

Phonetic convergence and language talent within native-nonnative interactions

Natalie Lewandowski, Travis Wade and Grzegorz Dogil
Institute for Natural Language Processing, University of Stuttgart, Germany
<https://doi.org/10.36505/ExLing-2008/02/0035/000094>

Abstract

The notion of phonetic convergence covers all adaptations in articulatory and acoustic features towards those of a communicative partner, or in other terms an increase in segmental and suprasegmental similarity between them (Pardo 2006). Up until now most of the experiments on convergence were designed for monolingual dyads, with very few investigations of convergence in native-nonnative interactions in a foreign language learning environment. We tried to analyze the convergent behavior of nonnative speakers of English in dialog with native speakers and the persistence of the effect in relation to their rated phonetic talent. In this paper we present first results for a global measurement of convergence - the comparison of amplitude envelope signals.

Key words: phonetic convergence, talent, foreign language, native-nonnative interaction

Introduction

Accommodating strategies (such as convergence, divergence, maintenance and complementarity) may be used to achieve solidarity with or dissociation from a partner, in a dynamic setting with online feedback (Giles et al. 1991). As Wedel (submitted) points out, convergence and divergence can also be part of a simultaneous process, with some features underlying an automatically driven positive accommodation and some functioning as identity markers for group membership and therefore subject to divergence. In order to account for such apparently countervailing and still simultaneously observed strategies, they need to be embedded in a dynamical systems framework of language that accounts amongst others for more fine-grained usage-induced changes.

Thus in this paper we assume a usage-based account that allows for the storage and processing of fine phonetic detail and additional social information - an exemplar theoretical account (e.g. Johnson 1997, Pierrehumbert 2001). It has been suggested that normalization processes are not needed in an exemplar model of perception because “the model retains the variability encountered in speech [and] it is able to cope with the variability that it encounters in new tokens” (Johnson 1997:162). An exemplar-based account could also provide an explanation for social accommodation processes, since more recent and more frequently heard

exemplars could guide the typical productions within a speech community and therefore lead to the adaptation of the prominent speech patterns. Although exemplar-based production models also provide a straightforward explanation for the nature of the perception-production loop, with both processes relying on the same pool of exemplars (e.g. Pierrehumbert 2001), this does not imply that a phonetic target is necessarily realized as a perfect match (e.g. due to noise in the motor control and execution).

Since pronunciation seems to have a separate and special status within second language acquisition, we hypothesize that for pronunciation convergence towards a foreign language speaking partner, be it conscious or not, not only a certain proficiency in that language is crucial but also some degree of phonetic talent.

Method

All participants of the present study have also been tested for their phonetic language talent in another ongoing project (see Jilka et. al 2008 for details). Their pronunciation talent based on their performance in various perception and production tests has been rated on a scale from 1 to 6, 1.0 standing for exceptionally talented and 6.0 for absolutely not talented.

Experimental design

The participants were 30 native speakers of German that were involved in two dialogs with native speakers of General American (male) and Southern Standard British English (female). The elicitation technique used was a Diapix-task (Bradlow et al. 2007), a picture matching game in which participants had to identify ten differences between their pictures without seeing the partner's picture. The control task consisted of reading a list of words that contained words from the two picture-sets and unrelated filler words before and after each Diapix-dialog. The whole experimental session thus consisted of the following steps: first reading of the word list, Diapix-dialog A with the GA speaker, second reading of the word list, Diapix-dialog B with the SSBE speaker, third reading of the word list.

Analysis

Participants' and native speakers' word-level productions were transformed to amplitude envelope signals over 4 log-spaced frequency ranges, equalized for total amplitude, and compared by cross-correlation to estimate a match. Match values range from 0 to 1, with 1 indicating a perfect match and 0 no match. Since these signals do not explicitly include any specific phonetic cues such as formant frequencies but are a (relatively) compact and transparent representation of information that is present in various forms

throughout the auditory system, the match value provides a rough global measurement of spectrotemporal similarity (Wade et al. submitted).

Results

We first compared mean match values from a test person and a native speaker early and late in the dialog. Figure 1 shows the match values across the dialog for two female participants (BS and BR) with high scores for phonetic talent (respectively 1,4 and 1,5) and their partners.

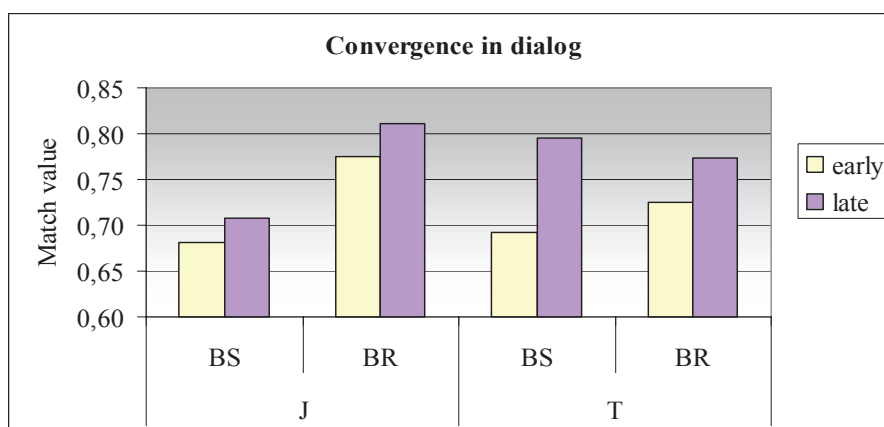


Figure 1. Convergence in dialog. Comparison of raw match values (mean across all target words) between an early and a late point in the conversations for two participants: BR and BS and two native speakers: J and T.

The match values are rising for all displayed conditions, indicating convergence in pronunciation between the participant and the respective English native dialog partner (J or T).

Table 1. Results of a bivariate correlation analysis between the word list task and the talent score.

		T2-T1	J3-J1
Talent	Pearson's correlation	0,893	0,880
	significance	0,107	0,120
	covariance	0,053	0,082

The second analysis (Table 1) aims at determining how persistent the convergent effect observed in the dialogs is. Compared were the mean match values between the word list readings before and after the dialogs, for four

subjects with varying talent scores (ranging from 4,2 up to 1,4) and the native speakers. The positive correlation results suggest that more talented subjects could indeed converge more than less talented speakers, although the effect did not reach statistical significance due to the small number of participants considered.

Discussion

Although the described results are based on a limited amount of data and therefore of course have to be treated with caution, they nevertheless point to some promising tendencies that need to be further explored on the whole data set. Some important factors that were disregarded now (such as the direction of convergence in the dialog, possible gender differences or frequency effects within the target words) will then also be taken into account to provide a full picture of the relation between talent and convergence in native-nonnative interactions.

References

- Bradlow, A. R., Baker, R. E., Choi, A., Kim, M. and van Engen, K. J. 2007. The Wildcat Corpus of Native- and Foreign-Accented English. *Journal of the Acoustical Society of America*, 121(5), Pt.2, 3072.
- Giles, H., Coupland, N. and Coupland, J. 1991. Accommodation Theory: Communication, context, and consequence. In H. Giles, J. Coupland and N. Coupland (Eds.), *Contexts of Accommodation: Developments in Applied Sociolinguistics*, 1-68. Cambridge, Cambridge University Press.
- Jilka, M., Anufryk, V., Baumotte, H., Lewandowski, N., Rota, G. and Reiterer, S. 2008. Assessing Individual Talent in Second Language Production and Perception. In: Rauber, A. S. et al. (Eds.). *New Sounds 2007: Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech*. Florianópolis, Federal University of Santa Catarina, 224-239.
- Johnson, K. 1997. Speech perception without speaker normalization: An Exemplar Model. In K. Johnson and J. W. Mullennix (Eds.), *Talker variability in speech processing*, 145–165. San Diego: Academic Press.
- Pardo, J. S. 2006. On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America*, 119, 2382-2393.
- Pierrehumbert, J. B. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee and P. Hopper (Eds.), *Frequency and the emergence of linguistic structure*, 137–157. Amsterdam, Benjamins.
- Wade, T., Dogil, G., Schütze, H., Walsh, M. and Möbius, B. (submitted). Syllable frequency effects in a context-sensitive segment production model.
- Wedel, A. and Van Volkinburg, H. (submitted): Modeling simultaneous convergence and divergence of linguistic features between differently-identifying groups in contact.