ACOUSTIC MEASURES OF NON-NATIVE ADDRESSEE REGISTER FOR MID TO HIGH PROFICIENT ENGLISH LEARNERS OF GERMAN

Bianca Sell, Megumi Terada, Malte Belz, Christine Mooshammer

Humboldt-Universität zu Berlin, Germany
bianca.maria.sell@hu-berlin.de

ABSTRACT

Up to now, most studies investigating speech adaption towards non-native addressees focused on L2 speakers with a low proficiency of English. Here, we examine task-based conversations when speaking to mid to high proficient L2 speakers of German. We compare articulation rate, vowel duration, and vowel space area of 20 German L1 speakers when conversing with L1 German speakers and with L2 German speakers of L1 English. As expected, we find no effect of addressee on vowel duration, while articulation rate is slower for non-native addressees, regardless of proficiency, compared to native addressees. However, vowel space area does not change for different addressees, contrary to enlargement reported for English non-native addressee register. The results will be discussed with respect to L2 proficiency and phonetic convergence.

Keywords: Foreigner-directed speech, German, hyperarticulation, register

1. INTRODUCTION

Speakers adapt their linguistic behaviour to the specific demands of the communicative situation – they employ different registers [1]. Besides situational variables such as setting or topic, the addressee's characteristics are another factor influencing register choices. Previous studies [2, 3, 4] show that speakers accommodate their linguistic behaviour to the (presumed) traits of the addressee. One of these traits is the addressee's linguistic background: Speakers shift into another register when communicating with a non-native interlocutor, which is called *non-native addressee register* (NNAR) or *foreigner directed speech* (FDS) [5]. On the phonetic level, NNAR features lower articulation rates, louder speech, and hyperarticulations [6, 7, 5]. The latter has often been accessed by the vowel space area (VSA), i.e. the area spanned by the first and second format of the corner vowels /iː, aː, uː/, and was mainly found for English NNAR, probably as a means to extend the perceptual distance of vowel categories in order to avoid confusion between them [5]. Vowel space extension was not consistently found in all languages investigated (e.g., not for Arabic NNAR, see [8]) and might be dependent on the size of the vowel inventory [5]. Results on VSA extension remain inconclusive and mostly focussed on the English language. As has been found by [9, 10], hyperarticulation in vowels is not related to vowel length. A much more consistent adaptation towards non-native speakers is a lower articulation rate, cf. [9, 7, 11]. While [9] report that adaptations in speech rate are mostly a result of longer pauses instead of a slower articulation rate, [11] find a longer mean word duration for English NNAR.

For German, research on NNAR has been focused mostly on qualitative parameters adapted towards low proficiency addressees [12, 6]. It is, however, difficult to differentiate in these studies between the addressee's low-level proficiency, the lower prestige of the addressee's L1, and possible power imbalances that might create a negative bias influencing linguistic behaviour [12]. It remains largely unknown what the features of German NNAR are when elicited by mid to high proficiency addressees whose L1 is (ideally) not subject to a negative bias. In the present study, we address this issue: We investigate whether German L1 speakers adapt their vowel space area (VSA), vowel duration (VD), and articulation rate (AR) when talking to non-native addressees with mid to high proficiency levels in German, with English as their L1 (considered as presitigous in Germany [13]), when compared to the same setting, but with German L1 interlocutors. We additionally assess the influence of the addressee's proficiency (self-reported mid vs. high proficiency).

For this purpose, we conduct a corpus-based study comparing participants' VSA, VD, and AR when talking to German native addressees or non-native addresses with mid to high proficiency. We expect to find an increased VSA when speakers are conversing with non-native interlocutors with a stronger effect for interlocutors with a lower proficiency (Hyp. 1). Likewise, it is predicted that the participants decrease their AR when speaking with non-native interlocutors, with a stronger effect for lower L2 proficiency (Hyp. 2), as shown in previous studies [5]. We do not expect any adaptions concerning vowel duration (Hyp. 3), as indicated by [9, 10]. The present study thus allows us to gain a better understanding of the phonetic level of NNAR in German for mid-to-high proficiency interlocutors.

2. METHOD

2.1. Corpus and annotation

In the Corpus of Non-Native Addressee Register (CoNNAR) we compare the data of 20 German L1 speakers (10 female; age 20–38 years, mean age = 26 years, sd = 4.5 years) who passed through our experiment twice with instructed interlocutors (confederates): Once with another German L1 speaker (L1 confederate, n = 4) and once with an English L1 speaker (L2 confederate, n = 4) with mid to high proficiency in German (self reported proficiency levels of B1 to C1 CEFR [14]). Two of the L2 confederates originate from the US, two from Southern England. Both varieties of English can be considered similarly familiar and prestigious in German [15]. L1 and L2 confederates were matched in age and gender, creating confederate pairs differing in the non-natives' German proficiency. Each confederate conversed with five experimental participants. The order of L1 and L2 confederate experiments was counterbalanced.

Each experimental session (n = 40) consisted of a word list with the sentence *Sage X bitte* 'Say X please' with X being a bisyllabic word containing all monophthongs of German in stressed position. The word list was read in a break-out room with only the participant and the experimenter. Next, participants conversed for 8 minutes about a topic of their choice and then solved two Diapix tasks, 8 minutes each (a spot-the-difference task where both interlocutors are given slightly different versions of the same picture with the goal of finding all differences, [16, 17]). The original Diapix picture materials by Baker & Hazan [17] were translated into German and adapted to elicit as many corner vowels without being overly obvious [18]. In the second Diapix confederates were instructed to ask for a clearer pronunciation after 3 to 4 found differences. Free conversation and the Diapix tasks were held in break-out rooms without the experimenter. Lastly, participants read the word list again in a break-out room and filled out a questionnaire after the second session.

Due to the COVID-19 pandemic the interlocutors were placed in different rooms and connected via Zoom. The data was recorded using directional microphones (Sennheiser) via Audacity [19] in a stereo file. The recordings lasted about one hour per session, participants were paid 11€/h. The collected data was transliterated using Praat [20], the segmentation of sounds was carried out automatically using the WebMAUS service [21]. We measured the frequencies of the first and second formant and vowel durations for the target vowels [aː], [iː], and [uː] in stressed and accented position within content words. In the Diapix task, we focused on target words in the pictures' written parts. If participants produced less than 2 target vowels per Diapix from reading those written parts, additional target vowels in stressed and accented position were annotated manually, avoiding diphthongisation due to r-vocalisation or following vowels. In case there were not enough target words following these conditions syllables without word stress were also accepted (7 cases). All target vowels were annotated manually. The beginning of each vowel was marked at the rising zero crossing of the first complete period, its end at the zero crossing of F2 fading.

2.2. Data preparation and measurements

The annotated word lists and Diapix tasks were converted into an EMU database [22] with the R package *emuR* version 2.3.0 [23]. The Praat formant tracker was used to calculate formants [24], using gender-specific formant ceilings (5000 Hz for male, 5500 Hz for female speakers). All formant trajectories were checked and manually corrected for obvious tracking errors in the EMU database.

For measuring the three corner vowels /iː, aː, uː/, five measure points in the center of the vowel were extracted for F1 and F2 and normalised using Lobanov's normalisation method [25]. Based on the normalised F1 and F2, we calculated the VSA using the following formula for triangle area in a Euclidean plane: $VSA = |\frac{1}{2} * ((F1_i * (F2_a - F2_u) + F1_a * (F2_u - F2_i) + F1_u * (F2_i - F2_a))|$

For articulation rate the syllable number is automatically detected based on the orthographic transcription by using the R-package *sylly.de* version 0.1.2 [26]. The total number of syllables per speaker in each Diapix task was summed and divided by total length of articulation (in sec.) excluding silent pauses, turn transitions, and extra- and paralinguistic events such as laughing, clicks and background noises as well as pseudonymised tokens.

3. RESULTS

According to Hypothesis 1, we expected an increase in VSA for non-native addressees compared to L1 addressees with a stronger effect for mid proficient learners (B1-B2) than high proficient learners (C1).

As can be seen in Figure 1, our data does not confirm this hypothesis. Even though participants show a larger VSA when speaking to non-natives with B2 proficiency than to the L1 confederate (cf. L1B2 in Figure 1), the results did not show any significant effects of addressee or their L2 proficiency on the VSA, as evaluated by a linear mixed-effects model with normalised VSA as dependent variable and confederate pair (L1B1, L1B2, L1C1_1, L1C1_2) as well as language (L1 vs. L2) as independent variable, including random intercepts for participants.

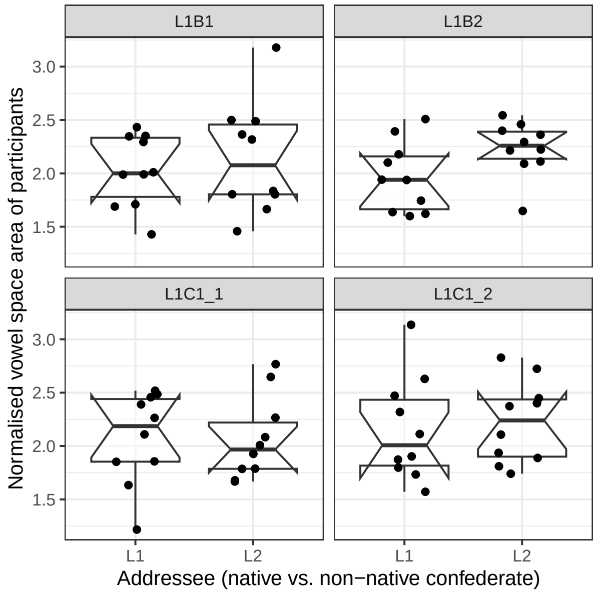

Figure 1: Normalised vowel space area of the L1 participants per confederate pair (L1B1, L1B2, L1C1_1, L1C1_2) and addressee (L1 vs. L2) for each of the two Diapix tasks.

In Hyp. 2 we expected to find a slower AR for non-native addressees compared to L1 addressees with a stronger effect for mid than high proficient learners. Figure 2 shows the AR per speaker sorted by confederate pairs. It can be seen that participants tend to decrease their AR with non-native interlocutors with mid level L2 proficiency (B1, B2) compared to L1 interlocutors. These speakers tend to behave similarly though individual speakers deviate from the described pattern. In contrast, participants maintain or increase their AR when conversing with advanced L2 speakers (C1_1, C1_2). We calculated a linear mixed-effects model with AR as dependent variable and confederate pair (L1B1, L1B2, L1C1_1, L1C1_2) as well as language (L1 vs. L2) as independent variable, including random intercepts for participants. Participants speak slower when talking to an L2 confederate than to an L1 confederate ($\beta = .1, se = .05, p < .05$). There was no significant effect for confederate pair.

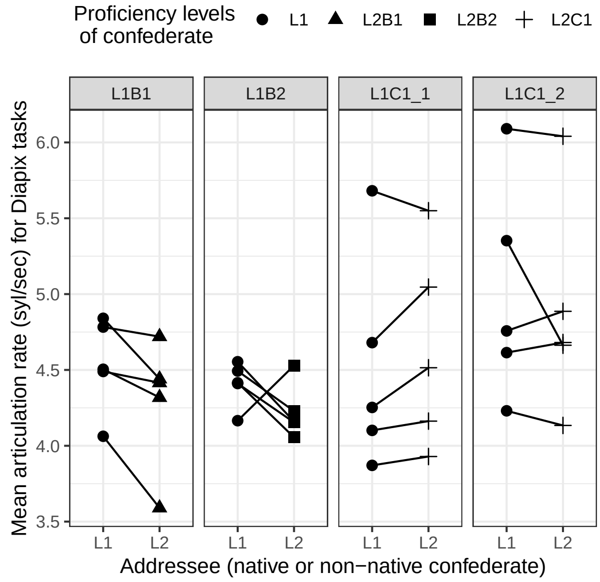

Figure 2: Mean articulation rate of L1 participants per confederate pair (L1B1, L1B2, L1C1_1, L1C1_2) and addressee (L1 vs. L2) for all Diapix tasks

Figure 3 shows a scatterplot with regression lines of articulation rates and VSA per participant sorted by the interlocutors' proficiency (L2B1-2 vs. L2C1 vs. L1). The correlations of normalised VSA and AR for every proficiency level did not reach significance, as calculated by a Pearson's correlation test.

Figure 4 shows boxplots for vowel duration per target vowel and proficiency. Speakers do not change vowel duration depending on the addressee. For none of the target vowels /iː aː uː/ the differences in duration based on the confederates' L1 or L2 proficiency reached significance, based on a linear mixed-effects model with duration as dependent variable and vowels as well as confederate pairs as independent variables, with participants as random intercepts. The null result of vowel duration confirm Hypothesis 3.

4. DISCUSSION

We investigated whether speakers change their articulation rate, vowel space area, and vowel duration when conversing with a non-native addressee with mid to high German L2 proficiency compared to German L1 addressees. In summary, VSA and vowel duration show no effect of the addressee while articulation rate is slower for non-native addressees compared to L1 addressees independent of

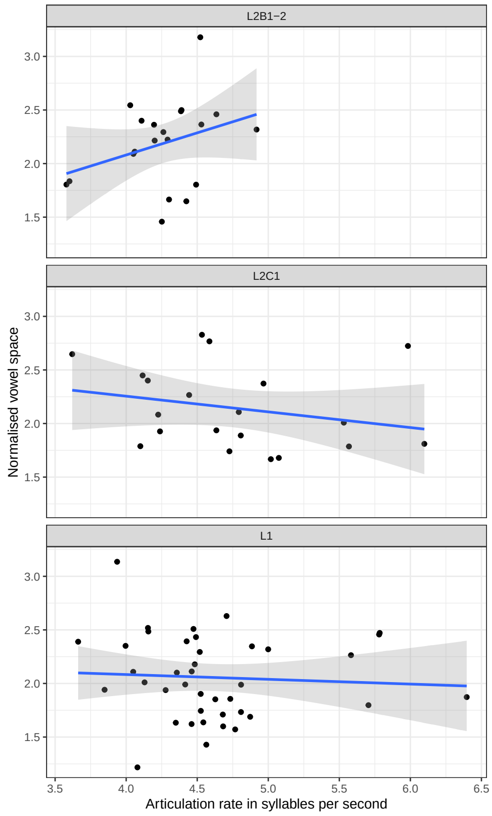

Figure 3: Vowel space area and articulation rate in both Diapix tasks per participant speaking to L1 or L2 confederates with varying proficiency levels.

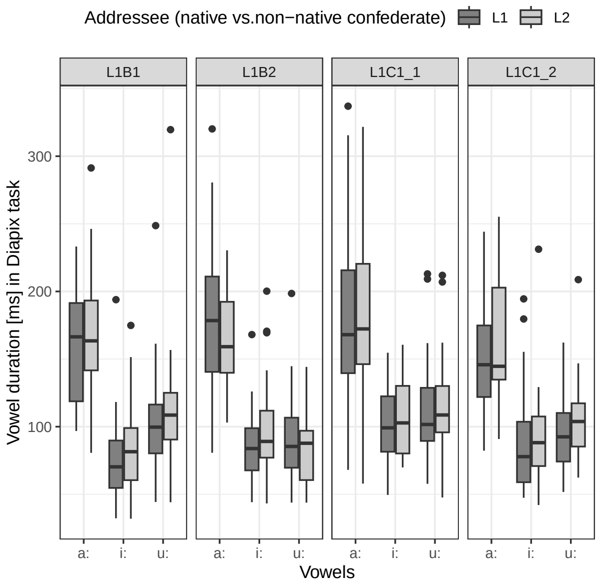

Figure 4: Duration of target vowels in Diapix tasks per confederate pair and the non-native's L2 proficiency.

the non-native addressees' proficiency. [27] report that hyperarticulation, measured as VSA extension, might be language-specific: A larger vowel inventory might lead to more hyperarticulations in order to minimise overlap between categories. German and English have a similarly large vowel inventories [28] (with 15 full vowels in German [29]). However, while hyperarticulations are regularly reported for English [5, 7], we find none for German. The absence of VSA extension in our data questions hyperarticulation as a feature dependent on the vowel inventory – at least in NNAR for mid to high proficiency addressees. Concerning vowel duration, our results confirm previous studies [9, 10] that find no influence of the addressee on vowel duration. As indicated by [11], articulation rate tended to be slower for non-native addressees though we do not find an effect of the non-natives' proficiency. This could be due to the addressees' higher proficiency or the fact that we examine *articulation rate* instead of *speech rate* (where slower speech rate was a consequence of longer pauses rather than vowel lengthening). It is possible that pauses are longer depending on the non-natives' proficiency in German NNAR (similarly to English NNAR, cf. [9]). The differences between our results and previous studies could stem from varying situational factors such as the absence of an actual interlocutor in favour of an imaginary one in [9] which [30] found to result in a stronger effect than a natural interlocutor, non-native addressees with different L1 or proficiency levels that might be subject to a bias [11], or different conversational tasks (e.g., Diapix task compared to free conversations [10]).

For a more in depth analysis, more data on German NNAR are necessary to statistically confirm the tendencies found in the present study. Additionally, variation in AR could be due to phonetic convergence: [31] report that Dutch L1 speakers converge to L2 speakers (inter alia in speech rate) though the effect was stronger for non-native addressees with lower accent ratings. This indicates that our participants' AR measurements might be influenced by the confederates AR. Future studies should take alignment into account and include the confederates' data as well. Therefore, we are planning to include effects of phonetic convergence in our analyses.

FUNDING

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – SFB 1412, 416591334.

5. REFERENCES

[1] Lüdeling, A., Alexiadou, A., Adli, A., Donhauser, K., Dreyer, M., Egg, M., Feulner, A. H., Gagarina, N., Hock, W., Jannedy, S., Kammerzell, F., Knoeferle, P., Krause, T., Krifka, M., Kutscher, S., Lütke, B., McFadden, T., Meyer, R., Mooshammer, C., Müller, S., Maquate, K., Norde, M., Sauerland, U., Solt, S., Szucsich, L., Verhoeven, E., Waltereit, R., Wolfsgruber, A., Zeige, L. E. 2022. Register: Language users' knowledge of situational-functional variation. *Register Aspects of Language in Situation (REALIS)* 1(1), 1–58.

[2] Bell, A. 1984. Language style as audience design. *Language in Society* 13(2), 145–204.

[3] Campbell-Kibler, K. 2010. The sociolinguistic variant as a carrier of social meaning. *Language Variation and Change* 22(3), 423–441.

[4] Hay, J., Jannedy, S., Mendoza-Denton, N. 1999. Oprah and/ay: Lexical frequency, referee design and style. *Proc. 14th ICPhS*, 1389–1392.

[5] Piazza, G., Martin, C. D., Kalashnikova, M. 2022. The acoustic features and didactic function of foreigner-directed speech: A scoping review. *Journal of Speech, Language, and Hearing Research* 65(8), 2896–2918.

[6] Roche, J. 1998. Variation in xenolects (foreigner talk). *Sociolinguistica* 12(1).

[7] Bradlow, A. R., Bent, T. 2002. The clear speech effect for non-native listeners. *Journal of the Acoustical Society of America* 112(1), 13.

[8] Kendi, A. A., Khattab, G. 2019. Acoustic properties of foreigner directed speech. *Proc. 19th ICPhS*, 810–814.

[9] Biersack, S., Kempe, V., Knapton, L. 2005. Fine-tuning speech registers: A comparison of the prosodic features of child-directed and foreigner-directed speech. *Ninth European Conference on Speech Communication and Technology*, 2401–2004.

[10] Uther, M., Knoll, M. A., Burnham, D. 2007. Do you speak E-NG-LI-SH? A comparison of foreigner-and infant-directed speech. *Speech Communication* 49(1), 2–7.

[11] Hazan, V., Uther, M., Grunlund, S. 2015. How does foreigner-directed speech differ from other forms of listener-directed clear speaking styles? *Proc. 18th ICPhS*.

[12] Hinnenkamp, V. 1982. *Foreigner Talk und Tarzanisch: eine vergleichende Studie über die Sprechweise gegenüber Ausländern am Beispiel des Deutschen und des Türkischen*. Hamburg: Buske Verlag.

[13] Schroedler, T., Purkarthofer, J., Cantone, K. F. 2022. The prestige and perceived value of home languages. insights from an exploratory study on multilingual speakers' own perceptions and experiences of linguistic discrimination. *Journal of Multilingual and Multicultural Development*, 1–18.

[14] Council of Europe, 2001. *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge University Press.

[15] Meer, P., Hartmann, J., Rumlich, D. 09 2021. Attitudes of German high school students toward different varieties of English. *Applied Linguistics* 43(3), 538–562.

[16] Van Engen, K., Baese-Berk, M., Baker, R. E., Choi, A., Kim, M., Bradlow, A. R. 2010. The Wildcat Corpus of native-and foreign-accented English: Communicative efficiency across conversational dyads with varying language alignment profiles. *Language and Speech* 53(4), 510–540.

[17] Baker, R., Hazan, V. 2011. DiapixUK: task materials for the elicitation of multiple spontaneous speech dialogs. *Behavior Research Methods* 43, 761–770.

[18] Bullock Oliveira, M., Sell, B. PDF and PSD files of DiapixGEtv picture materials – German version adapted to elicit tense vowels. Type: dataset.

[19] Audacity team, 2020. Audacity (version 2.4.2). https://www.audacityteam.org/.

[20] Boersma, P. 2001. Praat, a system for doing phonetics by computer. *Glot International* 5, 341–345.

[21] Schiel, F. 1999. Automatic Phonetic Transcription of Non-Prompted Speech. *Proc. 14th ICPhS*, 607–610.

[22] Winkelmann, R., Harrington, J., Jänsch, K. 2017. EMU-SDMS: Advanced speech database management and analysis in R. *Computer Speech & Language*, 392–410.

[23] Winkelmann, R., Jaensch, K., Cassidy, S., Harrington, J. 2020. emur: Main package of the emu speech database management system.

[24] Winkelmann, R. 2015. Praattoformants2asspdataobj.r.

[25] Lobanov, B. M. 1971. Classification of Russian vowels spoken by different speakers. *Journal of the Acoustical Society of America* 49(2B), 606–608.

[26] Michalke, M. 2017. *sylly.de: Language Support for 'sylly' Package: German*. (Version 0.1-2).

[27] Al-Tamimi, J.-E., Ferragne, E. 2005. Does vowel space size depend on language vowel inventories? Evidence from two Arabic dialects and French. *9th European Conference on Speech, Communication and Technology*. Newcastle University.

[28] Maddieson, I. 2013. Vowel quality inventories. *The world atlas of language structures online*.

[29] Kohler, K. J. 1999. German. In: *Handbook of the International Phonetic Association*. Cambridge: Cambridge Univ. Press, 86–89.

[30] Scarborough, R., Dmitrieva, O., Hall-Lew, L., Zhao, Y., Brenier, J. 2007. An acoustic study of real and imagined foreigner-directed speech. *Journal of the Acoustical Society of America* 121(5), 3044–3044.

[31] Wagner, M. A., Broersma, M., McQueen, J. M., Dhaene, S., Lemhöfer, K. 2021. Phonetic convergence to non-native speech: Acoustic and perceptual evidence. *Journal of Phonetics* 88, 101076.