PHONETIC SELECTIVITY IN ACCOMMODATION: THE EFFECT OF CHRONOLOGICAL AGE

Ildikó Emese Szabó

New York University
ildi.szabo@nyu.edu

ABSTRACT

This paper examines speech rate and f_0 data from the Switchboard corpus [11] to investigate how a speaker's chronological age affects the extent of their accommodation to their interlocutor. In terms of speech rate, I demonstrate that older speakers slow down less for a slow interlocutor than younger speakers, but they also speed up more for a fast interlocutor. I argue that these effects are due to older speakers' initial speech rate being slower than younger speakers', and thus the accommodation pattern attested by old speakers is compensatory, rather than a sign of decreased willingness or ability to accommodate later in life. This explanation is corroborated by the fact that accommodation for f_0 is not affected by the chronological age of the speaker.

Keywords: phonetic accommodation, convergence, chronological age, f_0, speech rate

1. INTRODUCTION

Accommodation is the changing of one's speech in response to an interlocutor. The extent of accommodation varies individually, but it is also linked to various social factors, such as gender of participants [24, 25, 2], sympathy [22, 1, 31], social status [14], and even the difference between the speaker's and the interlocutor's accent [18]. The effect of chronological age on accommodation is yet to be explored.

We might expect less accommodation from older speakers than younger ones for two reasons. Since several phonetic properties of a speaker change with physiological aging (e.g. slower production and worse hearing), we might expect that some of the age-related changes decrease the ability of older speakers to accommodate via worsened perception or more limited production. Moreover, accommodation has been linked to wanting to bridge a social gap [10, 9]. As older speakers exit the "linguistic marketplace" [5] they use more non-standard forms associated with lower status social group [19], which could indicate that they are less invested in social appearances. This could lead to a decreased willingness to accommodate to others.

This paper examines whether younger and older speakers differ in their accommodation trajectories regarding two phonetic dimensions: speech rate and fundamental frequency. Speakers accommodate to their interlocutors along both variables. Accommodation has been demonstrated for speech rate [30]. Moreover, a corpus-study on Switchboard showed that this effect is in fact due to the speaker converging to the interlocutor's speech rather than both of them converging to a conversation-specific value—fast speech for animated conversations or slow for more solemn ones [6]. F_0 accommodation has also been documented [12, 13, 14, 24, 2]. While accommodation is well-attested along both variables, they differ in terms of age-grading. Speech rate shows a clear pattern of age-grading—younger speakers talk faster than older speakers [8, 32, 15, 16, 17, 6]. In a longitudinal study, Reubold et al. [28] find that their female and male speaker do not follow the same trajectory: their male speaker's f_0 falls up to the age of about 85, and then sharply rises. Their female speaker, who was observed until the age of 76, shows a consistent decrease in f_0. However, the oldest speaker in this study is 62, and until that age there is a fairly uniform decrease in f_0 with age across the sexes Reubold et al. examined. While a uniform decrease in f_0 should be observable in the age range of this paper, it is blurred by sex-effects.

By comparing two different phonetic variables, the present study can add to the discussion on the uniformity of accommodation—i.e. how and whether accommodation looks different depending on phonetic dimension. A comparison of the two variables also allows us to tease apart general age effects in accommodation from variable-specific ones. While a change in accommodation for only one of the variables would mean a change in the production or perception of the given variable, a change in accommodation along both variables might indicate that either the willingness or the cognitive ability to accommodate is affected by aging. A lack of age effects would indicate that older speakers possess the same willingness and ability to accommodate as younger speakers.

2. SPEECH RATE STUDY

2.1. Methods

Following Cohen Priva et al. (2017), I study speech rate accommodation in Switchboard [11], a corpus of two-sided telephone conversations between strangers on a given topic. The dataset is limited to conversations of at least 5 minutes, utterances of at least 3 words, and to speakers whose *interlocutors* (`IL`) spoke enough early on for a baseline (`IL baseline`) to be established—3 utterances of 3+ words in the first 40 seconds. The final dataset consists of 50,854 utterances from 206 unique participants (116 female, 90 male), aged 16–62 (mean: 37.8, median: 35). Speech rate (*sr*) was measured as the actual duration of the utterance divided by its duration as predicted based on corpus-specific average durations for each word it contains, following Cohen Priva et al. [6]. A higher value for *sr* indicates slower speech, a lower *sr* means faster speech.

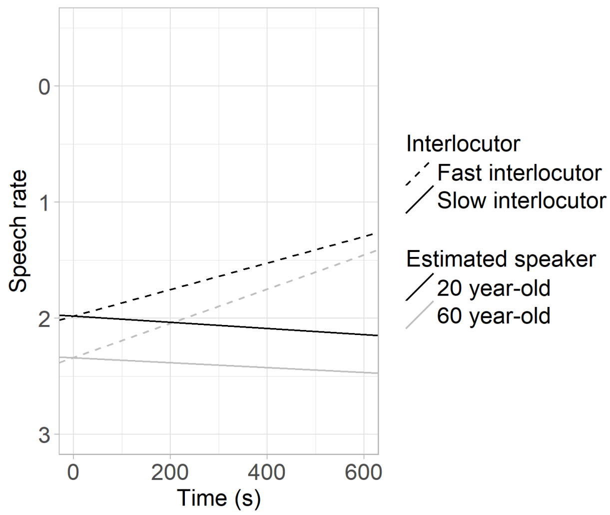
A linear mixed-effect model is fitted to the speaker's (SP) speech rate by utterance as predicted by `time` (elapsed in conversation), `IL baseline`, `age of SP`, and `age of IL`, with all interactions and a random by-speaker intercept. The interlocutor's baseline (`IL baseline`) for each conversation is obtained as the averaged *sr* of the interlocutor's first 3 utterances in the conversation. This baseline reflects the initial values of the interlocutor. The threshold of significance was adjusted by the number of tests run in the model with Bonferroni-correction (α=0.05/15=0.0033). Accommodation can be reflected in a positive main `IL baseline` effect or a positive `time`×`IL baseline` interaction—i.e. either the interlocutor's baseline and the speaker's *sr* correlate or with time the effect of the interlocutor's baseline on the speaker's *sr* increases, respectively. Age effects in accommodation would show up as an `IL baseline`×`SP age` or as a `time`×`IL baseline`×`SP age` interaction.

2.2. Results

Speakers' speech speeds up (we find lower values for *sr*) as more time elapses in the conversation (`time` β=$-3.29*10^{-3}$, p=0.0004), which is likely a task effect arising from the initial unnaturalness of having to strike up a conversation on a given topic. The older the speaker is, the less they speed up as the conversation goes on `time`×`age of SP` β=$8.38*10^{-5}$, p=0.0002). `time`×`IL baseline` is significant (β=$1.96*10^{-3}$, p=0.0008), which suggests that the longer the conversation goes on, the more the speaker's speech correlates with the interlocutor's. A three-way interaction between `time`, `IL baseline`, and `age of SP` (β=$-5.081*10^{-5}$, p=0.0003) reveals that speakers of different ages accommodate to their interlocutors differently: older speakers are less quick to accommodate to slow interlocutors than young speakers.

In order to understand whether this is part of a more complex pattern or indicative of older speakers accommodating less categorically, model predictions for a young and an old speaker were plotted with both slow and fast interlocutors (Figure 1). Estimates for speakers conversing with a fast interlocutor are in dashed, estimates with a slow interlocutors are in solid, age is reflected in the color of lines (black for the 20 year-old's estimates, gray for the 60 year-old's one). We can see that while old and young speakers have different speech rates at the beginning—younger speakers have a significantly faster speech rate than older ones to begin with—they reach similar target values after 10 minutes (600s). This in turn means that the rate with which they accommodate is different: older speakers speed up more to accommodate to a fast interlocutor, but they also slow down less to accommodate to a slow one. The difference in accommodating to slower interlocutors is less apparent, but still present, possibly due to the fact that it is outshadowed by the general effect of speeding up throughout the conversation (see the effect of `time` above).

Figure 1: Predicted for accommodation patterns for younger and older speakers (Black: age 20; Gray: age 60). Faster speech rates are vertically towards the top.



2.3. Interim conclusions

Both age-grading and speech rate accommodation have been replicated in the dataset here. Results further indicate that older and younger speakers follow different patterns in accommodation. Older speakers speed up more for fast interlocutors than younger speakers, but younger speakers also slow down more to accommodate to slow interlocutors than older people do. I argue that this is a compensatory pattern and can be explained by the difference in the starting speech rates of older vs. younger speakers. For example, in order to accommodate to a fast interlocutor, older speakers on average have a longer way to go than younger speakers who already talk faster on average. This is corroborated by the fact that old and young speakers reach comparable target values in the end.

However, the compensatory nature of this effect needs to be confirmed by looking at another phonetic dimension. If the age effect in speech rate accommodation is compensatory, then for a phonetic variable where age-grading is secondary (like f_0) we do not expect the age of the speaker to significantly affect the observed pattern of accommodation.

3. FUNDAMENTAL FREQUENCY STUDY

3.1. Methods

A study on f_0 accommodation was also conducted on the Switchboard corpus to test for compensatory effects. The same corpus was used, but because of the nature of measurements, a different set of items and speaker were excluded. Target vowels for f_0 measurements were restricted to high vowels only (/i ɪ u ʊ/), since vowel height has been shown to slightly but significantly influence f_0: high vowels have up to ∼25Hz higher f_0 values than low vowels [7, 27, 20, 23]. F_0 measurements were obtained from each monosyllabic word containing a high vowel at 45%, 50%, and 55%. Measurements above 300Hz were considered errors and were discarded. Further outliers were filtered out using the median-centered mean average deviation (MAD) [21]—outliers were identified as values beyond the median +/−3 * MAD, and were discarded. To exclude data points influenced by intonation patterns, all datapoints above 170 Hz for male participants and all datapoints above 280 Hz for female participants were also excluded (reference frequencies in [29, 4]). For every vowel, the remaining valid measurements (out of 45-50-55%) were averaged. Data were filtered for the baseline requirement (3 tokens from the interlocutor in the first 40 seconds), and speakers whose interlocutors did not speak enough early on, were discarded. Only same-sex dyads were considered in order to eliminate issues of physiologically infeasible targets. Female-female (*ff*) and male-male (*mm*) conversations were separated.

The dataset consists of 86,173 monoysllabic high monophthong tokens (*ff*: 54,164, *mm*: 32,029), from 713 conversation sides (*ff*: 442, *mm*: 271). There were 201 unique speakers (112 female, 89 male), whose age range was 16–62 (mean of 38.44, and a median of 35), which is comparable to the speech rate dataset. The 54,164 tokens were protductions of 533 different words, uttered 0.14 to 600.72 seconds into the conversation.

The female dyads and male dyads were analysed separately, with identical linear mixed-effect regression models fitted to the f_0 values. Independent variables were `time` (elapsed in conversation), `IL baseline`, `age of SP`, and `age of IL` in years, with all interactions as well as random by-speaker and by-word intercepts. The interlocutor's baseline (`IL baseline`) for each conversation is obtained as the averaged f_0 of the first 3 high vowels of the interlocutor. The threshold of significance was adjusted via Bonferroni-correction (α=0.0033). Again, accommodation can be reflected in a positive main `IL baseline` effect or a positive `time`×`IL baseline` interaction—i.e. either the interlocutor's baseline and the speaker's fundamental frequency correlate or with time the effect of the interlocutor's baseline on the speaker's f_0 increases. Age effects in accommodation would show up as an `IL baseline`×`SP age` or as a `time`×`IL baseline`×`SP age` interaction.

3.2. Results

The *ff* model shows no significant effects, there is a trend to converge (`IL baseline`: β=2.966*10−1, p=0.0289 > α=0.0033), which is not significant after Bonferroni-correction—likely an artifact of the number of tests run. The same is true for the interaction between `IL baseline` and `SP age`, which is not statistically significant either (p=0.0454 > α=0.0033). The interactions of these effects with time are not significant either (`time`×`IL baseline`: p=0.4855, `time`×`IL baseline`×`SP age`: p=0.2309). The fact that no f_0 accommodation was observed for *ff* dyads could be due to dataset-specific properties such as sparsity of the data rather than gender differences in accommodation, but it is also consistent with findings indicating that men accommodate more for f_0 [2].

In contrast, we do find evidence of accommodation in the *mm* model (`IL baseline`: β=5.190*10^{-1}, p=0.0001). The interaction between `IL baseline` and `SP age` does not reach significance after correcting for the number of tests run (p=0.0093 > α=0.0033). This indicates that the higher the f_0 of the interlocutor, the higher the f_0 of the speaker throughout the conversation, which suggests instantaneous accommodation—in contrast with the accommodation that we saw for speech rate that happened over the course of the conversation.

There also is a significant effect of the age of interlocutor (β=1.388, p=0.0018), which indicates that the older the interlocutor was, the higher the pitch of the speaker went. However, there is a weak positive correlation between an interlocutor's age and their baseline for interlocutors older than 35 years old (Kendall's tau-b τ=0.0312), which means that for the older speakers in this dataset the older they were the higher the baseline that they provided was. Therefore, the effect of speakers using higher pitch when talking to older interlocutors could indirectly indicate accommodation as well. There is an almost significant negative interaction between an interlocutor's age and their baseline under the adjusted p-threshold (β=–0.9776*10^{-3}, p=0.0038 >0.00333). This interaction points in the direction that the `IL baseline` effect and the `age of IL` effect are not independent of one another—i.e. that the latter is also an effect of accommodation.

4. SUMMARY AND CONCLUSIONS

This paper presented two corpus studies on accommodation—one on speech rate, the other on f_0 accommodation—and contributes to the literature on accommodation. Speech rate accommodation was found to vary with age: older speakers accommodated less to slower interlocutors, and more to faster ones, which indicates that the age effect is compensatory in nature. Each age group changed their speech more to accommodate to the group whose speech was further away from their own, which means they appproached similar target values even though initially older speakers talked slower than younger speakers. Since accommodation to faster interlocutors was no less for older speakers than for younger speakers, we can conclude that speakers retain both the ability and the willingness to accommodate (until at least the age of 60 years old). As for f_0 accommodation, no age effects were found, which further supports the retention of the ability and willingness to accommodate. However, research targetting fundamental frequency and other phonetic variables should be conducted to gather further evidence in support of this claim.

Phonetic selectivity of accommodation has been demonstrated—not all sounds are equal targets to accommodation [3] and accommodation along different phonetic dimensions does not covary [25, 26]. This paper, however, draws attention to another way in which accommodation is phonetically selective. The effect of social factors on accommodation is phonetically selective—only accommodation along certain phonetic variables is affected by a given social variable. Since age-grading of speech rate accommodation seems compensatory in nature, we can stipulate that social stratification of a given variable is a prerequisite of social sensitivity in its accommodation. That is, we only observe social groups accommodating differently along some phonetic variable if there was a difference between those social groups to begin with. Such a theory would correctly predict that, for instance, f_0 accommodation is sensitive to gender [24, 25, 2], which outlines interesting areas of further research.

The exact course and timing of accommodation not only depends on the social identities and physiological properties of the talkers but also on the phonetic variables involved. While accommodation can be characterised as a holistic process, it is far from uniform, and the present paper demonstrates that the effect of social factors can be conditional on the phonetic dimension.

ACKNOWLEDGEMENTS

I would like to thank my two anonymous reviewers at ICPhS, along with Lisa Davidson, Laurel MacKenzie and Gregory Guy for their feedback on this project and paper, as well as members of the NYU PEP Lab. All remaining errors are my own.

5. REFERENCES

[1] Babel, M. 2010. Dialect divergence and convergence in New Zealand English. *Language in Society* 39(4), 437–456.

[2] Babel, M., Bulatov, D. 2012. The role of fundamental frequency in phonetic accommodation. *Language and Speech* 55(2), 231–248.

[3] Babel, M. E. 2009. *Phonetic and social selectivity in speech accommodation*. PhD thesis University of California, Berkeley.

[4] Baken, R. J., Orlikoff, R. F. 2000. *Clinical measurement of speech and voice*. London: Taylor and Francis.

[5] Bourdieu, P. 1991/1982. *Language and Symbolic Power*. Harvard University Press. Original printed in 1982.

[6] Cohen Priva, U., Edelist, L., Gleason, E. 2017. Converging to the baseline: Corpus evidence for convergence in speech rate to interlocutor's baseline. *The Journal of the Acoustical Society of America* 141(5), 2989–2996.

[7] Crandall, I. B. 1925. The Sounds of Speech. *Bell*

Labs Technical Journal 4(4), 586–626.

[8] Duchin, S. W., Mysak, E. D. 1987. Disfluency and rate characteristics of young adult, middle-aged, and older males. *Journal of Communication Disorders* 20(3), 245–257.

[9] Giles, H., Coupland, N., Coupland, J. 1991. Accommodation theory: Communication, context, and consequence. *Contexts of accommodation: Developments in applied sociolinguistics* 1.

[10] Giles, H., Mulac, A., Bradac, J. J., Johnson, P. 1987. Speech accommodation theory: The first decade and beyond. *Annals of the International Communication Association* 10(1), 13–48.

[11] Godfrey, J. J., Holliman, E. C. 1992. SWITCHBOARD: Telephone speech corpus for research and development. *IEEE International Conference on Acoustics, Speech, and Signal Processing* volume 1 517–520.

[12] Goldinger, S. D. 1997. Words and voices: Perception and production in an episodic lexicon. *Talker variability in speech processing* 33–66.

[13] Gregory, S., Webster, S., Huang, G. 1993. Voice pitch and amplitude convergence as a metric of quality in dyadic interviews. *Language & Communication* 13(3), 195–217.

[14] Gregory, S. W., Webster, S. 1996. A nonverbal signal in voices of interview partners effectively predicts communication accommodation and social status perceptions. *Journal of Personality and Social Psychology* 70(6), 1231–1240.

[15] Harnsberger, J. D., Shrivastav, R., Brown, W., Rothman, H., Hollien, H. 2008. Speaking rate and fundamental frequency as speech cues to perceived age. *Journal of Voice* 22(1), 58–69.

[16] Horton, W. S., Spieler, D. H., Shriberg, E. 2010. A corpus analysis of patterns of age-related change in conversational speech. *Psychology and Aging* 25(3), 708.

[17] Kendall, T. 2013. *Speech Rate, Pause and Sociolinguistic Variation: Studies in Corpus Sociophonetics*. Palgrave Macmillan.

[18] Kim, M. 2012. *Phonetic accommodation after auditory exposure to native and nonnative speech*. PhD thesis Northwestern University.

[19] Labov, W. 2001. *Principles of Linguistic Change: Social Factors* volume 2. Oxford: Blackwell.

[20] Lehiste, I., Peterson, G. E. 1961. Some basic considerations in the analysis of intonation. *The Journal of the Acoustical Society of America* 33(4), 419–425.

[21] Leys, C., Ley, C., Klein, O., Bernard, P., Licata, L. 2013. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology* 49(4), 764–766.

[22] Natale, M. 1975. Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *Journal of Personality and Social Psychology* 32(5), 790.

[23] Ohala, J. J., Eukel, B. W. 1987. Explaining the Intrinsic Pitch of Vowels. *In honor of Ilse Lehiste* 207–215.

[24] Pardo, J. S. 2006. On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America* 119(4), 2382–2393.

[25] Pardo, J. S. 2010. Expressing oneself in conversational interaction. In: Morsella, Ezequiel, , (ed), *Expressing oneself/expressing ones̛ self: Communication, cognition, language, and identity*. 183–196.

[26] Pardo, J. S., Urmanche, A., Wilman, S., Wiener, J. 2017. Phonetic convergence across multiple measures and model talkers. *Attention, Perception, & Psychophysics* 79(2), 637–659.

[27] Peterson, G. E., Barney, H. L. 1952. Control methods used in a study of the vowels. *The Journal of the acoustical society of America* 24(2), 175–184.

[28] Reubold, U., Harrington, J., Kleber, F. 2010. Vocal aging effects on F0 and the first formant: A longitudinal analysis in adult speakers. *Speech Communication* 52(7-8), 638–651.

[29] Titze, I. 1994. *Principles of voice production*. Englewood Cliffs, NJ.

[30] Webb, J. T. 1970. Interview synchrony: An investigation of two speech rate measures in an automated standardized interview. In: *Studies in dyadic communication New York*. Pergamon, New York 55–70.

[31] Yu, A., Abrego-Collier, C., Baglini, R., Grano, T., Martinovic, M., Otte III, C., Thomas, J., Urban, J. 2011. Speaker attitude and sexual orientation affect phonetic imitation. *University of Pennsylvania Working Papers in Linguistics* 17(1), 26.

[32] Yuan, J., Liberman, M., Cieri, C. 2006. Towards an Integrated Understanding of Speaking Rate in Conversation. *Ninth International Conference on Spoken Language Processing*.