Speech Accommodation in Health-Care Interactions: Evidence Using a Mixed-Reality Platform

Rose Baker[1], Susan C. Bobb[2], Dai'Sha Dowson[1], Elisha Eanes[1], Makyah McNeill[1], Hannah Ragsdale[1], Audrey Eaves[1], Joseph G. Lee[1], Kathrin Rothermich[1]

[1]East Carolina University, [2]Gordon College
[1]E 5th Street, Greenville, NC 27858, [2]255 Grapevine Road, Wenham, MA 01984
{bakerr22, eanese23, mcneillm17, ragsdaleh18}@students.ecu.edu, susan.bobb@gordon.edu,
{dowsond23, eavesa16,leejose14, rothermichk17}@ecu.edu

Abstract

Many people in the US use more than one language at home, yet English remains the dominant (L1) language in US society, which can complicate medical encounters. In this study we ask in what ways effective communication can be ensured in health care settings when speakers differ in language proficiency. One strategy people use is second language (L2) speech accommodation, which is characterized by slowed speech, less complex words, and clearer enunciation. We employ a mixed-reality platform called MURSION to document how a group of Physician Assistant students use speech accommodation during a healthcare encounter. MURSION is a computer-based virtual environment where participants interact with an Avatar controlled by a human interactor in a standardized environment. We record 5-minute interactions between the student and a high or low English proficiency Avatar. Our analyses evaluate lexical choices in L1-L2 interactions with SCOPE (South Carolina Psycholinguistic Metabase) and acoustic properties with PRAAT. Results show that clinical students use slower speech and high frequency words when speaking to a low proficiency virtual patient, indicating a sensitivity for the communicative needs of L2 English users. Speech accommodation results will contribute to communication training modules for clinicians to interact efficiently with linguistically diverse populations.

Keywords: communicative needs, foreigner-directed speech, patient care, MURSION, PRAAT

1. Introduction

In the US, many people speak English as a second language, and minority language usage is on the rise. Even so, English remains the standard language in most places in US society, including healthcare environments. In these contexts, minority language users are especially vulnerable to miscommunications. While access to translators can mitigate misunderstandings, limited funding and resources restrict their availability.

One strategy that dominant English speakers (L1) apply to increase the chances of successful communication with second-language users (L2) is speech accommodation (Ferguson, 1975). L2 speech accommodation is characterized by slowed speech, use of easier words, and exaggerated enunciation. Self-reports suggest that L1 speakers often use speech accommodation with L2 speakers to ensure smooth communication. In a recent survey study, Rothermich et al. (2023) found that half of L1 respondents used accommodative speech to facilitate understanding. Though L1 speakers may intend to facilitate communication, L2 impressions suggest the effect of accommodation is not always positive. In the same study by Rothermich and colleagues (2023), almost half of L2 respondents said they do not appreciate the use of speech accommodation. More research is needed on the actual effectiveness of L2 speech accommodation. One of the few studies that has

directly investigated whether accommodation supports comprehension is a recent study by Piazza et al. (2023). Results suggested that language learners exposed to accommodative speech recognized new words faster and had better pronunciation than learners who received non-accommodated speech. Accommodated speech may be helpful for L2 speakers, but little is known about how speech accommodation helps L2 speakers successfully navigate the world outside the language learning classroom.

It is necessary to extend research on speech accommodation to the healthcare environment. In what ways do clinicians adapt their language when speaking to a person with limited proficiency? How do multifaceted features of L1 accommodation work together in conversation to sensitively meet the needs of the L2 speaker? We will investigate how L1 students with clinical training adapt their speech to be more accessible to L2 speakers at multiple level levels of language simultaneously.

2. Related Work

Previous research in English has shown that L1 speakers modulate their accommodation depending on perceived level of L2 English proficiency. Speech may be adjusted lexically or acoustically. Specifically, lexically accommodated speech may include lower average word frequency or fewer idiomatic expressions(Scarborough et al., 2007), and L1

speakers may repeat themselves more, clarify more often, and produce speech with a higher type/token ratio and lower mean length of utterance generally (Henzl, 1979). Relatively few studies have addressed lexical accommodation (see Rothermich et al., 2019 for a recent review), but acoustic L2 speech accommodation is well-documented: L1 speakers will speak more slowly, pause for longer (Pardo, 2006), adjust their fundamental frequency (Babel & Bulatov, 2012), duration, and intensity (Rodriguez-Cuadrado et al., 2018) and converge to the amplitude envelope of their L1 interlocutor (Lewandowski & Jilka, 2019). Additionally, the lexical and acoustic aspects of accommodation interact; vowel duration changes according to the semantic density of a word as well as its contextual predictability (Scarborough, 2010).

Accommodation is not uniform across speakers and situations. L1 speakers tend to change their speech more dramatically with L2 speakers who are less proficient (Chaudron, 1983; Dahl, 1981). L1 speakers also self-report evaluating the L2 proficiency level of their interlocutor to gauge their need for accommodation (Margić, 2017). Kühnert and Antolík (2017) found that L1 speakers used slower speech rates with less proficient L2 speakers, while they used a faster speech rate with more proficient L2 speakers. The properties of speech accommodation may therefore vary based on perceived L2 proficiency.

L1 accommodation may not always be appropriate due to the fact that spoken L2 proficiency is an unreliable indicator for comprehension ability. Previous research confirms that L2 learners frequently comprehend the target language at a higher level than their speaking ability (Litcofsky et al., 2016). Due to a discrepancy between receptive and expressive language skills, L1 speakers may use speech accommodation unnecessarily. Consequently, accommodation could be negatively perceived as condescending by the L2 speaker if it is not needed (Zuengler, 1991).

3. Current Study

With the advent of new technologies, such as virtual reality tools, it is now possible to capture speech accommodation dynamics and control for variables such as speaker characteristics. In this experiment, we innovate a new way of studying real-time language use by using a mixed-reality platform called MURSION. In MURSION, the study participant to interact with an avatar, controlled by a human interactor. It has mostly been used in immersive classroom simulations (Hudson et al., 2019) and to study clinical counseling (Harper, 2019). One advantage of the virtual environment is that features of the avatar, such as age, gender, race, accent, or medical history, can be standardized across sessions. We use this technology to document L2 speech accommodation in a healthcare setting. To our knowledge, we are the first to apply it to investigate linguistic accommodations.

We invited Physician Assistant and Nurse Practitioner students (n=17) to take part in the study. We targeted clinical students for two main reasons: 1) to assess their interaction with the MURSION tool to pilot it as a method of clinical education, and 2) to test a population of emerging clinicians who might have limited clinical experience with L2 speakers. We recorded their brief interactions with the avatar patients using Zoom and analyzed their speech based on acoustic and lexical features. We expected L1 speakers to accommodate low proficiency L2 speakers by raising their mean pitch, increasing their pitch range and speaking more loudly and slowly.

Figure 1. Screenshots of the virtual avatar patients. Left = Avatar 1, high proficiency; Right = Avatar 2, low proficiency.

4. Methods

Data collection is ongoing to meet our anticipated sample size of 60. So far, we have analyzed data from 17 L1 clinical students (age range = 24 – 42; 3 male/ 14 female) who interacted with two avatars over two sessions in MURSION, (see **Figure 1**). Our MURSION paradigm simulates a patient-provider interaction using two curated scripts developed by a clinical associate professor. Students are instructed to perform an "intake" interview with their new patients, the avatars. These interviews include standardized questions to establish care, known health issues, and family medical history. The order of encounters is counterbalanced by L2 proficiency level across participants. All L1 student participants are paired with both avatars. Interactions take about 5-10 minutes per participant. The participants are not informed about the proficiency manipulation until after the study.

Both avatars had the same medical issue in each respective student encounter. Avatar 1 displays high L2 proficiency (subtle Hispanic accent, advanced vocabulary, little to no grammatical errors) and complained of heartburn. Avatar 2 displays low L2 proficiency (heavy Hispanic accent, limited vocabulary, frequent grammatical errors) and presents with a sprained ankle. The same interactor controlled Avatar 1 and 2. Both avatars present as females and appear to be in their 20s.

Recording transcriptions
To analyze acoustic and lexical parameters of interest, the recordings need to be broadly transcribed at the word level. Each recorded interaction between the L1 speaker and the two avatars is independently transcribed by 8 transcribers (RB, DD, HR, and trained research assistants).

Data analysis
Data were analyzed considering acoustic and lexical aspects of language. Statistical analyses will uncover significant acoustic and lexical accommodation patterns among L1 speakers that vary according to L2 proficiency.

Acoustic data from 17 subjects were extracted from recordings with custom scripts in Praat, a speech analytical software (Boersma, 2002). We extracted acoustic features that are known to reflect speech accommodation, such as pitch measures (mean pitch, pitch range), intensity, and speech rate. We hypothesize that L1 speakers will accommodate low proficiency L2 speakers by maintaining their mean pitch, increasing their pitch range, and speaking more loudly and slowly.

Lexical data from 15 subjects will be extracted from transcription files with the South Carolina Psycholinguistic metabase (SCOPE; Gao et al., 2023). These data will consist of lexical frequency, contextual diversity, familiarity, and semantic diversity. We expect that L1 speakers will use words that are more common, familiar, and semantically flexible with low proficiency L2 speakers.

5. Preliminary Results

Preliminary findings from the high and low proficiency encounters suggest emerging L2 speech accommodation patterns in the MURSION testing environment.

Figure 2. *Results for mean pitch over time. Red = low proficiency, black = high proficiency.*
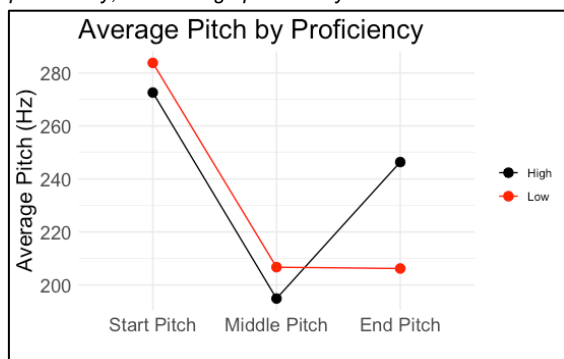
Acoustic results. A one-way ANOVA was conducted to examine the relationship between proficiency levels and four speech characteristics: speech rate, mean pitch, intensity, and word fluency (see Table 1). The results indicated a significant effect of proficiency levels on speech rate ($F(1, 28) = 4.69$, $p = 0.039$), but no significant effects on mean pitch ($F(1, 28) = 0.635$, $p = 0.432$), pitch range ($F(1, 28) = 0.721$, $p = 0.403$), or intensity ($F(1, 28) = 2.396$, $p = 0.133$). Consistent with our hypotheses, speech rate was lower when participants spoke with the low proficiency avatar.

Table 1. *Results for the acoustic analysis .*

	Avatar Proficiency	
	High	**Low**
Mean Pitch (Hz)	201.17	212.19
Pitch Range (Hz)	575.12	579.82
Mean Intensity (dB)	60.57	56.88
Speech rate (syll/sec)	4.58	4.11

An analysis of pitch over time revealed an interesting pattern (**Figure 2**). While pitch patterns are similar at the beginning and the midpoint of the interaction, participants seem to speak with a higher pitch towards the end when speaking to a high proficiency avatar. An analysis of the end pitch values revealed no significant differences.

Lexical results. A one-way ANOVA was conducted to examine the relationship between proficiency levels and four lexical characteristics: word frequency, contextual diversity, familiarity, and semantic diversity (see Table 2). The results indicated a significant effect of proficiency levels on word frequency ($F(1,19) = 60.07$, $p < .001$), contextual diversity ($F(1,19) = 30.21$, $p < .001$), and semantic diversity ($F(1,19) = 12.45$, $p < .001$) but no significant effects on familiarity ($F(1, 19) = 0.669$, $p = 0.424$). Consistent with our hypotheses, the participants used significantly more common, contextually and semantically flexible words with the low proficiency avatar.

Table 2. *Results for the lexical analysis.*

	Avatar Proficiency	
	High	**Low**
Frequency (Lg10)	4.23	4.46
Contextual Diversity (Lg10)	3.48	3.6
Familiarity (z value)	2.31	2.29
Semantic Diversity	2.00	2.04

6. Conclusions and Ongoing Work

Our hypotheses were partially confirmed, since the data collected here demonstrate that health care providers in training adjust their speech when interacting with the avatar who displays low levels of English proficiency. Specifically, we found providers using slower speech rates in these health care encounters with the low

proficiency avatar than with the high, ostensibly to improve comprehension in this medical context. No other acoustic measures (i.e., pitch and intensity) suggested differences. Similar to our results, previous research has also shown reduced speech rates in accommodative speech (Scarborough et al., 2007). Although not significant, the data patterns for mean pitch over time suggest that speakers adjust in real-time to their speech partner. High pitch is typically associated with friendliness and conversational engagement (Chen et al., 2001), so it is interesting that we note an overall decrease in pitch toward the lower proficiency avatar across the interaction compared to the higher proficiency avatar. It is too early to tell how this change interacts with other measures of accommodation, but one possibility is that interactions across proficiency levels are effortful, taxing working memory by increasing cognitive load as the more proficient speaker monitors comprehension and speech of their speech partner while adjusting what they are saying to be comprehensible. The observed decrease in pitch could indicate disengagement with the conversation due to the required effort of the interaction. As data collection continues, we expect that our acoustic findings will grow.

Results from our lexical measures further support our hypotheses. We found that student word choice became more frequent, contextually diverse, and semantically diverse with a low proficiency English speaker. These adjustments reflect sensitivity to the vocabulary range of the patient. Previous inquiry into lexical accommodation among language teachers measured word frequency, neighborhood density, and simplification supports our findings that L1 speakers use structurally more simple language with their L2 interlocutors. Lexical choices can also be intentional with the aim of supporting effective communication (Quené, 2008). Lexical adjustments are also highly related to discourse management in L1-L2 interactions, which may contribute to changes of the structure and content of L1 language (Henzl, 1979). The intake task assigned to our participants required them to ask questions of their patient and provide clinical advice, so they tended to use clarifying questions and check for understanding often. These discourse tasks may be similar to the social role of a language teacher who tends to use longer wait times for responses and more frequent conversational repair (Weizheng, 2019). Our study is one of the first to extend lexical accommodation results to the healthcare environment, showing that clinicians use accommodation strategies to communicate effectively and manage discourse with low-proficiency L2 speakers of English.

Our results are consistent with the previous literature, including studies with more naturalistic settings, capturing some of the modulations of speech patterns that occur when more proficient speakers of English direct language toward less proficient speakers. They suggest that virtual reality environments such as MURSION could create potentially effective training environments to document emerging cross-linguistic and cross-cultural interactional behaviors for healthcare providers who are still in training. A full sample size is needed to confirm the patterns of our initial analyses. Future work will assess whether educational interventions for clinicians can shift conversational patterns toward accommodation styles that are more effective in supporting conversations across different levels of language proficiency.

7. Acknowledgments

8. Bibliographical References

Babel, M., & Bulatov, D. (2012). The Role of Fundamental Frequency in Phonetic Accommodation. Language and Speech, 55(2), 231–248. https://doi.org/10.1177/0023830911417695

Boersma, P. (2002). Praat, a system for doing phonetics by computer. *Glot International 5.9/10*, 341–345. https://doi.org/10.1097/AUD.0b013e31821473f7

Chaudron, C. (1983). Simplification of Input: Topic Reinstatements and Their Effects on L2 Learners' Recognition and Recall. *TESOL Quarterly*, *17*(3), 437. https://doi.org/10.2307/3586257

Chen, A., Rietveld, T., & Gussenhoven, C. (2001). *Language-specific effects of pitch range on the perception of universal intonational meaning*. 1403–1406. https://doi.org/10.21437/Eurospeech.2001-362

Dahl, D. A. (1981). The role of experience in speech modifications for second language learners. *Minnesota Papers in Linguistics and Philosophy of Language*, *7*(2), 78–93.

Ferguson, C. A. (1975). Toward a Characterization of English Foreigner Talk. *Anthropological Linguistics*, *17*(1), 1–14.

Gao, C., Shinkareva, S. V., & Desai, R. H. (2023). SCOPE: The South Carolina psycholinguistic metabase. *Behavior Research Methods*, *55*(6), 2853–2884. https://doi.org/10.3758/s13428-022-01934-0

Harper, L. R. (2019). *Simulation of Patient Caregiver Counseling in Speech Language Pathology*.

Henzl, V. M. (1979). Foreigner Talk in the Classroom. *IRAL : International Review of Applied Linguistics in Language Teaching*, *17*(2), 159–167.

Hudson, M. E., Voytecki, K. S., Owens, T. L., & Zhang, G. (2019). Preservice teacher experiences implementing classroom management practices through mixed-reality simulations. *Rural Special Education Quarterly*, *38*(2), 79–94.

Kühnert, B., & Antolík, T. K. (2017). Patterns of articulation rate variation in English/French tandem interactions. In J. Volin & R. Skarnitzl (Eds.), *Pronunciation of English by Speakers of Other Languages*. Cambridge Scholars Publishing.

Lewandowski, N., & Jilka, M. (2019). Phonetic Convergence, Language Talent, Personality and Attention. *Frontiers in Communication*, *4*. https://www.frontiersin.org/articles/10.3389/fcomm.2019.00018

Litcofsky, K. A., Tanner, D., & van Hell, J. G. (2016). Effects of language experience, use, and cognitive functioning on bilingual word production and comprehension. *International Journal of Bilingualism*, *20*(6), 666–683. https://doi.org/10.1177/1367006915579737

Margić, B. (2017). Communication courtesy or condescension? Linguistic accommodation of native to non-native speakers of English. *Journal of English as a Lingua Franca*, *6*(1). https://doi.org/10.1515/jelf-2017-0006

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America*, *119*(4), 2382–2393. https://doi.org/10.1121/1.2178720

Piazza, G., Kalashnikova, M., & Martin, C. (2023). Phonetic accommodation in non-native directed speech supports L2 word learning and pronunciation. *Scientific Reports*, *13*. https://doi.org/10.1038/s41598-023-48648-7

Quené, H. (2008). Multilevel modeling of between-speaker and within-speaker variation in spontaneous speech tempo. *The Journal of the Acoustical Society of America*, *123*(2), 1104–1113. https://doi.org/10.1121/1.2821762

Rodriguez-Cuadrado, S., Baus, C., & Costa, A. (2018). Foreigner talk through word reduction in native/non-native spoken interactions. *Bilingualism: Language and Cognition*, *21*(2), 419–426. https://doi.org/10.1017/S1366728917000402

Rothermich, K., Baker, R., Falkins, S. M., Kum, J., Capps, M. R., & Bobb, S. C. (2023). Reported Use of Second-Language Speech Accommodation in Everyday Interactions: The Role of Individual Differences. *Journal of Speech, Language, and Hearing Research*, *66*(9), 3307–3327. https://doi.org/10.1044/2023_JSLHR-23-00167

Rothermich, K., Harris, H. L., Sewell, K., & Bobb, S. C. (2019). Listener impressions of foreigner-directed speech: A systematic review. *Speech Communication*, *112*, 22–29. https://doi.org/10.1016/j.specom.2019.07.002

Scarborough, R. (2010). Lexical and contextual predictability: Confluent effects on the production of vowels. In *Lexical and contextual predictability: Confluent effects on the production of vowels* (pp. 557–586). De Gruyter Mouton. https://doi.org/10.1515/9783110224917.5.557

Scarborough, R., Dmitrieva, O., Hall-Lew, L., Zhao, Y., & Brenier, J. (2007). An acoustic study of real and imagined foreigner-directed speech. *The Journal of the Acoustical Society of America*, *121*(5), 3044–3044. https://doi.org/10.1121/1.4781735

Weizheng, Z. (2019). Teacher-Student Interaction in EFL Classroom in China: Communication Accommodation Theory Perspective. *English Language Teaching*, *12*(12), 99–111.

Zuengler, J. (1991). Accommodation in native-nonnative interactions: Going beyond the "what" to the "why" in second-language research. In H. Giles, J. Coupland, & N. Coupland (Eds.), *Contexts of accommodation: Developments in applied sociolinguistics* (pp. 223–244). Cambridge University Press. https://doi.org/10.1017/CBO9780511663673

9. Language Resource References

Boersma, P. (2002). Praat, a system for doing phonetics by computer. Glot International 5.9/10, 341–345. https://doi.org/10.1097/AUD.0b013e31821473f7

Gao, C., Shinkareva, S. V., & Desai, R. H. (2023). SCOPE: The South Carolina psycholinguistic metabase. *Behavior Research Methods*, *55*(6), 2853–2884. https://doi.org/10.3758/s13428-022-01934-0