

INDIVIDUAL DIFFERENCES IN VOWEL COMPACTNESS PERSIST UNDER INTOXICATION ACROSS FIRST AND SECOND LANGUAGES

Charles B. Chang, Kevin Tang, & Andrew Nevins*

Boston Univ.; Heinrich-Heine-Universität Düsseldorf & Univ. of Florida; Univ. College London
cbchang@post.harvard.edu, Kevin.Tang@hhu.de, a.nevins@ucl.ac.uk

ABSTRACT

Alcohol intoxication facilitates inhibition of one's first language (L1) ego, which may lead to reduced individual differences among second language (L2) speakers under intoxication. This study examined whether, compared to speaking while sober, speaking while intoxicated would reduce individual differences in the acoustic compactness of vowel categories in sequential bilinguals exemplifying diverse L1–L2 pairs (German–English, Korean–English). Vowel compactness in $F_1 \times F_2$ space varied by language (German, Korean, English) and by vowel, and was generally lower in intoxicated compared to sober speech, both across languages and throughout a bilingual's language repertoire. Crucially, however, there was still a wide range in compactness under intoxication; furthermore, individuals with more compact vowels while sober also produced more compact vowels while intoxicated, in both L1 and L2. Taken together, these findings show patterned variability of vowel compactness, suggesting that articulatory precision is an individual-difference dimension that persists across speaking conditions and throughout the repertoire.

Keywords: compactness, individual differences, vowels, bilingualism, alcohol.

1. INTRODUCTION

How does drinking affect one's speech? A growing literature on alcohol intoxication and spoken language production suggests that there may not be a simple answer to this question. On the one hand, several studies have found, consistent with the general decline of motor control under intoxication, that intoxication degrades speech, in terms of more errors, less precise gestural coordination, more consonant lenition, slower speech rate, greater pitch vari-

ability, and/or lower intelligibility [1, 2, 3, 4, 5, 6, 7, 8]. On the other hand, studies that have examined intoxication effects on speech in a bilingual's second language (L2), in isolation or in comparison to speech in their first language (L1), have often observed that intoxication affects L2/nonnative speech differently from L1 speech [7, 9]. For example, intoxication improved L1 German speakers' pronunciation in Dutch [10], which can be explained in terms of intoxication making a speaker's "language ego" less L1-biased, facilitating nonnative speech [9].

The possibility of facilitative intoxication raises interesting questions for L2 speech research. For one, if intoxication improves L2 production, might it "level the playing field" among L2 learners, who often show marked individual differences even from the same L1 background [11, 12, 13, 14, 15]? L2 speech research has been increasingly concerned with accounting for individual differences, linking them to factors such as cue-weighting preferences, memory variation, and perceptual and neural differences [16, 17, 18, 19, 20]. In this study, we focused on individual differences in *compactness*, the acoustic consistency of a phonetic category's production [21], thought to reflect the "consistency of [speakers'] phonological-motor mapping" [22, p. 826].

Part of a larger project examining intoxication effects on bilingual speech, the current study asked how individual differences among bilinguals in vowel compactness would be affected by intoxication, across L1 and L2. Based on findings of intoxication degrading L1 but facilitating L2 speech, we hypothesized that intoxication would result in less compact L1 vowels, yet—by enhancing access to L2 phonological-motor mappings—more compact L2 vowels. Given the persistence of variation in compactness across L1 and L2 [21, 22], intoxication-induced changes in compactness were not predicted to reduce individual differences in L1; that is, habitually "less compact" speakers were predicted to become even less compact under intoxication along with "more compact" speakers. In contrast, if intoxication does indeed have an L2-specific facilitation effect (and there is an upper limit on compactness), it could allow less-compact speakers to "catch up"

*We gratefully acknowledge the data contributions and research assistance of Kai Xin Bao, Marissa Carl, Erik Duchowski, Jenny Geng, Sam Green, Michael Hindley, Young Shin Kim, Peyton Krinsky, Ji Hye Kwon, Wen Jia Liu, Eliana Mugar, Yin Wang, Yifan Wu, and Steven Zhang, as well as the helpful feedback of two anonymous reviewers.

to more-compact speakers in terms of compactness, leading to reduced individual differences in L2. To explore these predictions, we compared the sober and intoxicated speech of bilinguals exemplifying diverse L1–L2 pairs.

2. METHODS

2.1. Participants

Participants comprised two groups of sequential bilinguals: Germans (L1 German, L2 English; $N = 8$; 4f, 4m; $M_{age} = 27.1$ yr) and Koreans (L1 Korean, L2 English; $N = 8$; 8f, 0m; $M_{age} = 27.1$ yr). The groups were born and raised in Germany or South Korea, respectively, and were tested in the UK, where most participants had been living for 1–2 years for study abroad. They reported no history of hearing or speech problems, and their average scores on the IELTS [23] (7.0 band) and LexTALE [24] (60s), which did not differ significantly between groups, suggested that their English proficiency was generally “upper intermediate” (B2) or higher.

2.2. Materials

Speech materials were compiled for each language on the basis of published dialogues (e.g., from a play), which were edited for emotional and gender neutrality, turn length, and archaisms. The German dialogue contained a total of 55 turns; the Korean dialogue 123 turns; and the English dialogue 95 turns. The full materials with source references are available open-access at osf.io/y2r87/.

2.3. Procedure

The main task was a dialogue-reading task, which participants completed in a sound-insulated room four times (2 languages \times 2 speaking conditions) across two test sessions spaced no more than 14 days apart. Each dialogue was uttered in two speaking conditions (sober, intoxicated), completed on separate days in the order sober–intoxicated in the Korean group and in counterbalanced order in the German group. Participants were told not to eat, drink, or use mouthwash starting 2 h before each session and not to smoke starting 0.5 h before.

Before beginning the dialogue-reading task, participants were instructed to read the dialogues naturally. Seated in front of a microphone facing the experimenter, the participant went through each dialogue with the experimenter, each person reading one character’s lines. The session was audio-recorded at 44.1 kHz with 16-bit resolution in stereo;

the recordings were then converted to mono.

Participants’ BAC was tested in both conditions using a breathalyzer (AlcoMate Premium AL-7000). First, BAC was measured to ensure that participants had no alcohol in their system. In the intoxicated condition, the target BAC was 0.12%. To reach this BAC, participants consumed, at their own pace, an amount of alcohol (vodka or rum, mixed with juice) determined on the basis of their self-reported weight and BAC charts [25]. BAC was tested 15 min after about 75% of the mixture was consumed and then every 3–5 min until it went over 0.12% and dropped back down to 0.12%; if needed, a small top-up was given from the remaining alcohol. Once BAC had reached 0.12%, participants were taken into the sound-insulated room for the dialogue-reading task.

2.4. Analysis

Forced-aligned TextGrids (produced using the Montreal Forced Aligner [26]) were hand-corrected and then submitted, along with the audio recordings, to acoustic analysis in Praat [27]. For each of the vowel intervals marked in the TextGrids, the analysis extracted both the first and second formants (F_1 , F_2) over the middle 50% of the vowel’s duration using the default settings of Praat’s linear predictive coding method; the maximum formant parameter was set to 5,000 Hz for males and 5,500 Hz for females.

Because German and English, in contrast to Korean, are described as having unstressed vowel quality reduction, the final dataset for statistical analysis included only primary-stressed monophthongs in German and English, as well as the monophthongs of Korean. The F_1 and F_2 measures on this set of vowels were checked for errors separately by participant and by language, using the 3- SD criterion for outliers; tokens with one or more errors (comprising 1.2% of the dataset) were removed. Formant measures on the remaining tokens ($N = 64,382$ total; 9,081 in German, 27,450 in Korean, 27,851 in English) were then Nearey-normalized [28] using the `vowels` package [29] in R [30]. The full dataset is available open-access at osf.io/s9vkw/.

The normalized formant measures were analyzed both qualitatively and quantitatively. First, we inspected vowel plots to observe variation in vowel compactness among participants and languages and within the repertoire (L1, L2). Then we calculated a metric of vowel area (i.e., the inverse of compactness) in terms of the area of the 95% confidence ellipse around tokens. Finally, we built a linear mixed-effects model on the vowel area data, using the `lmerTest` package [31]. The model included fixed effects for Group (treatment-coded; ref = German),

Condition (treatment-coded; ref = sober), RepLang (treatment-coded; ref = L1), Vowel (sum-coded; all monophthongs including German /i e ε a u o y ʏ ø œ/, Korean /i e ε u o ʌ i/, and English /i e ε æ u u o ɔ ʌ ɹ/), all interactions among Group, Condition, and RepLang, a Condition \times Vowel interaction, as well as random intercepts by Participant.¹

3. RESULTS

Vowel area varied significantly by language, with tokens of a vowel category generally more diffusely (i.e., less compactly) distributed in the Korean monophthongs than the German monophthongs in sober speech [$\beta = 0.226, t = 4.575, p < .001$]. On the other hand, vowel area was smaller in the German group's L2 English monophthongs as compared to their L1 German [$\beta = -0.122, t = -4.019, p < .001$], and this L2 effect was even stronger in the Korean group [$\beta = -0.138, t = -3.161, p = .002$].²

Examination of individual differences in vowel area revealed substantial variation in both groups. Table 1 summarizes the ranges in average vowel area by group, language, and condition. Although the minimum and maximum values within each group are different (consistent with the language effect above), there is a wide range in every case.

Group	Language, Condition	Min	Max
German	German, sober	0.303	0.500
German	German, intoxicated	0.316	0.637
German	English, sober	0.204	0.487
German	English, intoxicated	0.308	0.497
Korean	Korean, sober	0.556	0.885
Korean	Korean, intoxicated	0.565	0.953
Korean	English, sober	0.332	0.483
Korean	English, intoxicated	0.308	0.515

Table 1: Range in average vowel area (min, max values) in Nearey-normalized $F_1 \times F_2$ space.

Examples of individual differences are depicted in Fig. 1–4, which show the vowel spaces for some of the most-compact (G005, K004) and least-compact (G006, K006) participants by language and condition on the same scale, with 95% confidence ellipses around tokens of the same vowel (plots for all participants are at osf.io/s9vkw/). Note the marked difference in compactness between G005 and G006 and between K004 and K006, as well as the broad similarity among the four panels in each figure.

As seen in Fig. 1–4, the area for a given vowel tended to be larger in intoxicated than in sober speech for all languages. Thus, instead of showing an L2-specific effect (in the interaction coefficients),

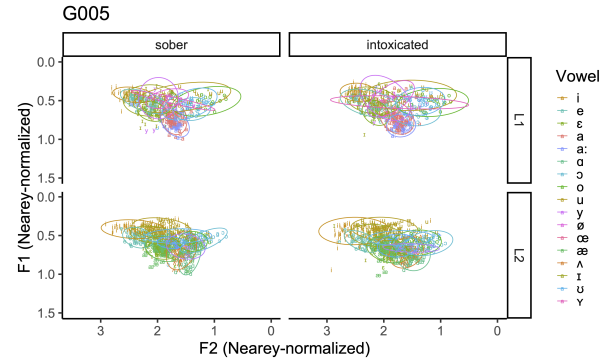


Figure 1: Vowel spaces by language (L1 German, L2 English) and condition in participant G005.

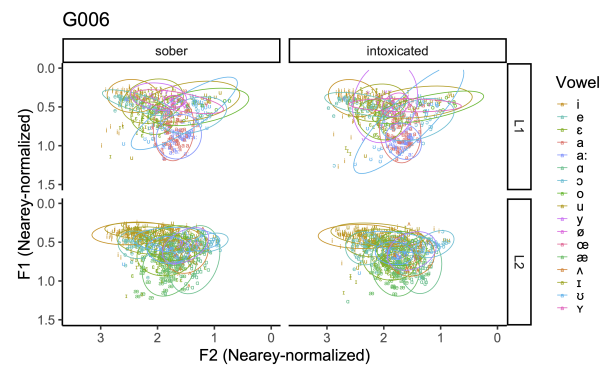


Figure 2: Vowel spaces by language (L1 German, L2 English) and condition in participant G006.

the model showed that the marginal intoxication effect in German [$\beta = 0.049, t = 1.743, p = .082$] did not differ in Korean [$\beta = 0.030, t = 0.625, p = .532$] or L2 English [$|\beta|s < 0.099, |t|s < 1.586, ps > .1$]. That is, we found no evidence of L2-specific facilitation of vowel production. Rather, the intoxication effect was general, resulting in expanded vowel areas at the group mean level in L1 (Germans: $M_{int} = 0.460, M_{sob} = 0.391$; Koreans: $M_{int} = 0.743, M_{sob} = 0.677$) and L2 (Germans: $M_{int} = 0.395, M_{sob} = 0.316$; Koreans: $M_{int} = 0.414, M_{sob} = 0.404$).

In order to further examine the relationship between vowel areas in intoxicated and sober speech at the individual level, we correlated participants' by-language average vowel areas for the intoxicated condition against those for the sober condition. The two sets of values were highly correlated overall [$r(30) = 0.926, t = 13.448, p < .001$], in German [$r(6) = 0.747, t = 2.752, p = .033$], in Korean [$r(6) = 0.902, t = 5.120, p = .002$], and in the German group's English [$r(6) = 0.845, t = 3.866, p = .008$], although not in the Korean group's English [$r(6) = 0.480, t = 1.342, p = .228$]. Thus, in general, bilinguals with more compact vowels while

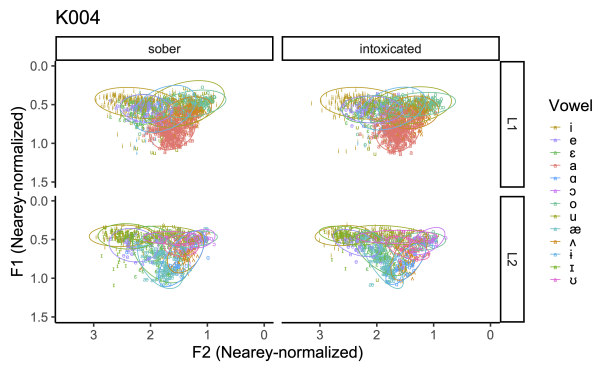


Figure 3: Vowel spaces by language (L1 Korean, L2 English) and condition in participant K004.

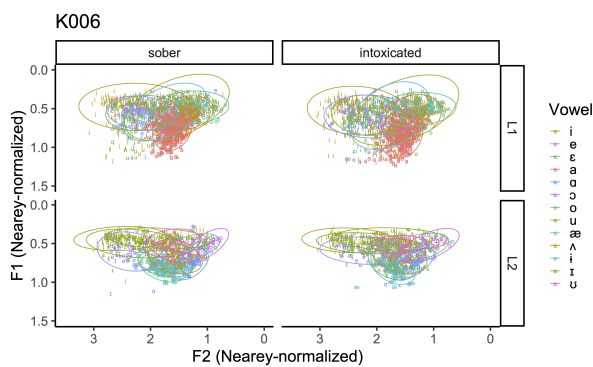


Figure 4: Vowel spaces by language (L1 Korean, L2 English) and condition in participant K006.

sober also produced more compact vowels while intoxicated, whether speaking in L1 or L2.

Finally, there were significant differences among vowels in terms of area, but not in terms of intoxication effects. Compared to the average vowel area over all vowels, the vowels /i ɪ u ʊ/ (in German and English), /y ʌ:/ (in German), /i/ (in Korean), and /ɑ/ (in English) were significantly more diffusely distributed, whereas /e ε o/ (in German, Korean, and English), /a/ (in German and Korean), /ɔ/ (in German and English), /ʌ/ (in Korean and English), /ø œ/ (in German), and /æ/ (in English) were significantly less diffusely distributed (i.e., more compact). Crucially, none of the Condition \times Vowel coefficients were significant, meaning that different vowels did not vary significantly in terms of how their vowel area was affected by intoxication.

4. GENERAL DISCUSSION

Taken together, the results of this study do not support the hypothesis of an L2-specific facilitative effect of intoxication on speech production. Contrary to expectation, intoxication decreased vowel com-

pactness in L2 (English), just as it did in L1 for both German and Korean. Intoxication-induced changes in compactness also left a wide range in compactness at the individual level (Table 1). Individual variation in compactness, moreover, tended to be correlated between sober and intoxicated speech, in both L1 and L2. These findings provide evidence that individual differences in compactness persist across speaking conditions and throughout a bilingual’s repertoire, consistent with the view that the articulatory precision reflected in compactness is a stable dimension of individual differences.

The contrast between our results and those of [10] invites the question of why [10] observed facilitative effects of intoxication on L2 speech but we did not. There are two possible, and mutually compatible, explanations. First, our L2 speakers may have been more proficient than those in [10]. Second, our study analyzed a specific, objective aspect of segmental production, whereas [10] focused on a subjective measure of overall pronunciation. What we take away from these differences is that the effects of intoxication on speech may be multifaceted and variable over development. For example, intoxication may worsen acoustic *consistency* at the segmental level, but it may not necessarily affect *proximity* to native norms, and it may well improve suprasegmentals that contribute to perceived pronunciation quality, none of which were tested in this study.

While this study found persistent individual differences in the acoustic consistency of vowels, the findings are limited in a few ways that leave open several directions for future research. For one, as mentioned above, we examined only compactness, so we cannot say for sure whether more-compact speakers were also more “on target” vis-a-vis norms than less-compact speakers. Thus, further work investigating central tendencies of L2 phonetic categories in conjunction with their distributions would provide a fuller picture of individual differences in L2 speech. In addition, for reasons of space, we were only able to discuss the range of individual differences in compactness, glossing over potential variation in changes in compactness by vowel. In future research, it would therefore be insightful to examine, for example, whether speakers showing average compactness pattern differently from the most-compact and/or least-compact speakers in terms of intoxication effects on specific vowels. Finally, our results showed suggestive effects of language (German vs. Korean) and language status (L1 vs. L2) on compactness, which could be usefully elaborated in future studies crossing languages with different-size vowel inventories as L1 and L2.

5. REFERENCES

- [1] Trojan, F., Kryspin-Exner, K. 1968. The decay of articulation under the influence of alcohol and paraldehyde. *Folia Phoniatrica et Logopaedica* 20, 217–238.
- [2] Beam, S. L., Gant, R. W., Mecham, M. J. 1978. Communication deviations in alcoholics; A pilot study. *Journal of Studies on Alcohol and Drugs* 39, 548–551.
- [3] Chin, S. B., Pisoni, D. B. 1997. *Alcohol and Speech*. San Diego: Academic Press.
- [4] Pisoni, D. B., Martin, C. S. 1989. Effects of alcohol on the acoustic-phonetic properties of speech: Perceptual and acoustic analyses. *Alcoholism: Clinical and Experimental Research* 13(4), 577–587.
- [5] Johnson, K., Pisoni, D. B., Bernacki, R. H. 1990. Do voice recordings reveal whether a person is intoxicated? A case study. *Phonetica* 47, 215–237.
- [6] Hollien, H., DeJong, G., Martin, C. A., et al. 2001. Effects of ethanol intoxication on speech suprasegmentals. *JASA* 110, 3198–3206.
- [7] Wieling, M., Blankevoort, C., Hukker, V., et al. 2019. The influence of alcohol on L1 vs. L2 pronunciation. In: Calhoun, S., Escudero, P., Tabain, M., Warren, P. (eds), *Proceedings of the 19th ICPHS*. Canberra: ASSTA Inc., 3622–3626.
- [8] Tang, K., Chang, C. B., Green, S., et al. 2022. Intoxication and pitch control in tonal and non-tonal language speakers. *JASA Express Lett.* 2, 065202.
- [9] Guiora, A. Z., Beit-Hallahmi, B., Brannon, R. C. L., et al. 1972. The effects of experimentally induced changes in ego states on pronunciation ability in a second language: An exploratory study. *Comprehensive Psychiatry* 13, 421–428.
- [10] Renner, F., Kersbergen, I., Field, M., et al. 2018. Dutch courage? Effects of acute alcohol consumption on self-ratings and observer ratings of foreign language skills. *Journal of Psychopharmacology* 32, 116–122.
- [11] Dörnyei, Z., Skehan, P. 2003. Individual differences in second language learning. In: Doughty, C., Long, M. (eds), *The Handbook of Second Language Acquisition*. Malden: Blackwell, 589–630.
- [12] Chandrasekaran, B., Sampath, P. D., Wong, P. C. M. 2010. Individual variability in cue-weighting and lexical tone learning. *JASA* 128, 456–465.
- [13] Perrachione, T. K., Lee, J., Ha, L. Y. Y., et al. 2011. Learning a novel phonological contrast depends on interactions between individual differences and training paradigm design. *JASA* 130, 461–472.
- [14] Bowles, A. R., Chang, C. B., Karuzis, V. P. 2016. Pitch ability as an aptitude for tone learning. *Language Learning* 66, 774–808.
- [15] Chang, C. B., Kwon, S. 2020. The contributions of crosslinguistic influence and individual differences to nonnative speech perception. *Languages* 5, 49.
- [16] Baker, W., Trofimovich, P. 2006. Perceptual paths to accurate production of L2 vowels: The role of individual differences. *IRAL* 44, 231–250.
- [17] Raizada, R. D. S., Tsao, F.-M., Liu, H.-M., et al. 2010. Quantifying the adequacy of neural representations for a cross-language phonetic discrimination task: Prediction of individual differences. *Cerebral Cortex* 20, 1–12.
- [18] Idemaru, K., Holt, L. L., Seltman, H. 2012. Individual differences in cue weights are stable across time: The case of Japanese stop lengths. *JASA* 132, 3950–3964.
- [19] Darcy, I., Park, H., Yang, C.-L. 2015. Individual differences in L2 acquisition of English phonology: The relation between cognitive abilities and phonological processing. *Learning and Individual Differences* 40, 63–72.
- [20] Schertz, J., Cho, T., Lotto, A. J., et al. 2016. Individual differences in perceptual adaptability of foreign sound categories. *Attention, Perception, and Psychophysics* 78, 355–367.
- [21] Kartushina, N., Frauenfelder, U. H. 2014. On the effects of L2 perception and of individual differences in L1 production on L2 pronunciation. *Frontiers in Psychology* 5, 1246.
- [22] Kartushina, N., Hervais-Adelman, A., Frauenfelder, U. H., et al. 2015. The effect of phonetic production training with visual feedback on the perception and production of foreign speech sounds. *JASA* 138, 817–832.
- [23] British Council, IDP: IELTS Australia, Cambridge English. 2023. IELTS in CEFR scale. <https://www.ielts.org/about-ielts/ielts-in-cefr-scale>.
- [24] Lemhöfer, K., Broersma, M. 2012. Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods* 44, 325–343.
- [25] Miller, W. R., Muñoz, R. F. 1982. *How to Control Your Drinking*. Englewood Cliffs: Prentice Hall.
- [26] McAuliffe, M., Socolof, M., et al. 2017. Montreal Forced Aligner: Trainable text-speech alignment using Kaldi. In: *INTERSPEECH-2017*, 498–502.
- [27] Boersma, P., Weenink, D. 2022. Praat: Doing phonetics by computer. Version 6.2.23, praat.org.
- [28] Nearey, T. M. 1977. *Phonetic Feature Systems for Vowels*. PhD thesis, University of Alberta.
- [29] Kendall, T., Thomas, E. R. 2018. vowels: Vowel manipulation, normalization, and plotting. R package version 1.2-2. Available from <https://cran.r-project.org/web/packages/vowels/>.
- [30] R Development Core Team, 2022. R: A language and environment for statistical computing. Version 4.2.1, www.r-project.org.
- [31] Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B. 2017. lmerTest package: Tests in linear mixed effects models. *J. Stat. Softw.* 82, 1–26.

¹ Random slopes for Condition by Participant were explored but not included because they caused the model to be singular. Results remained the same regardless.

² The full model output showing all fixed-effect coefficients is available at osf.io/s9vkw/.