andreeva@coli.uni-saarland.de, moebius@coli.uni-saarland.de

ABSTRACT

We investigated the influence of information density (ID) on vowel space size in L2. Vowel dispersion was measured for the stressed tense vowels /iː, oː, aː/ and their lax counterpart /ɪ, ɔ, a/ in read speech from six German speakers, six advanced and six intermediate Bulgarian speakers of German. The Euclidean distance between center of the vowel space and formant values for each speaker was used as a measure for vowel dispersion. ID was calculated as the surprisal of the triphone of the preceding context. We found a significant positive correlation between surprisal and vowel dispersion in German native speakers. The advanced L2 speakers showed a significant positive relationship between these two measures, while this was not observed in intermediate L2 vowel productions. The intermediate speakers raised their vowel space, reflecting native Bulgarian vowel raising in unstressed positions.

Keywords: information density, vowel dispersion, L2, Bulgarian, German.

1. INTRODUCTION

Vowel dispersion is measured as the distance between vowel tokens and the center of a talker's vowel space as defined by the first two formants [7] and is widely used in sociophonetic studies [13, 23]: vowels in a space with a large dispersion are more distinct from each other than vowels produced with more central tongue height and frontness. Vowel space size is influenced by speech rate, phonological context, average vowel duration, sex of the speaker, and vowel identity [21, 19, 24, 23, 11, 22]. There are only few studies with a focus on information density (ID) factors and their impact on vowel dispersion. Word frequency [16, 14, 25, 13, 17] and language redundancy or predictability [10, 4, 9] have been identified as significant factors impacting on vowel dispersion in American English. Vowels are more dispersed in high information content. [18, 11] broadened this field by analyzing the effect and interaction of ID and prosodic structure on segmental variability in production studies from a cross-language perspective, including six languages (American English, German, French, Finnish, Czech, and Polish). In these studies, as in the present one, ID is defined as contextual predictability, or surprisal (Equation 1)

$$(1) \quad Surprisal(unit_i) = -log_2 P(unit_i|Context),$$

and estimated from language models (LMs) based on large text corpora. They found that vowels in high surprisal contexts were more dispersed than in low surprisal contexts.

In the present study, we widen the scope of previous analyses by investigating vowel dispersion of Bulgarian L2 speakers of German. We assume that native speakers of a language share the same LM, with some degree of individual variability due to idiolectal, sociolinguistic or regional factors. Due to their exposure to the L2, language learners presumably build mental models of the predictability of linguistic events in their target language. These models will vary as a function of the speaker's proficiency level and amount of exposure. We investigate if ID factors of the target language (German) can explain phonetic variability, and here specifically vowel dispersion, of Bulgarian L2 speakers at different proficiency levels. Our analysis thus introduces a new approach to investigating language learning from an information-theoretic perspective.

We predict that the relation between ID and patterns of vowel dispersion observed in L1 speakers is also apparent in advanced proficiency level (C2) language learners, but less pronounced, or even non-existent, in intermediate proficiency level (B2) learners.

2. MATERIAL

Six Bulgarian L2 speakers at an intermediate proficiency level (B2), age range 19–24 (M = 20), and six Bulgarian L2 speakers at an advanced proficiency level (C2), age range 36–54 (M = 43), were recorded in addition to six German native speakers (L1), age range 28–52 (M = 37). All subjects in this study were females.

The speakers were asked to read aloud German text passages from the EUROM-1 corpus [8] in a quiet environment with a head mounted microphone (AKG C520), digitized with an Audiobox (M-Audio Fast Track) using Praat [6] and its default settings for audio recordings (41 kHz, 16 Bit). Speakers were asked to read fluently and as if they were engaged in telephone conversations in a professional setting. Only vowels in accented position were analyzed in order to control for prosodic effects. Tense and lax vowels in the corner positions of the German vowel space were chosen for analysis: /aː, a, iː, ɪ, oː, ɔ/. In total, we analyzed 2,393 vowel tokens (L1 = 796, C2 = 797, B2 = 798).

3. METHOD

The German LM was based on the SDeWaC corpus, which was derived from the DeWaC corpus [5]. The web-crawled corpus contains 846,159,403 running words and 1,094,902 lexical types in a diverse range of genres. The corpus was transcribed using the g2p tool in German Festival [12]. The transcriptions of the 1,000 most frequent words in the corpus were manually verified by the first author. Systematic errors were identified and corrected for all lexical items in the corpus. The LM was calculated based on phonemes using the SRILM toolkit [20].

F1 and F2 were measured at the temporal midpoint in vocalic nuclei. Formant analysis was conducted with the Burg algorithm in Praat [6] with a maximum of five formants, window size of 0.025 sec, pre-emphasis from 50 Hz, and a maximum formant threshold of 5,500 Hz. Formant values were cleaned and manually checked before speaker-dependent normalization was applied to control for differences due to speaker identity [1]. Vowel dispersion was calculated as the Euclidean distance between the center of the vowel space and formant values for every vowel per speaker [7].

4. RESULTS

4.1. Descriptive statistics

On average, Bulgarian L2 speakers showed a larger vowel dispersion calculated for all vowels pooled than German natives ($M = 1.32; SD = 0.50$). B2 Bulgarian speakers ($M = 1.36; SD = 0.37$) were slightly more dispersed in their German vowel production than C2 speakers ($M = 1.34; SD = 0.43$). As can be seen in Figure 1, L2 speakers at both proficiency levels had similar vowel dispersion values for /aː, a/ as the native speakers. The back vowels /oː, ɔ/ were similarly dispersed in C2 and L1 speakers,

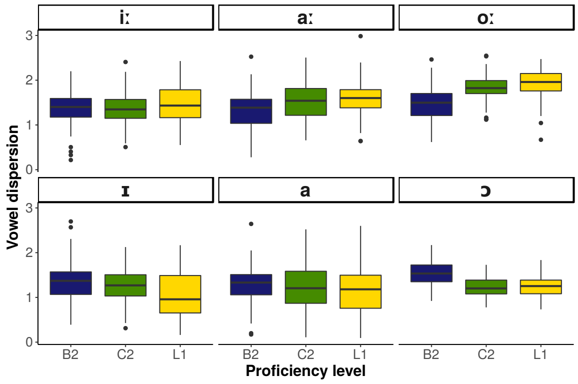

Figure 1: Vowel dispersion of German (L1) and Bulgarian speakers of German at intermediate (B2) and advanced (C2) proficiency levels per vowel phoneme.

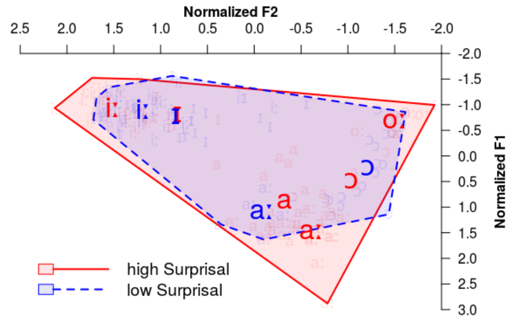

Figure 2: Vowel space of German L1 speakers in high and low surprisal contexts. Binning of surprisal was based on 10 % of the highest and lowest values in the data set.

whereas B2 speakers showed less dispersion for /oː/, and more for /ɔ/, compared to the other two speaker groups. With regard to the closed front vowels /iː, ɪ/ we found that neither the C2 nor the B2 learners reached the same level of dispersion as the German natives. While the target /iː/ was approached with a little less vowel dispersion than in the native speech, the L2 speakers showed much higher dispersion values for the lax vowel /ɪ/ .

Advanced Bulgarian speakers of German produced a pattern for vowel space expansion in different ID contexts similar to that observed for German L1 speakers (Figure 2).

For advanced Bulgarian L2 speakers, German back and mid vowels were less dispersed in low than in high surprisal contexts. The front vowels /iː/ and /ɪ/ approached a similar position in the vowel space under low surprisal, while they were clearly separated under high surprisal (Figure 3, left). In contrast, intermediate speakers did not show the expected pattern of vowel space reduction under low

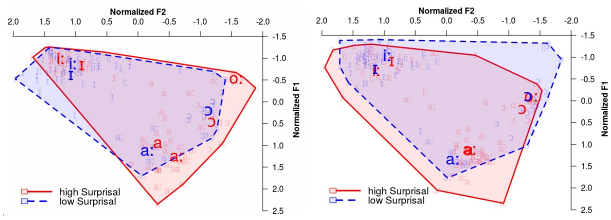

Figure 3: Vowel space of Bulgarian L2 speakers under high and low surprisal at advanced (left) and intermediate (right) proficiency level. Binning of surprisal was based on 10 % of the highest and lowest values in the data set.

compared to high surprisal. Instead, we found that low surprisal vowels were raised relative to high surprisal vowels (Figure 3, right). This pattern reflects the native Bulgarian pattern of vowel reduction [2].

4.2. Linear mixed-effects models

We calculated Pearson's r correlations between vowel dispersion and surprisal per speaker group. Vowel dispersion and triphone of the preceding context were significantly correlated for the L1 speakers ($r = 0.23; t(794) = 6.58; p < 0.001$) and the L2 speakers at C2 level ($r = 0.14; t(795) = 3.86; p < 0.001$). There was no significant correlation for the B2 speakers ($r = -0.02; t(796) = -0.69; p < 0.49$).

Following the results of the correlation analysis we calculated three different LMMs for each speaker group using triphone surprisal of the preceding context as an ID measure. We decided to control for vowel tenseness (tense vs. lax) in the statistical model. We also included average vowel duration based on the production data as a control variable. Because of the small number of data points per group we included both content and function words in the statistical analysis, while adding the factor word class to the model (Table 1).

A collinearity analysis was performed to identify potential dependencies between the factors. Word frequency and surprisal were moderately negatively correlated ($r = -0.62$). Word frequency and word class were strongly positively correlated ($r = 0.80$), with function words showing higher frequency values than content words. Word class and surprisal, on the other hand, showed a weaker negative correlation ($r = -0.50$) than word frequency and surprisal. Average vowel duration and surprisal were only weakly correlated ($r = 0.26$). Higher surprisal values were correlated with longer vowel duration. A similar relationship was observed for vowel dispersion and average vowel duration ($r = 0.23$). Vowel tenseness was correlated with average vowel duration as well ($r = 0.50$), indicating that tense vowels were longer than lax vowels.

As a result of the collinearity analysis surprisal, word class, average vowel duration, and vowel tenseness were included as fixed factors. Word frequency was excluded as a predictor in this model because it showed strong correlations with word class and a moderate correlation with surprisal. The random structure of the model consisted of random intercepts for speaker and word. LMMs with a larger random structure did not converge because of the small amount of data points per model. Vowel tenseness was sum-coded, word class was treatment-coded, and both continuous predictors were log-transformed. The model structure is given in Equation 2.

$$\begin{aligned} VowelDispersion \sim\ & TriSur + WordClass + \\ & Tenseness + DurAv \\ & (1|Speaker) + (1|Word) \end{aligned} \tag{2}$$

In the LMM for German L1 speakers (Table 1), we found the expected significant effects for the control factors. Long, tense vowels were more dispersed than short, lax vowels. Vowels in function words were less dispersed than vowels in content words. However, we only found a tendency for a positive effect of surprisal on vowel dispersion. In the model for the C2 speakers, there were significant effects of tenseness and duration in the expected directions, and a tendency for a negative effect of word class. Since we did not find a significant effect for surprisal on vowel dispersion in the model for German natives, we did not expect to observe a significant effect in the models for L2 speakers, because correlation values between surprisal and vowel dispersion were lower or non-significant for these two groups. In the LMM for Bulgarian B2 speakers of German, there were no significant effects of any of the fixed effects to explain vowel dispersion.

We calculated effect sizes for the three LMMs and their significant effects separately. The largest overall effect size of the entire model was found for the German natives ($Var = 63.39\%$). The same model structure explained only 37.71% of the variance in the data of L2 advanced speakers, and even less variance in the vowel dispersion of L2 intermediate speakers ($Var = 24.11\%$). For both L1 ($Var = 18.72\%$) and L2 advanced speakers ($Var = 10.17\%$) average vowel duration was the strongest predictor of vowel dispersion. Vowel tenseness added 6.42% explained variance for the German L1 data, and 1.48% in the model for C2 vowel dispersion. While word class was not significant in the L2 models, it explained 2.03% of data variance in German L1 vowel dispersion.

Table 1: Vowel dispersion of L1 and L2 speakers: regression coefficients, standard error (SE), and statistical output of LMM analyses including triphone surprisal of the preceding context (TriSur), tenseness (lax–tense), average vowel duration (DurAv), and word class (function–content words). *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Group	Terms	Coeff.	SE	t-value
L1	TriSur	0.02	0.08	0.25
	Tenseness	-0.21	0.04	-5.27***
	Word class	-0.19	0.07	-2.53*
	DurAv	0.42	0.11	3.81***
C2	TriSur	-0.03	0.07	-0.50
	Tenseness	-0.13	0.03	-4.13***
	Word class	-0.11	0.06	-1.98
	DurAv	0.31	0.12	2.58*
B2	TriSur	-0.05	0.06	-0.79
	Tenseness	-0.04	0.03	-1.36
	Word class	-0.001	0.05	-0.03
	DurAv	0.11	0.11	0.99

5. DISCUSSION

This study investigated whether Bulgarian L2 speakers of German behave similarly to German native speakers in their vowel dispersion in different surprisal contexts, and whether their vowel productions depended on their proficiency level of German.

German vowels were more dispersed when they were difficult to predict from their preceding context [18, 11]. Advanced L2 speakers showed a tendency to modulate their vowel productions in the same way as German natives with regard to ID factors, whereas intermediate L2 speakers were not able to make these distinctions. This finding indicates that the proficiency level of L2 speakers can be expressed as the degree of familiarity with the target (German) language structures and their predictabilities on a sub-word level.

Although we found significant positive correlations between surprisal and vowel dispersion for L1 and C2 speakers, this effect was not significant in a more complex LMM analysis including other control factors. This may be due to the small amount of data points per LMM, the restricted number of surprisal contexts because of the short length of the text passage, and the specific nature of vowel phonemes and their behavior under high and low surprisal.

The advanced L2 speakers were able to differentiate their vowel productions with regard to differences in tenseness and vowel duration. They also showed a tendency to produce native-like differences between vowel tokens in function or content words. These effects were not found in B2 speakers. We can therefore clearly separate the two proficiency levels using vowel dispersion as an acoustic measure. Interestingly, the amount of German competence of the three groups was also mirrored in the effect sizes of the corresponding LMMs. The effect size of the model decreased with decreasing proficiency level of German. Average vowel duration was the strongest predictor for vowel dispersion for both German L1 and advanced L2 speakers. However, this effect size should be interpreted with caution because vowel duration and tenseness were positively correlated ($r = 0.50$).

L2 competence varied between vowel phonemes (Figure 1). Bulgarian natives do not mark stress differences for /i/ in vowel height [2]. We speculate, they therefore show difficulties to differentiate German /iː/ and /ɪ/, irrespective of proficiency level. The main difference between German /oː/ and /ɔ/ is vowel height. Bulgarian natives also differentiate vowel height for unstressed and stressed /ɔ/ in their native tongue. While Bulgarian C2 speakers are able to adapt this difference in height to German tense/lax differences for this vowel pair, B2 speakers fail to make this difference. For Bulgarian /a/, the F1 and F2 values are not statistically separable from short German /a/ [3] and German tense and lax /a/ are not statistically different [15] which is potentially why Bulgarian L2 speakers are able to approach German native productions successfully.

We found that Bulgarian L2 speakers, in particular at B2 level, showed more vowel dispersion than German L1 speakers. This finding is not surprising considering that B2 speakers raised their vowel space under low surprisal, thus reflecting Bulgarian L1 vowel raising in unstressed condition [2] (Figure 3, right). Although all analyzed vowels were stressed we found this characteristic reduction pattern for low surprisal vowels in the B2 speakers. This can be interpreted as a certain degree of awareness in Bulgarian B2 speakers of the German phonological structures and their predictabilities. But they were not able to produce the target-language reduction pattern for vowels in low surprisal context and instead relied on their L1 reduction pattern.

6. ACKNOWLEDGEMENT

This research was funded by the German Research Foundation (DFG) as part of SFB 1102 ‘Information Density and Linguistic Encoding’ at Saarland University.

7. REFERENCES

[1] Adank, P., Smits, R., van Hout, R. 2004. A comparison of vowel normalization procedures for language variation research. *The Journal of the Acoustical Society of America* 116(5), 3099–3107.

[2] Andreeva, B., Barry, W., Koreman, J. 2013. The Bulgarian stressed and unstressed vowel system. A corpus study. *Proceedings of Interspeech* Lyon, France. 345–348.

[3] Andreeva, B., Barry, W. J., Pützer, M., Tanchev, A. 2015. L2 stressed vowel production by Bulgarian learners of German. *Proceedings of the XVIIIth ICPhS* Glasgow. 350–354.

[4] Aylett, M., Turk, A. 2006. Language redundancy predicts syllabic duration and the spectral characteristics of vocalic syllable nuclei. *Journal of the Acoustical Society of America* 119(1), 30–48.

[5] Baroni, M., Kilgarriff, A. 2006. Large linguistically-processed web corpora for multiple languages. *Proceedings of EACL* Stroudsburg, PA, USA. Association for Computational Linguistics 87–90.

[6] Boersma, P., Weenink, D. 2017. Praat: doing phonetics by computer.

[7] Bradlow, A. R., Torretta, G. M., Pisoni, D. B. 1996. Intelligibility of normal speech. I. Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication* 20, 255–272.

[8] Chan, D., Fourcin, A., Gibbon, D., Granstrom, B., Huckvale, M., Kokkinakis, G., Kvale, K., Lamel, L., Lindberg, B., Moreno, A., Mouropoulos, J., Senia, F., Trancoso, I., Veld, C., Zeiliger, J. September 1995. Eurom - a spoken language resource for the EU. *Proceedings of Eurospeech* Madrid. 867–870.

[9] Clopper, C. G., Pierrehumbert, J. B. 2008. Effects of semantic predictability and regional dialect on vowel space reduction. *The Journal of the Acoustical Society of America* 124(3), 1682–1688.

[10] Jurafsky, D., Bell, A., Gregory, M., Raymond, W. D. 2001. Probabilistic relations between words: Evidence from reduction in lexical production. In: Bybee, J., Hopper, P., (eds), *Frequency and the Emergence of Linguistic Structure*. Amsterdam: Benjamins 229–254.

[11] Malisz, Z., Brandt, E., Möbius, B., Oh, Y., Andreeva, B. 2018. Dimensions of segmental variability: interaction of prosody and surprisal in six languages. *Frontiers in Communication – Language Sciences* 3, 1–18. article 25.

[12] Möhler, G., Schweitzer, A., Breitenbücher, M., Barbisch, M. 2000. IMS German Festival (Version: 1.2-os).

[13] Munson, B. 2007. Lexical access, lexical representation, and vowel production. *Laboratory phonology* 201–228.

[14] Munson, B., Solomon, N. P. 2004. The effect of phonological neighborhood density on vowel articulation. *Journal of Speech, Language, and Hearing Research* 47, 1048–1058.

[15] Pätzold, M., Simpson, A. P. 1997. Acoustic analysis of German vowels in the Kiel Corpus of Read Speech. *The Kiel Corpus of Read/Spontaneous Speech Acoustic data base, processing tools and analysis results. Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel (AIPUK)* 32, 215–247.

[16] Pierrehumbert, J. B. 2000. Exemplar dynamics: word frequency, lenition and contrast. In: Bybee, J., Hopper, P., (eds), *Frequency effects and the emergence of linguistic structure*. Amsterdam: John Benjamins 1–19.

[17] Scarborough, R. 2010. Lexical and contextual predictability: Confluent effects on the production of vowels. In: Fougeron, C., Kuehnert, B., Imperio, M., Vallee, N., (eds), *Laboratory phonology 10* volume 10. Berlin: Mouton de Gruyter 557–586.

[18] Schulz, E., Oh, Y. M., Malisz, Z., Andreeva, B., Möbius, B. 2016. Impact of prosodic structure and information density on vowel space size. *Proceedings of Speech Prosody 2016* Boston. 350–354.

[19] Simpson, A. P., Ericsdotter, C. 2007. Sex-specific differences in f0 and vowel space. *Proceedings of the XVIth ICPhS* number August Saarbrücken. 6–10.

[20] Stolcke, A. 2002. SRILM - an extensible language modeling toolkit. *Proceedings of Interspeech* 2(Denver, Colorado), 901–904.

[21] Turner, G. S., Tjaden, K., Weismer, G. 1995. The influence of speaking rate of vowel working space and intelligibility for individuals with amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research* 38, 1001–1013.

[22] Wedel, A., Nelson, N., Sharp, R. 2018. The phonetic specificity of contrastive hyperarticulation in natural speech. *Journal of Memory and Language* 100, 61–88.

[23] Weirich, M., Simpson, A. P. 2014. Differences in acoustic vowel space and the perception of speech tempo. *Journal of Phonetics* 43(1), 1–10.

[24] Weiss, B. 2007. Rate dependent vowel reduction in German. *Proceedings of the 12th SPECOM* Moscow.

[25] Wright, R. 2004. Factors of lexical competition in vowel articulation. In: Local, J. J., Ogden, R., Temple, R., (eds), *Papers in Laboratory Phonology VI*. Cambridge: Cambridge University Press 75–87.