L2 Speakers' Use and Maintenance of Clear Speech in Natural Conversations

Yilin Hao

MA Program in the Social Sciences, University of Chicago

Author Note

Author roles were classified using the Contributor Role Taxonomy (CRediT; https://credit.niso.org/) as follows: Yilin Hao: conceptualization, writing, formal analysis

Correspondence concerning this article should be addressed to Yilin Hao, MA Program in the Social Sciences, University of Chicago, 1155 E 60th St., Chicago, IL 60637, USA, Email: y2hao@uchicago.edu

Abstract

This study investigates L2 speakers' use and maintenance of clear speech in natural conversations.

Keywords: L2 speakers, speech production, R programming, ggplot2, data communication

L2 Speakers' Use and Maintenance of Clear Speech in Natural Conversations

Introduction

Literature Review

Native English speakers often tend to use clear speech to make their speech more intelligible in their conversations with people who may have difficulty understanding their speech, such as non-native speakers. This raises the question: do second language (L2) speakers of English also use clear speech when communicating with other L2 speakers? This proposed study may provide insight into L2 speakers' use of clear speech in natural conversations, especially whether they will spontaneously adopt this speech style when they are talking to other L2 speakers and their maintenance of this speech style throughout the conversation. Literature review "Clear speech" refers to speakers' modification of their speech in order to increase the intelligibility for the listener, which includes various acoustic-phonetic changes like slower speech and linguistic adaptations like higher frequency words (Tuomainen et al., 2022). It is usually used by speakers to accommodate noisy environments, listeners with hearing impairments, and non-native listeners (Mattys et al., 2012). In Hazan et al.'s (2018) study of speech adaptation in conversations with adults with hearing loss, they found some significant differences in acoustic features between clear speech and normal speech style, which is also referred to as plain speech, including acoustic cues like articulation rate, mid-frequency range, and vocal effort which reflected by mean energy within the mid-frequency region. The benefit of clear speech has been found to be significant when addressing listeners with impaired hearing (Picheny et al., 1985) and in noisy environments (Calandruccio et al., 2020). By modifying the acoustic cues, clear speech can greatly improve native speakers' intelligibility gain (Bradlow & Bent, 2002). Although non-native speakers face different challenges in understanding speech compared to native speakers with hearing impairments or in noisy environments—difficulty accessing the underlying linguistic code rather than the speech signal—researchers found that non-native speakers can still benefit from clear speech with features such as a slower speaking rate and a wider dynamic pitch range (Bradlow & Bent, 2002). Recent studies have shown that clear speech has an overall effect in increasing

intelligibility gain regardless of language background, and the benefit difference between l2 speakers and native speakers is not significant (Jung & Dmitrieva, 2023). In addition, L2 speakers with higher fluency in English demonstrate a clear speech benefit that is more comparable to that of native speakers. (Smiljanić & Bradlow, 2011). Previous research has shown that besides perceptually benefiting from clear speech, fluent L2 speakers also exhibit the ability to produce effective acoustic-phonetic modifications in their speech and achieve the effect of clear speech (Kato & Baese-Berk, 2022). Although the extent of acoustic modification in their speech differs from that of native speakers, L2 speakers are still able to successfully use clear speech (Jung & Dmitrieva, 2023). By analyzing L2 speakers' acoustic characteristics while reading given texts in clear and plain speech styles, Jung and Dmitrieva (2023) found that L2 speakers also demonstrated the same clear speech features as native speakers. In other words, L2 speakers exhibited their ability to employ clear speech when they were prompted to do so.

Present Study

While the prior studies examined the effectiveness of L2 speakers' clear speech and the perceptual benefit of clear speech for L2 speakers using artificial methods where there's no actual conversation, there are few studies that examined L2 speakers' use of clear speech in natural conversations (Bradlow & Bent, 2002; Smiljanić & Bradlow, 2011). Previous studies have illustrated that native speakers' clear speech exhibits greater hyperarticulation under artificial methods, where participants are instructed to read texts in a clear speech style, compared to their clear speech produced in natural conversations (Hazan & Baker, 2011). Therefore, in natural conversations, L2 speakers' use of clear speech and its effectiveness may also differ from their performance in artificial contexts. The present study aims to explore whether L2 speakers will spontaneously use clear speech while talking to someone who may have difficulty understanding them— in this study, other non-native speakers— and whether they maintain this use throughout the conversation. Specifically, the present study aims to answer the following questions: whether L2 speakers will show clear speech characteristics like lower articulation rate, higher F0, and higher mean energy of mid-frequency region (Hazan et al., 2018) in their conversation with other

L2 speakers in comparison with their conversation with native speakers, and whether these characteristics persist from the early to the later stages of the conversation. We hypothesize that L2 speakers will use clear speech when interacting with another L2 speaker, indicating their purpose of making their speech more intelligible to non-native listeners (Mattys et al., 2012). If the hypothesis is true, then native speakers' speech is expected to show modifications including slower speech rate, greater F0, and higher mean energy in the mid-frequency region during their conversations with L2 speakers compared to their conversations with native speakers (Hazan et al., 2018). Furthermore, if L2 speakers tend to maintain using clear speech throughout the conversation to meet the intelligibility needs of listeners, they will continue showing these features throughout the speech. Lee and Baese-Berk(2020) found that native speakers' clear speech became less intelligible in the later portions of a conversation but reset to clear speech at the beginning of the next conversation, which suggests that even though the use of clear speech is oriented by listeners' needs, native speakers do not monitor their listener throughout the conversation. Therefore, it is possible that L2 speakers may use clear speech at an earlier stage of their conversation and become less clear later, which can suggest that the maintenance of clear speech is speaker-driven even though the use of clear speech is mostly listener-driven (Lee & Baese-Berk, 2020). An alternative hypothesis is that they will use plain speech style when they communicate with L2 speakers, thus their acoustic cues and speech signals would not have significant differences when talking to both L2 and native speakers, which can reflect their assumption that their listeners have no problem understanding them.

Methods

Participants Participants were students recruited from Northwestern University, including native speakers of English and L2 speakers. L2 speakers were included if they were non-native speakers who could also speak English and demonstrate proficiency in English. These participants had been in the U.S. for no more than 3 years and had achieved TOEFL scores of at least 600 for the paper-based test, 250 for the computer-based test, or 100 for the internet-based test (Van Engen et al., 2010). Materials This study will analyze conversations recorded during Diapix tasks, drawn

from the Wildcat Corpus (Van Engen et al., 2010). The Diapix task requires two participants to identify differences in a pair of pictures. Each pair of pictures presents the same general scene but with 10 differences: each picture has three items that are missing from the other, and there are four slight differences between the two pictures ("change" items, such as differences in color or other details) (Van Engen et al., 2010). Each participant will be able to see one of a pair of pictures and they are asked to identify all the differences by verbal communication with their partner.

Procedure Participants completed the Diapix task in two groups. One group involves pairs of two L2 speakers, and the second group involves pairs of one L2 speaker with a native English speaker. Participants sat back-to-back in a recording room where they could communicate but could not see each other's pictures. They were instructed to work collaboratively through conversation to identify all the differences in their pictures within 20 minutes and mark them by drawing circles or notes, etc. They wore headsets so their speeches will be recorded separately. Data Analysis Recording of participants' speech will be transcribed using the phonetic analysis software Praat. The articulation rate will be calculated by the number of syllables produced in speech divided by the total duration of the speech. The mean energy and fundamental frequency will be measured by Praat. The independent variables will be different kinds of participant pairing, including one native speaker and one L2 speaker or two L2 speakers, and different stages of conversation (early or late). We will compare the F0, mean energy, and articulation rate of L2 speakers' speech between these two groups, either talking to L2 speakers or native speakers. In addition, we will examine how these features vary between the early and late stages of the conversation. The greater F0, mean energy, and slower articulation rate are going to be the indicators of clear speech. The results will be analyzed using R with separate linear mixed-effects models.

Results

The linear mixed-effects regression model reveals that L2 speakers' F0 range is larger when their conversation partner is an L2 speaker with a different L1 with them (M = 31.42125) compared to when the conversation partner is native speaker (M = 21.77625) ($\beta = 9.64$, SE = 4.05, $p = 0.0235$). In addition, there is a tendency for L2 speakers' F0 to be greater when they

Table 1

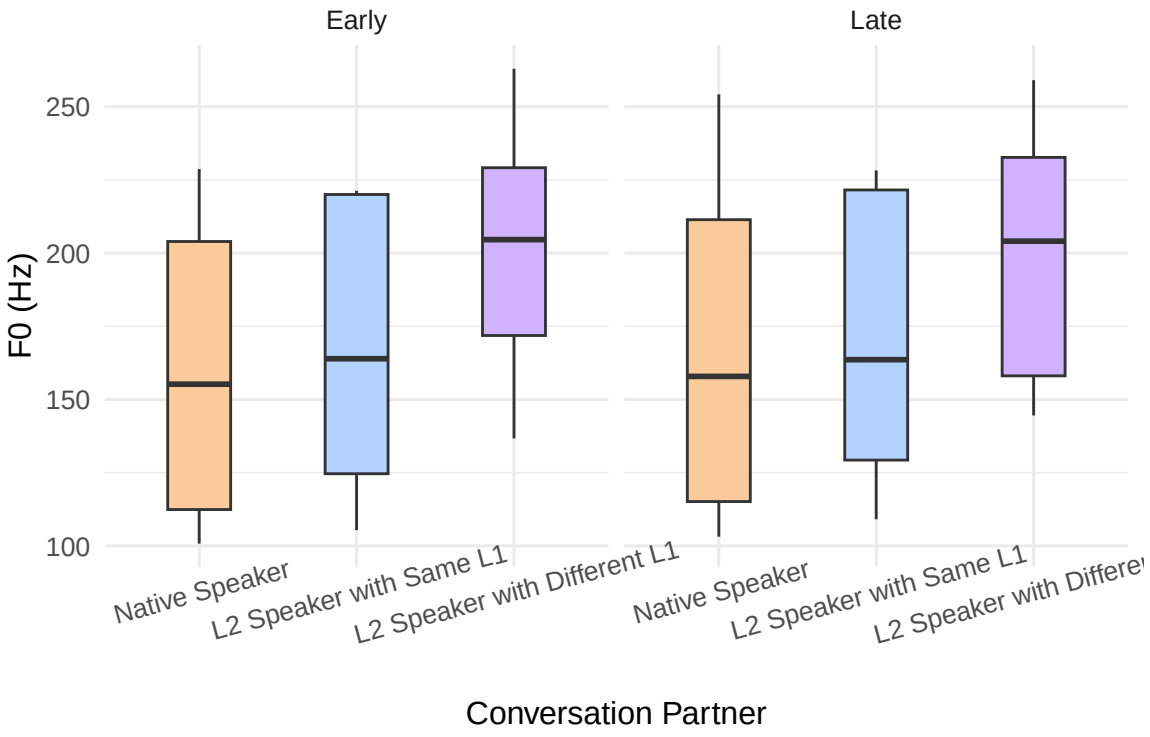
F0, F0 Range, Speech Rate, and Vowel Duration by Conversation Partner

Conversation Partner	F0	F0 range	Speech Rate	Vowel Duration
L2 Speaker with Different L1	200.91	31.42	3.17	0.14
Native Speaker	158.64	21.78	2.89	0.14
L2 Speaker with Same L1	167.86	23.74	3.17	0.14

Note. F0 (Hz), F0 Range (Hz), Speech Rate (syllables per second), Vowel Duration (s)

Figure 1

F0 by Conversation Partner

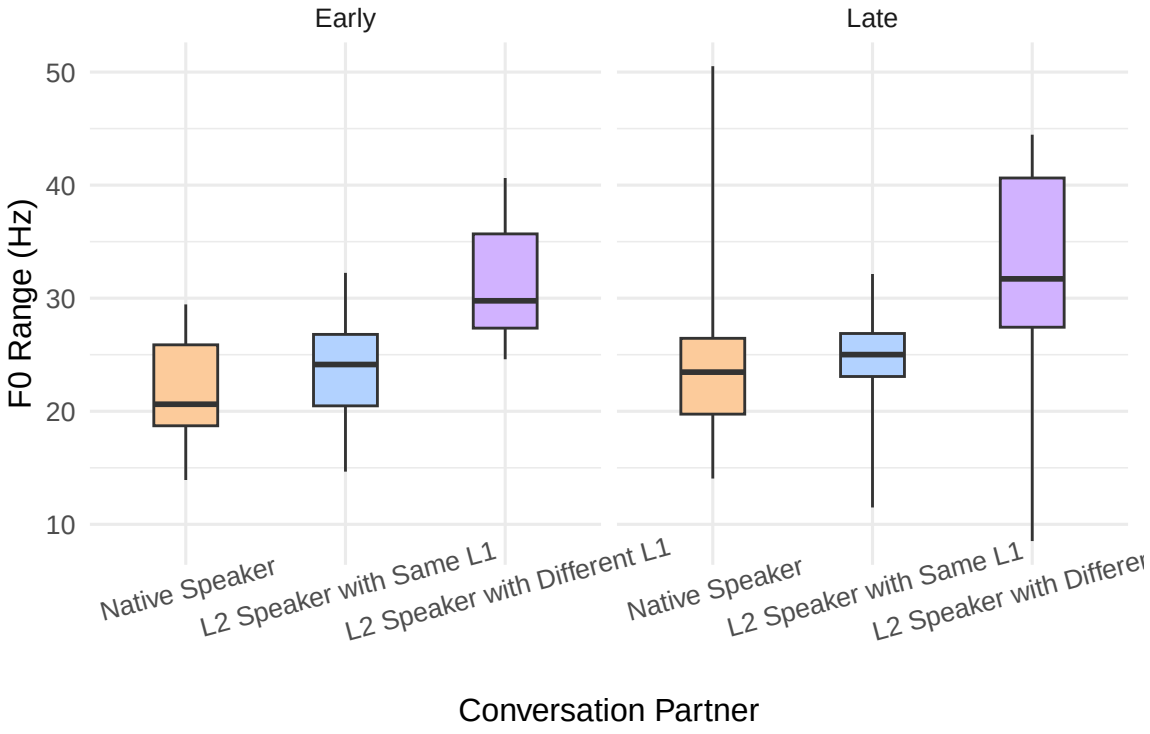


Note. F0 (Hz) for each group.

speak with an L2 speaker with a different L1 (M = 200.905) than when speaking with a native speaker (M = 158.64375), but the difference is not significant (β = 42.26, SE = 25.15, p = 0.1073).

Figure 2

F0 Range by Conversation Partner

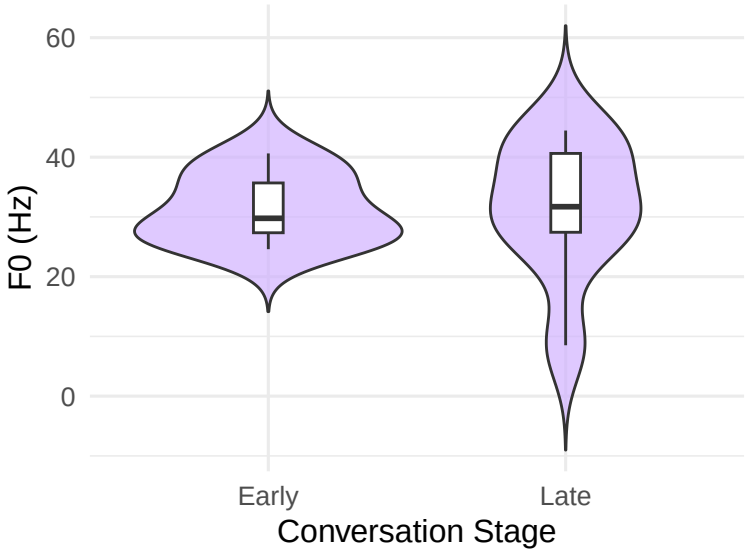


Note. F0 Range (Hz) for each group.

There is no significant difference in speech rate between the L2 listener (M = 3.17125) and native listener conditions (M = 2.89375) (β = 0.28, SE = 0.2, p = 0.1747). Similarly, vowel duration does not differ significantly between the L2 listener (M = 0.1375) and native listener conditions (M = 0.14375) (β = -0.01, SE = 0.01, p = 0.4971). Further, there is no significant difference in F0 or F0 range between early and late portion of the conversation as shown in Figure 1, Figure 2, and Figure 3.

Figure 3

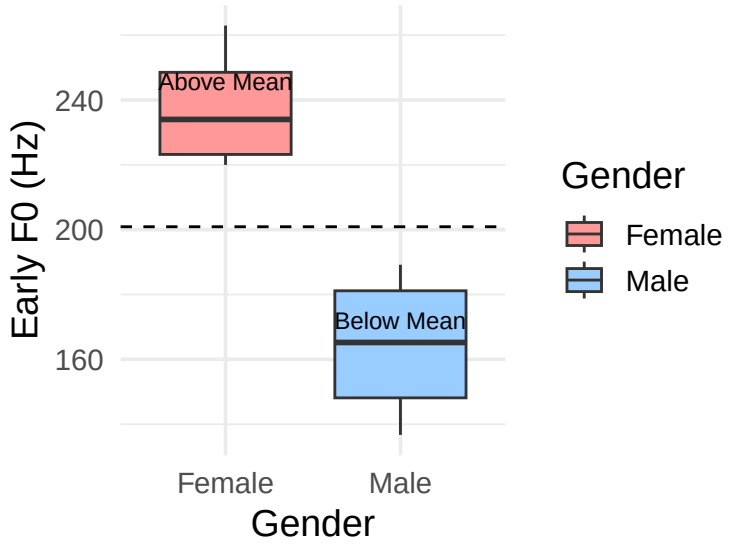
F0 Range by Conversation Stage



Note. F0 Range (Hz).

Figure 4

F0 Comparison by Gender in L2 Speakers



Note. F0 (Hz).

Discussion

This study examines L2 speakers' use of **clear speech**[1] in natural conversations. Specifically, we investigate whether L2 speakers employ clear speech when communicating with another L2 speaker who has a different L1. The results indicate that their F0 range is significantly larger when speaking to another L2 speaker compared to a native speaker. Although the mean difference in F0 between these two conditions is large, it does not reach statistical significance, possibly due to the high variance in F0 between female and male speakers in our dataset[2]. As shown in Figure 4, female and male speakers inherently exhibit different F0 values due to physiological differences. However, the significant increase in F0 range suggests that L2 speakers actively hyperarticulate their speech.

There is no significant difference in speech rate and vowel duration, which may be because L2 speakers generally speak at a slower rate[3]. Since they do not expect their listeners to have trouble understanding them, they may not adjust their speech in these ways. Additionally, the lack of significant differences between the early and late portions of speech suggests that L2 speakers consistently monitor their speech throughout the conversation.

---

[1] **Clear speech** refers to speech modifications that enhance intelligibility, including:

- Greater F0
- Greater F0 range
- Slower articulation rate
- Longer vowel duration

[2] Female speakers generally have a higher F0 than male speakers.

[3] The participants' speech rate is approximately 3 syllables per second, which is considerably slower than the typical speech rate of native speakers.

References

Bradlow, A., & Bent, T. (2002). The clear speech effect for non-native listeners. *The Journal of the Acoustical Society of America*, *112*(1), 272–284. https://doi.org/10.1121/1.1487837

Calandruccio, L., Porter, H., Leibold, L., & Buss, E. (2020). The clear-speech benefit for school-age children: Speech-in-noise and speech-in-speech recognition. *Journal of Speech, Language, and Hearing Research*, *63*(12), 4265–4276. https://doi.org/10.1044/2020_JSLHR-20-00353

Hazan, V., & Baker, R. (2011). Acoustic-phonetic characteristics of speech produced with communicative intent to counter adverse listening conditions. *The Journal of the Acoustical Society of America*, *130*(4), 2139–2152. https://doi.org/10.1121/1.3623753

Hazan, V., Tuomainen, O., Kim, J., Davis, C., Sheffield, B., & Brungart, D. (2018). Clear speech adaptations in spontaneous speech produced by young and older adults. *The Journal of the Acoustical Society of America*, *144*(3), 1331–1346. https://doi.org/10.1121/1.5053218

Jung, Y.-J., & Dmitrieva, O. (2023). Acoustic properties of non-native clear speech: Korean speakers of English. *Speech Communication*, *154*, 102982. https://doi.org/10.1016/j.specom.2023.102982

Kato, M., & Baese-Berk, M. (2022). Perceptual consequences of native and non-native clear speech. *The Journal of the Acoustical Society of America*, *151*(2), 1246–1258. https://doi.org/10.1121/10.0009403

Lee, D.-Y., & Baese-Berk, M. (2020). The maintenance of clear speech in naturalistic conversations. *The Journal of the Acoustical Society of America*, *147*(5), 3702–3711. https://doi.org/10.1121/10.0001315

Mattys, S., Davis, M., Bradlow, A., & Scott, S. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, *27*(7-8), 953–978. https://doi.org/10.1080/01690965.2012.705006

Picheny, M., Durlach, N., & Braida, L. (1985). Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech. *Journal of Speech,*

Language, and Hearing Research, *28*(1), 96–103. https://doi.org/10.1044/jshr.2801.96

Smiljanić, R., & Bradlow, A. (2011). Bidirectional clear speech perception benefit for native and high-proficiency non-native talkers and listeners: Intelligibility and accentedness. *The Journal of the Acoustical Society of America*, *130*(6), 4020–4031. https://doi.org/10.1121/1.3652882

Tuomainen, O., Taschenberger, L., Rosen, S., & Hazan, V. (2022). Speech modifications in interactive speech: Effects of age, sex and noise type. *Philosophical Transactions of the Royal Society B*, *377*(1841), 20200398. https://doi.org/10.1098/rstb.2020.0398

Van Engen, K., Baese-Berk, M., Baker, R., Choi, A., Kim, M., & Bradlow, A. (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. https://doi.org/10.1177/0023830910372495