



# Can intensive exposure to foreign language sounds affect the perception of native sounds?

Jian Gong<sup>1</sup>, María Luisa García Lecumberri<sup>2</sup>, Martin Cooke<sup>3,2</sup>

<sup>1</sup>School of Foreign Languages, Jiangsu University of Science and Technology, Zhenjiang, China

<sup>2</sup>Language and Speech Lab, University of the Basque Country, Vitoria-Gasteiz, Spain

<sup>3</sup>Ikerbasque, Bilbao, Spain

j.gong@just.edu.cn, garcia.lecumberri@ehu.es, m.cooke@ikerbasque.org

## Abstract

A possible side-effect of exposure to non-native sounds is a change in the way we perceive native sounds. Previous studies have demonstrated that native speakers' speech production can change as a result of learning a new language, but little work has been carried out to measure the perceptual consequences of exposure. The current study examined how intensive exposure to Spanish intervocalic consonants affected Chinese learners with no prior experience of Spanish. Before, during and after a training period, listeners undertook both an adaptive noise task, which measured the noise level at which listeners could identify native language consonants, and an assimilation task, in which listeners assigned Spanish consonants to Chinese consonant categories. Listeners exhibited a significantly reduced noise tolerance for the Chinese consonants /l/ and /w/ following exposure to Spanish. These two consonants also showed the largest reductions in Spanish to Chinese category assimilations. Taken together, these findings suggest that Chinese listeners modified their native language category boundaries as a result of exposure to Spanish sounds in order to accommodate them, and that as a consequence their identification performance in noise reduced. Some differences between the two sounds in the time-course of recovery from perceptual adaptation were observed.

**Index Terms:** consonants, foreign language, perception

## 1. Introduction

Language learners are required to adapt to sounds that differ from those in their native inventory. This dynamic process has been the subject of a number of models [1, 2, 3], all of which argue that perceived similarities between native and non-native phonetic categories play a key role in the perception of non-native sounds. Thus, when faced with phonetic categories from another language, learners may interpret them in terms of their own phonological system or as separate categories to those they already possess. Exposure and perceptual training [4, 5] are important for listeners' adaptation to non-native categories, which can ultimately result in foreign language perceptual categories becoming more native-like [6].

The finding that listeners show perceptual adaptation to foreign language (FL) sounds raises the issue of whether exposure to such exemplars might also affect *native* categories, perhaps leading to a loss of native (L1) competence – a form of the more general phenomenon of linguistic attrition. A series of speech production experiments has demonstrated two related phenomena, assimilation and dissimilation [3, 7], in which bilinguals' L1 and second language (L2) categories depart from both L1

and L2 norms, either because a compromise category is being used for the two languages or, if a separate category has been created for the L2 sound, in order to increase the difference between the L2 and the L1 similar sounds [7, 8].

Other studies have shown that adults' L1 phonetic categories can be changed in FL immersion environments e.g., [9, 10]. Although changes in learners' L1 categories might relate to the dominance of FL usage – usually the case in immersion settings – several recent studies show FL influence on the L1 in inexperienced learners. For instance, [11] found that even with low fluency in English, young Japanese learners' L1 VOT showed dissimilation-like drift from L1 norms. Similarly, adult native English speakers demonstrated some drift in L1 plosive and vowel production during the first few weeks of Korean learning [12]. Phonetic drift was greater for inexperienced FL learners than for experienced learners [13], suggesting that the influence of the FL on a learner's L1 might be stronger in the early stages of learning.

Fewer studies have examined possible *perceptual* effects of FL exposure on L1 sounds. While brief exposure to ambiguous or dialectally-different L1 sounds is known to affect the perception of L1 phonetic categories [4, 5, 14], to our knowledge only one study has investigated whether brief exposure to FL sounds can also affect perception of L1 categories. In [15], a group of novice English learners of French from an intensive French course participated in a 6-week longitudinal study in which their perceptual boundaries for VOT in voiced/voiceless plosives were monitored weekly using phoneme categorisation and semantic priming tests. Learners showed perceptual boundary shifts towards French after only a few weeks, providing some evidence that brief but intense non-native exposure can alter native category perception.

One approach to assessing modifications to perceptual boundaries from FL learning is via a perceptual assimilation task (during which listeners are asked to identify FL sounds in terms of L1 categories) by examining whether assimilation patterns change with further exposure to FL sounds. Recent cross-language vowel mapping experiments [16] found that Japanese learners of English with more English experience, as measured by a larger vocabulary size, had more concentrated English to Japanese vowel assimilation patterns (i.e., individual English vowels were assimilated to fewer Japanese categories) than learners with less English experience. The authors claim that this outcome is evidence that more experienced Japanese learners of English were 're-phonologising', either by stretching their L1 categories to accommodate the FL vowel system, or through the formation of new categories for FL sounds. A study of French to English assimilation [17] also showed an

experience-based difference. Similar assimilation concentrations have also been reported for non-native vowel perception as a result of training [18].

The current study explored possible changes in the perception of L1 (Chinese) consonants as a result of intensive exposure to FL (Spanish) consonants in intervocalic context, using a pre-test/training/post-test design, via two tasks. First, L1 category identification was assessed using a consonant-specific adaptive noise level procedure. Our hypothesis is that any loss in L1 performance will show up as a lower tolerance to noise for specific consonants. Second, an assimilation experiment measured the assignment of Spanish sounds to Chinese categories. Given previous results on FL to L1 vowel assimilation patterns [16, 17, 18], we hypothesise that consonants will show similar perceptual assimilation changes due to listeners' adaptation to FL sounds as a result of FL exposure.

## 2. Methods

The current analysis is based on data collected for a larger study into the rapid learning of FL consonants [19]. Here, we focus on the L1 category identification task and the comparison with the outcome of the assimilation task. The current section reviews the testing and training protocols and materials used in [19] and introduces the L1-in-noise and assimilation tasks.

### 2.1. Participants

Twenty native Mandarin Chinese listeners (11 females; mean age 20.5) from the Southern Medical University in China took part in the study. Selection criteria were (i) no prior study of Spanish or any other Romance language; (ii) no time spent outside China; (iii) no hearing problems; (iv) Northern Mandarin dialect. Listeners were paid for taking part. A cohort of eight listeners with a similar profile were recruited as a control group.

### 2.2. Speech materials

Both L1 (Chinese) and FL (Spanish) materials were consonants presented in an intervocalic context i.e. VCVs. The same vowel contexts were used in both languages, namely all 9 combinations of the 3 corner vowels /a,i,u/.

#### 2.2.1. Spanish

The 18 consonant phonemes /p, b, t, d, k, g, tʃ, f, θ, s, x, m, n, ɲ, l, r, r, j/ were recorded in VCVs with both initial and final vowel stress (324 combinations in all: 18 consonants × 9 vowel contexts × 2 stresses). Sixteen male native Spanish speakers each produced the 324 combinations.

#### 2.2.2. Chinese

Chinese tokens were drawn from an earlier study [20] which used the 24 consonants /p<sup>h</sup>, p, t<sup>h</sup>, t, k<sup>h</sup>, k, ts<sup>h</sup>, ts, tʃ<sup>h</sup>, tʃ, tɕ<sup>h</sup>, tɕ, f, s, ʃ, ɕ, x, m, n, ŋ, ɹ, l, j, w/. Speakers produced both initial and final vowels in the first lexical tone. Since lexical stress is not a feature of Chinese, the corpus does not differentiate the two stress types. Seventeen male native Chinese speakers read each of the 216 tokens of the corpus (24 consonants × 9 vowel contexts).

Both corpora were manually screened to identify and remove mis-produced or noisy items, leading to final corpora of 4931 VCV tokens for Spanish and 3331 for Chinese. Tokens from 10 speakers were used in the training sessions, while mate-

rial from the remaining 6 speakers was used for the L1-in-noise and assimilation tests.

### 2.3. Tasks

Listeners underwent a sequence of tasks prior to (pre-test), during (mid-test), and following (post-test) a period of intensive exposure to Spanish consonants. At each of the three test phases participants identified Chinese consonants in noise (L1-in-noise test), categorised Spanish consonants using Chinese categories (assimilation test); and identified Spanish consonants in quiet (FL test). The focus of the current article is on the L1 test and its relation with perceptual assimilation; the FL test is not described further here. Listeners in the control undertook the pre- and mid-tests with a gap of two days, and no training.

#### 2.3.1. Exposure with feedback

The exposure period consisted of 16 sessions spread over four days. In each session, listeners identified 180 Spanish VCV tokens (10 examples of each of the 18 consonants); no token was repeated during the 16 sessions. In this way, each listener had a total exposure of 2880 different tokens. Listeners responded to each VCV by selecting one of 18 consonants from an on-screen keyboard, receiving feedback and the obligation to listen to the stimulus exactly once more for any erroneous responses.

#### 2.3.2. L1-in-noise test

Two considerations motivated the use of a consonants-in-noise task. First, since we are considering possible perceptual changes at the phonetic level, it is reasonable to assume that any effects will show up in a forced-choice task devoid of higher-level factors. Second, masking noise is required to remove ceiling effects; further, an adaptive noise threshold task is sensitive enough to reveal subtle distinctions. For instance, [21] demonstrated perceptual deficits for highly-competent bilinguals in a speech-in-noise task, effects that were not evident when higher-level contextual information was available.

Per-consonant signal-to-noise ratios (SNRs) for a criterion level of identification performance were estimated by presenting L1 tokens in speech-shaped noise whose level was changed using a 2-down 1-up adaptive procedure [22]. We refer to the converged SNRs as Consonant Reception Thresholds (CRT) by analogy with Speech Reception Thresholds [23]. Consonants were presented in a random order during the test, so that thresholds were estimated for all consonants in parallel. The initial SNR was set at +4 dB for all consonants, and a fixed step size of 2 dB was used in the adaptive procedure. The CRT was estimated as the mean of the SNRs used in the final five presentations.

Stimuli for the L1 test were 20 examples of each of the 24 consonants, obtained by random sampling from the 17 speakers of the Chinese VCV corpus. Since a previous study [20] found that Latin alphabet representations of Chinese phonetic categories produced orthographic biases on perceived categories, listeners selected their response from an onscreen grid which displayed Chinese characters with the corresponding consonant in initial position. An exception was /ŋ/, which was presented in final position to meet phonotactic constraints. The task required around 35 minutes.

#### 2.3.3. Assimilation test

The assimilation task involved mapping Spanish VCVs to Chinese consonant categories. A total of 360 Spanish VCVs were

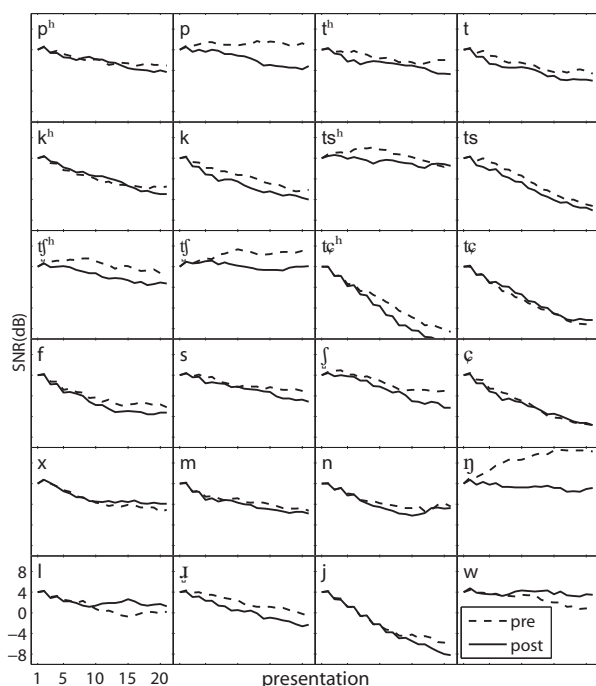


Figure 1: Per-consonant SNR evolution in pre and post-test.

categorised along with 48 Chinese VCV tokens which served as control items. Listeners were presented with VCVs in quiet and responded by selecting the nearest Chinese category using an onscreen keyboard.

## 2.4. Procedure

All tasks took place in a quiet computer laboratory at the Southern Medical University in China. Participants were tested simultaneously. Stimuli were presented using AC 97 sound cards and SALAR A522 headphones. Stimulus presentation and response collection was controlled by custom MATLAB programs. All stimuli were normalised to have equal RMS energy prior to presentation. Participants were allowed to adjust the volume to a comfortable listening level.

## 3. Results

### 3.1. L1-in-noise test

Figure 1 shows the change in presentation SNR for each consonant as a function of the stimulus presentation number during pre- and post-test. Consonant reception thresholds were estimated by computing the across-listener mean threshold at each point in the presentation sequence, and then taking the mean over the final five presentations as the CRT value. For the majority of consonants the CRT was well-estimated by this procedure. However, for the consonants /ts, tʃ^h, tʃ, ɕ, ʃ, j/ it was not possible to estimate a reliable consonant reception threshold due to lack of convergence.

For 15 of the remaining 18 consonants, converged thresholds remained stable or fell from pre- to post-test, indicating no change or an improvement in identifying first language consonants in noise. However, the three consonants /x, l, w/ showed an elevated threshold in the post-test, although the effect was not statistically-significant for /x/. Figure 2 displays SNR evolu-

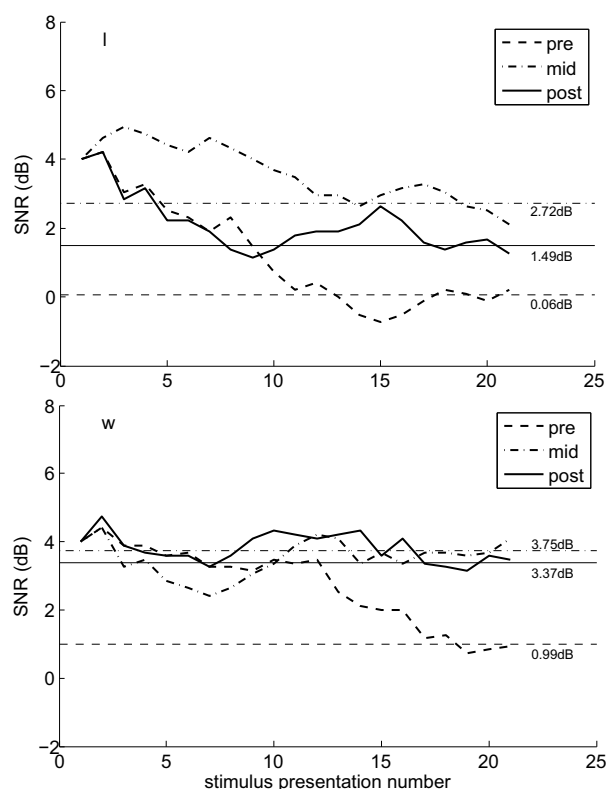


Figure 2: SNR evolution and final consonant reception thresholds (horizontal lines) for /l/ and /w/ in pre-, mid- and post-tests.

tion and estimated thresholds for /l/ and /w/ during the three test phases. For /l/, only the 2.7 dB change from pre-test to mid-test exceeded Fisher's Least Significant Difference (LSD) of 2.1 dB. Here, there is a clear difference between the pre- and mid-test SNR evolution throughout the entire presentation range. For /w/, the threshold obtained in pre-test was significantly lower than either the mid- or post-tests, exceeding Fisher's LSD of 2.2 dB. Listeners in the control group showed no significant difference in CRTs from pre- to mid-test for either sound (rise of 0.55 and drop of 0.69 dB for /l/ and /w/ respectively). None of the consonants that failed to converge show elevated CRTs, although the possibility cannot be ruled out for /tʃ/.

### 3.2. Assimilation test

Figure 3 depicts changes in assimilations from pre- to post-test. One striking feature is that the largest reductions in assimilations occur for the Chinese phonemes /l/ and /w/. This is confirmed by Table 1, which shows all reductions of 10 percentage points or more; the six largest reductions occur for /l/ and /w/. Chinese /l/ is a frequent target category in the assimilation test, occurring among the dominant responses for several Spanish categories in the initial assimilation task, and remaining in four categories in the post-test. It seems likely that exposure to the diverse collection of FL sounds which are deemed to be potential sharers of a single L1 category for /l/, might well exert some influence on the perceptual boundaries of the L1 category.

Some of the large assimilation reductions resulted in a Spanish sound strongly increasing its assimilation to one other Chinese category. For example, /b/ and /g/, strongly assimilated to /w/ in the pre-test, were assimilated to /p/ and /k/ respectively

Table 1: *Largest pre-post reductions in assimilation.*

Spanish	Chinese	pre (%)	post (%)	$\Delta$
b	w	65	17	-48
g	w	29	-	-29
r	l	59	32	-27
r	l	84	62	-22
g	l	21	-	-21
d	l	30	13	-17
s	s	53	42	-11
d	ts	38	28	-10

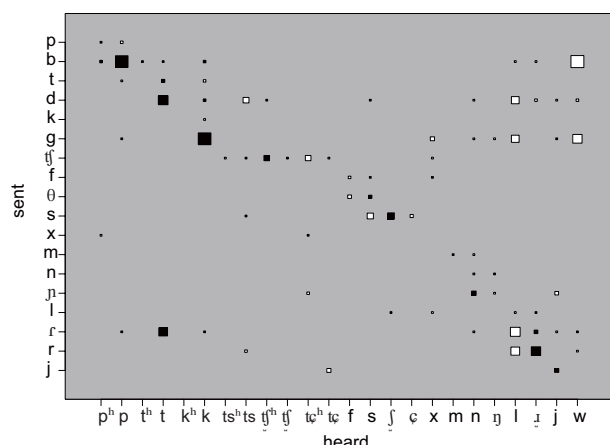


Figure 3: *Pre-post difference in assimilation percentages. The area of each square is proportional to the percentage points pre-post change in assimilation for a FL (sent), L1 (heard) pair (white: reductions; black: increases).*

in the post-test, reaching a level at which they can be considered to be well-categorised within those sounds (i.e., over 70 % of the time, following the stringent criterion adopted in [16]). In the case of the five Spanish sounds originally assimilated to /l/, even though after training there is a significant reduction in their assimilation to /l/, only Spanish /l/ and /d/ can be said to have been well-categorised according to this criterion.

## 4. Discussion

The purpose of the current study was to determine whether the identification of L1 segments is affected by intensive perceptual exposure to FL sounds. Many L1 consonants displayed a higher noise tolerance following training, as indicated by a reduction in CRTs from pre- to post-test. This fall may be due to task familiarity or exposure to the masker, reflecting procedural learning [24]. In this light, the finding of statistically-significant *increases* in CRT for the Chinese sounds /w/ and /l/ is highly-suggestive of an effect of intensive exposure to foreign language (Spanish) sounds on the perception of native categories.

Assimilation tests also led to changes in FL category identifications in terms of L1 Chinese consonants. Several sounds displayed a reduction in FL category dispersion from pre- to post-test and strong tendencies towards identification with a single L1 consonant, demonstrating the effectiveness of the training regime. Such concentrations of FL to L1 assimilations through experience have been found in other studies [16], and our results confirm this process is also present in consonants.

Equally interesting is the finding that the strongest dispersion reductions in FL to L1 consonant mappings were manifested by /l/ and /w/, the two consonants with a significant CRT increase.

Taken together, the CRT increase and decrease in FL to L1 assimilation for /w/ and /l/ provide strong indications that listeners were re-phonologising FL sounds, in part by stretching L1 categories to accommodate them, following training sessions that presumably increased awareness of FL consonant cues [16].

The change of CRTs from pre-test to post-test shows a different pattern for /l/ and /w/. For the lateral, the increase in threshold is highest at the point of the mid-test, and shows some recovery by the time of the post-test. This might indicate ongoing category modification that is incomplete by the mid-test. For /w/ the pattern of SNR evolution is quite consistent at mid- and post-test, suggesting either that any changes are fully in place by the time of the mid-test, or that further exposure is required for the CRT to revert to that seen in the pre-test. These findings suggest that short-term perceptual learning – perhaps of the kind seen in lexically-guided perceptual adaptation [14] – is at work even for the sub-lexical categories of the current study. The time-course of recovery from any perceptual changes to the L1 due to FL exposure remains an important open question. [15] found that in a phoneme categorisation task, the most significant L1 ‘drift’ occurred in the middle of a 6-week FL learning course, while learners had returned to their earlier behaviour by the latter part of the training regime.

As mentioned in the Introduction, other studies have demonstrated clear effects of early stage FL learning on L1 speech production [11, 12, 13]. While the principal distinction is that our study investigated perception, other differences between the studies may be relevant. Unlike the current *ab initio* study, the aforementioned works investigated FL learners engaged in acquiring the language for communicative purposes. Apart from type of exposure, learners age and competence have also been found to be important factors in L1 category changes. [11] tested children aged 4-6, and it is likely that their L1 phonological system was not fully mature, making FL influence more plausible. FL competence may also be relevant in L1 category effects. [15] showed that native perceptual boundaries of novice learners can be shifted during FL learning. [12] also demonstrates that this kind of phonetic drift in L1 category productions occurs in early stage adult FL learners, but less strongly for experienced FL learners [13]. Our study provides further evidence of L1 consonant perceptual adaptation to FL sounds by inexperienced learners. More work is needed to determine if further training increases the number of L1 sounds whose perception is affected by FL exposure.

## 5. Conclusions

Chinese listeners with no prior exposure to Spanish showed significant elevations of the noise threshold required to identify two native language consonants following an intensive period of exposure to Spanish consonants. These consonants also showed the strongest reductions in assimilations, suggesting that learners were modifying their L1 categories to accommodate the new sounds. These findings indicate that exposure to foreign language sounds can influence the perception of native language sounds, at least in the short term.

**Acknowledgements:** This study was supported by Jiangsu Social Science Fund grant No. 14YYC001, the Spanish Ministry MICIN DIACEX project and the Basque Government Consolidados grant.

## 6. References

- [1] P. K. Kuhl, "Innate predispositions and the effects of experience in speech perception: The native language magnet theory," in *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*, B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage, and J. Morton, Eds. Berlin: Springer, 1993, pp. 259–274.
- [2] C. T. Best, "A direct realist view of cross-language speech perception," in *Speech perception and linguistic experience: Theoretical and methodological issues*, W. Strange, Ed. Baltimore: York Press, 1995, pp. 171–204.
- [3] J. E. Flege, "Second-language speech learning: Theory, findings, and problems," in *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research*, W. Strange, Ed. Baltimore: York Press, 1995, pp. 233–277.
- [4] C. M. Clarke and M. Garrett, "Rapid adaptation to foreign-accented English," *Journal of the Acoustical Society of America*, vol. 116, no. 6, pp. 3647–3658, 2004.
- [5] A. R. Bradlow and T. Bent, "Perceptual adaptation to non-native speech," *Cognition*, vol. 106, no. 2, pp. 707–729, 2008.
- [6] W. F. L. Heeren and M. E. H. Schouten, "Perceptual development of phoneme contrasts: How sensitivity changes along acoustic dimensions that contrast phoneme categories," *Journal of the Acoustical Society of America*, vol. 124, no. 4, pp. 2291–2302, 2008.
- [7] J. E. Flege, C. Schirru, and I. R. A. MacKay, "Interaction between the native and second language phonetic subsystems," *Speech Communication*, vol. 40, no. 4, pp. 467–491, 2003.
- [8] G. Oh, S. Guion-Anderson, K. Aoyama, J. Flege, R. Akahane-Yamada, and T. Yamada, "A one-year longitudinal study of English and Japanese vowel production by Japanese adults and children in an English-speaking setting," *Journal of Phonetics*, vol. 39, no. 2, pp. 156–157, 2011.
- [9] J. E. Flege, "The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification," *Journal of Phonetics*, vol. 15, no. 1, pp. 47–65, 1987.
- [10] E. de Leeuw, M. S. Schmid, and I. Mennen, "The effects of contact on native language pronunciation in an L2 migrant setting," *Bilingualism: Language and Cognition*, vol. 13, no. 1, pp. 33–40, 2010.
- [11] N. Yusa, K. Nasukawa, M. Koizumi, J. Kim, N. Kimura, and K. Emura, "Unexpected effects of the second language on the first," in *New Sounds 2010: Proceedings of the 6th international symposium on the acquisition of second language speech*, K. Dziubalska-Kolaczyk, M. Wrembel, and M. Kul, Eds. Poznan, Poland: Adam Mickiewicz University, 2010, pp. 580–584.
- [12] C. B. Chang, "Rapid and multifaceted effects of second-language learning on first-language speech production," *Journal of Phonetics*, vol. 40, no. 2, pp. 249–268, 2012.
- [13] —, "A novelty effect in phonetic drift of the native language," *Journal of Phonetics*, vol. 41, no. 6, pp. 520–533, 2013.
- [14] D. Norris, J. M. McQueen, and A. Cutler, "Perceptual learning in speech," *Cognitive Psychology*, vol. 47, no. 2, pp. 204–238, 2003.
- [15] M. Tice and M. Woodley, "Paguettes & bastries: Novice French learners show shifts in native phoneme boundaries," in *86th Annual Meeting of the Linguistic Society of America*, 2012, pp. 72–75.
- [16] R. L. Bundgaard-Nielsen, C. T. Best, and M. D. Tyler, "Vocabulary size is associated with second-language vowel perception performance in adult learners," *Studies in Second Language Acquisition*, vol. 33, no. 3, pp. 433–461, 2011.
- [17] E. S. Levy, "Language experience and consonantal context effects on perceptual assimilation of French vowels by American-English learners of French," *Journal of the Acoustical Society of America*, vol. 125, no. 2, pp. 1138–1152, 2009.
- [18] P. Iverson and B. G. Evans, "Learning English vowels with different first-language vowel systems II: Auditory training for native Spanish and German speakers," *Journal of the Acoustical Society of America*, vol. 126, no. 2, pp. 866–877, 2009.
- [19] J. Gong, "Computational modelling of sound perception in a second language," Ph.D. dissertation, University of the Basque Country, 2013. [Online]. Available: [http://www.laslab.org/resources/Thesis\\_JianGong.pdf](http://www.laslab.org/resources/Thesis_JianGong.pdf)
- [20] J. Gong, M. Cooke, and M. L. García Lecumberri, "Towards a quantitative model of Mandarin Chinese perception of English consonants," in *New Sounds 2010: Proceedings of the 6th international symposium on the acquisition of second language speech*, K. Dziubalska-Kolaczyk, M. Wrembel, and M. Kul, Eds. Poznan, Poland: Adam Mickiewicz University, 2010, pp. 580–584.
- [21] L. Mayo, M. Florentine, and S. Buus, "Age of second-language acquisition and perception of speech in noise," *Journal of Speech, Language and Hearing Research*, vol. 40, pp. 686–693, 1997.
- [22] H. Levitt, "Transformed up-down methods in psychoacoustics," *Journal of the Acoustical Society of America*, vol. 49, no. 2B, pp. 467–477, 1971.
- [23] R. Plomp and A. Mimpén, "Improving the reliability of testing the speech reception threshold for sentences," *International Journal of Audiology*, vol. 18, no. 1, pp. 43–52, 1979.
- [24] D. L. Woods, Z. Doss, T. J. Herron, T. Arbogast, M. Younus, M. Ettlinger, and E. W. Yund, "Speech perception in older hearing impaired listeners: benefits of perceptual training," *PloS one*, vol. 10, no. 3, p. e0113965, 2015.