System 66 (2017) 142–155

Contents lists available at ScienceDirect

System

journal homepage: www.elsevier.com/locate/system

Ab initio perceptual learning of foreign language sounds: Spanish consonant acquisition by Chinese learners

Jian Gong [a, b, *], María Luisa García Lecumberri [b], Martin Cooke [b, c]

[a] School of Foreign Languages, Jiangsu University of Science and Technology, 2 Mengxi Road, 212003 Zhenjiang, China
[b] Language and Speech Laboratory, Universidad del País Vasco, Paseo de la Universidad 5, 01006 Vitoria, Spain
[c] Ikerbasque (Basque Science Foundation), Spain

ARTICLE INFO

Article history:
Received 12 January 2016
Received in revised form 17 March 2017
Accepted 19 March 2017

Keywords:
Phonetic training
Consonant perception
Spanish
Chinese
Second language sound acquisition

ABSTRACT

High-variability phonetic training is effective in the acquisition of foreign language sounds. Previous studies have largely focused on small sets of contrasts, and have not controlled for the quantity of prior or simultaneous exposure to new sounds. The current study examined the effectiveness of phonetic training in full-inventory foreign language consonant acquisition by listeners with no previous exposure to the language. Chinese adult listeners underwent an intensive training programme, bracketed by tests that measured both assimilation of foreign sounds to native categories, and foreign category identification rates and confusions. Very rapid learning was evident in the results, with initial misidentification rates halving by the time of the mid-test, and continuing to fall in subsequent training sessions. Changes as a result of training in perceptual assimilation together with improved identifications and reduced response dispersion suggest an expansion of listeners' native categories to accommodate the foreign sounds and an incipient process of foreign language category formation.

1. Introduction

Adult language learners have problems in perceiving and producing certain non-native phonemes. Together with age of acquisition and first language (L1) influence, the amount of foreign language (FL) exposure is crucial in learners' FL perception and production. However, FL exposure can be difficult to quantify. In immersion settings, length of residence is normally used as a proxy (Flege & Liu, 2001) since it is an easily measurable variable, although not necessarily highly predictive (DeKeyser & Larson-Hall, 2005), perhaps, as Flege (2009) suggests, because length of residence in itself does not necessarily reflect an individual's real amount or quality of input in the target language. Length of formal FL instruction maybe similarly unreliable. Increasingly, particularly for English, learners get exposed to the FL outside the classroom through music, films, internet and conversations. Under these circumstances, quantifying the total amount of exposure is challenging (Flege, 2009).

Training is one way of controlling and enhancing the amount of L2 input. High-variability phonetic training is effective in improving adults' foreign language perception (e.g., Lengeris & Nicolaidis, 2015; Lively, Logan, & Pisoni, 1993; Nishi & Kewley-Port, 2007). These studies have also shown that the use of a relatively small amount of high-variability speech material forces adult listeners to filter out natural variation (such as that produced by positional allophones and individual speaker

* Corresponding author. School of Foreign Languages, Jiangsu University of Science and Technology, 2 Mengxi Road, 212003 Zhenjiang, China.
E-mail address: j.gong@just.edu.cn (J. Gong).

http://dx.doi.org/10.1016/j.system.2017.03.018
0346-251X/

differences) and construct more robust phonetic categories for the target language, as attested by rapid improvement in their FL phoneme identification performance and generalisation to other contexts/voices. Most high-variability training studies have been carried out with experienced listeners who had been learning the target language for some time before the experiment, resulting in a large degree of variability in the amount and form of input prior to training. Even once training has started, the amount of additional exposure is difficult to control, especially in an immersion setting. Studies with adult listeners employing English as the target language further complicate quantification of prior exposure given the widespread use of English in the media. Prior knowledge and concurrent exposure are thus confounding factors in controlled studies of the role of input in phonetic category development.

An alternative approach is to train listeners on artificial language sounds or on languages previously unknown to listeners (e.g., Holt & Lotto, 2006). Among the few studies which have focused on the acquisition of entirely unknown phonetic categories, most have been based on a small set of contrasts. Golestani and Zatorre (2004) trained English speaking listeners to identify a Hindi dental–retroflex contrast. Similarly Pruitt, Jenkins, and Strange (2006) trained American English speakers and Japanese speakers to identify this contrast. Hirata, Whitehurst, and Cullings (2007) focused on training native English speakers to identify five Japanese vowel length contrasts. In the current study, we also investigate the acquisition of unknown phonetic categories through training, mainly as a way of controlling learners' competence and input as well as the intervention of higher-order linguistic processing. We include the full consonantal inventory of the target language (Spanish) to better observe mutual influences amongst phonetic categories within the Spanish inventory and across the two languages' sound systems.

More recently, high-variability phonetic training has been extended to a larger set of contrasts, mostly involving English vowels: Japanese learners, 9 vowels (Nishi & Kewley-Port, 2007); Spanish and German learners, 14 vowels (Iverson & Evans, 2009); French learners, 14 vowels (Iverson, Pinet, & Evans, 2012). Greek learners of English in a study by Lengeris and Nicolaidis (2015) did use consonants, but only for 7 exemplars. To the authors' knowledge, no studies have used the full set of consonants from an unknown target language, and consequently we are left with an incomplete picture of the mutual consonant confusions that can occur at an early stage of acquisition. The primary purpose of the current study is to examine, in a setting with tight control over input and prior exposure to the target language, the evolution of FL sound acquisition for a full consonant inventory, following the high-variability training paradigm. Here, we investigate the case of Chinese listeners exposed to high-variability training in an unknown language, viz. Spanish.

Few studies have investigated the perception of Spanish consonants for non-native listeners. Rose (2010, 2012) examined how native English learners of Spanish discriminate and assimilate four Spanish consonants (/r, r, t, d/) to English categories, finding Spanish /r-d/ to be the most difficult contrast for native English speakers. Spanish /r/ was assimilated to English /r/ most of the time, Spanish /r/ was assimilated to /r/ and /d/, and Spanish /d/ was assimilated to /l/, /d/ and /r/. More studies investigated the production of Spanish consonants by non-native learners. Diaz Campos (2004) compared the production of Spanish word-initial stops, intervocalic fricatives, word-final laterals and palatal nasals by native English speakers in L2 and FL settings. Face and Menke (2009) analyzed how native English learners of Spanish produce Spanish voiced plosives in intervocalic positions. Chen (2007) compared plosive productions by native Spanish speakers and Chinese L2 learners. Gonzalez Lopez and Counselman (2013) focused on the production of Spanish voiceless plosives by English speakers.

L1 influence is probably the single strongest factor in FL sound acquisition (Best, 1995; Flege, 1995, pp. 233–277; Kuhl, 1993) but a learner's phonological system is dynamic and can, in principle, change in its treatment of FL and L1 sounds (Best & Tyler, 2007; Heeren & Schouten, 2008). A learner's interpretation of the FL phonological system and FL to L1 assimilation patterns may also change as a function of experience (Major, 2001; Wu, Munro, & Wang, 2014). Changes in L1 vowel production have been observed as a result of short-term training on two unknown non-native vowels (Kartushina, Hervais-Adelman, Frauenfelder, & Golestani, 2016). In a cross-language vowel mapping experiment, Bundgaard-Nielsen, Best, and Tyler (2011) found that those Japanese learners of English with more English experience, as measured by a larger vocabulary size, had more concentrated English to Japanese assimilation patterns (i.e., individual English vowels were assimilated to fewer Japanese categories) than learners with less English experience. Bundgaard-Nielsen et al. (2011) claim that more experienced Japanese learners of English were 'reattuning and rephonologizing', by stretching their L1 categories to accommodate the FL vowel system, or through the formation of new categories for FL sounds. Similarly, in a study of French to English vowel assimilation, Levy (2009) found that native American English speakers with no French experience demonstrated a scattered assimilation pattern for French mid-front rounded /œ/, while those with extensive experience had more concentrated assimilation patterns for /œ/. Interestingly, similar assimilation concentrations were reported in a perceptual training study (Iverson & Evans, 2009), where native Spanish and German (L1) speakers' English (L2) vowel assimilation selections became more consistent on those target sounds with the largest assimilation percentages after high-variability auditory training. To explore the possibility that training can induce rephonologization[1], we examine here changes in assimilation patterns following intensive exposure.

The current study focuses on two research questions. First, is high-variability phonetic training effective *ab initio* in the challenging setting where all FL categories are present? Second, do listeners' non-native to native assimilation patterns change as a function of experience through intensive training? To ensure tight control over FL input prior to and during the

[1] According to Bundgaard-Nielsen et al. (2011), rephonologization means to "establish an L2 phonology by modification of or addition to the learner's existing L1 phonological system".