PERCEPTUAL ADAPTATION FOR L1 AND L2 ACCENTS IN NOISE BY MONOLINGUAL BRITISH ENGLISH LISTENERS

Melanie Pinet, Paul Iverson & Bronwen G. Evans

Department of Speech, Hearing and Phonetic Sciences, UCL, UK
m.pinet@ucl.ac.uk

ABSTRACT

Native (L1) listeners are more accurate at understanding speech spoken in their own accent in noise than they are at understanding speech spoken in other L1 or non-native (L2) accents. The present study investigated whether this accent advantage is affected by perceptual adaptation. Standard Southern British English (SE) listeners were presented with L1 and L2 accents in noise, either in single-talker blocks (i.e., one accent, designed to promote adaptation) or in mixed blocks. The results demonstrated that there was significant adaptation to accents, but that there was no interaction between the type of accent and the amount of learning. That is, listeners adapted similarly to all accents, regardless of whether they were familiar or unfamiliar, or L1 or L2. This suggests that listeners' advantage for familiar L1 accents is retained even when listeners have time to tune into the accents of a particular talker.

Keywords: accent adaptation, talker-listener interaction, speech perception, noise

1. INTRODUCTION

Speech recognition in noise critically depends on an interaction between the accents of the talker and the listener. Native listeners can accurately recognize a variety of native (L1) and non-native (L2) accents in quiet, but they have a clear advantage for L1 over L2 speech in noisy listening conditions. This effect is greatly enhanced when the listener's accent matches the talker's [2, 9, 10]. For instance, Pinet, et al. [9] presented Standard Southern British English 'SE' listeners with a mixed variety of L1- and L2-accented sentences in both quiet and noisy listening conditions. The listeners performed equally well on L1 accents in quiet, but in noise, their recognition processes became selectively tuned to their own accent. That is, they had similarly low levels of intelligibility for unfamiliar L1 and L2 accents, and only had an intelligibility advantage for L1 accents that matched their own spoken accent.

However, it is plausible that accent adaptation can overturn this selective tuning process. Previous work has shown that there is an initial processing cost associated with exposure to an unfamiliar accent, followed by decreased reaction times over the first few trials as the listeners adapt [4]. Intelligibility improves over a slightly longer time frames (e.g., 40 trials) than do reaction times, such that listeners typically improve in their recognition accuracy by about 5-15 percentage points [3, 7].

The aim of the present study was to investigate whether L1 listeners' accent selectivity [9] can be reversed when given the opportunity to tune into unfamiliar accents in single accent blocks, so that they can more easily adapt. Listeners were presented with an L1 accent that matched their own (SE), a relatively unfamiliar L1 accent (Northern Irish English, IE), and two L2 accents (French- and Korean-accented English, FE and KE). The accents were presented in single blocks as well as in a mixed accent block, in order to evaluate whether individuals are broadly improving on the task within the block or specifically adapting to single accents.

2. METHOD

2.1. Subjects

The subjects were 18 monolingual Standard Southern British English listeners (SE), aged 22 to 35 (mean: 28 years). None of the subjects reported any speech, hearing or learning difficulties.

One female talker of Standard Southern British English (SE), Northern Irish English (IE), French-accented English (FE) and Korean-accented English (KE) were recorded reading the Bamford-Kowal-Bench (BKB) sentences (e.g., *The green tomatoes were small*, *They laughed at his story*) [1]. The French talker had low English proficiency; she had learned English at school and was residing in France at the time of recording. The Korean talker was a low proficiency speaker who was residing in the UK. The digitized recordings were embedded in speech-shaped noise

with a signal-to-noise ratio from -1 to -5 dB; the exact values were selected for individual talkers based on previous data, in an attempt to equate intelligibility levels (a target of 70% correct words in sentences). The speech-shaped noise was generated for each individual talker such that it matched the smoothed long-term average spectrum of their speech.

2.2. Procedure

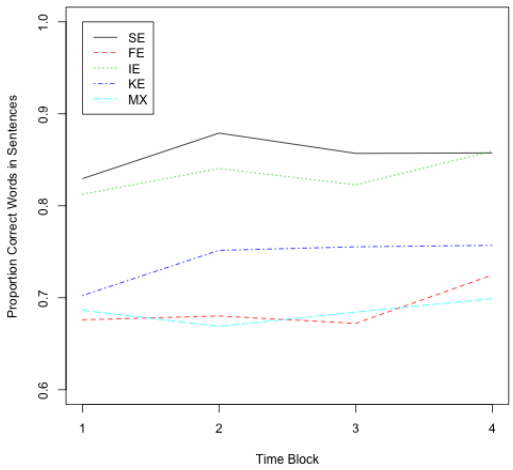
The subjects performed a sentence recognition task where they listened to the stimuli and repeated what they had heard. Responses were given verbally (i.e., the experimenter marked how many keywords were spoken correctly). Each block contained 56 sentences and each sentence was presented only once (i.e., they were not repeated within or across conditions). The stimuli were presented in a random order within each single and mixed accent block, and presentation order was counterbalanced between subjects.

3. RESULTS

Figure 1 displays recognition accuracy (i.e., the proportion of words correctly identified in the sentences) for the four accents and the mixed-accents block across time (divided into four time periods). A mixed-effects ANOVA was conducted with time and accent condition as within-subject factors. All analyses were conducted on arcsine-transformed scores. The results revealed significant main effects of time, $F(1, 685) = 6.28$, $p = .01$ and accent, $F(4, 685) = 51.13$, $p < .01$, but no interaction between the two. Listeners thus performed differently on the different accents (i.e., the levels of noise used did not fully equate intelligibility differences). The adaptation effect (i.e., change over time) was small (see Figure 1), but significant. However, the lack of interaction shows that they adapted to all accents in a similar manner. Therefore, the overall moderate effect of learning indicates no additional learning of the unfamiliar accents over their own.

Figure 2 displays boxplots of the listeners' performance on the four accents in single versus mixed accent blocks. The listeners performed worse on the mixed-accent block over the same single-accent blocks, showing that adaptation was facilitated by a continuous exposure to a single accent. A mixed-effects ANOVA was conducted with blocking and accent as within-subject factors. The results revealed significant main effects of blocking, $F(1, 115) = 35.97$, $p < .01$, and accent, $F(3, 115) = 17.76$, $p < .01$, and a significant interaction between the two, $F(3, 115) = 2.83$, $p < .05$. Overall, the results thus demonstrate that there were significant advantages for individuals listening to only a single talker and accent within each block, which demonstrates accent adaptation. The significant interaction likely occurred because there was a little less of a difference for the FE accent between the two types of blocking.

Figure 1: Proportion of correctly identified words for each accent condition over time. MX indicates the mixed accent block.

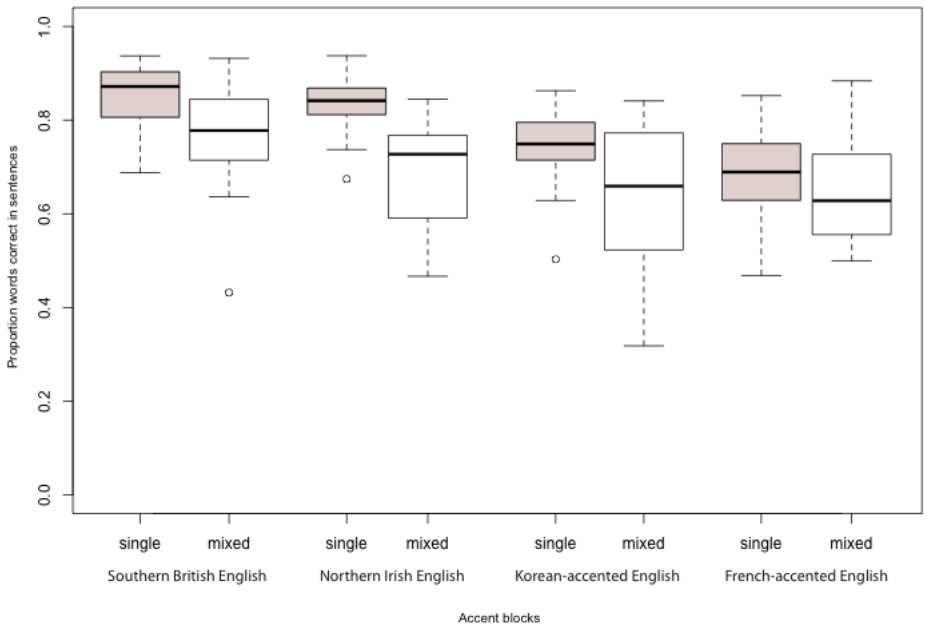


4. DISCUSSION

The results demonstrated a significant but moderate overall learning effect for all four accents. Interestingly, listeners showed similar amounts of learning for all accents, including their own accent, demonstrating that adaptation is a pervasive, if small in magnitude, effect [7, 8]. However, the uniformity in learning also indicates that the selectivity for an accent near one's own continues to occur even when listeners are allowed to adapt to individual accents and talkers.

One possible explanation for this homogenous learning effect is that the perceptual adaptation is talker-specific rather than accent-specific. Bradlow and Bent [3] found that listeners, at least under some conditions, are able to generalize their accent adaptation to new talkers with the same accent, as long as they are exposed to multiple talkers during adaptation. It is thus possible that accent-adaptation can occur in addition to more talker-specific adaptation. However, in the present study we were unable to find any interaction between the talkers' and listeners'accents.

Figure 2: Boxplots comparing the listeners' performance on the four accents in single versus mixed accent presentation conditions. Boxplots display the quartile ranges of scores.



Moreover, the results showed that the amount of adaptation to the accents was quite minimal. One possible explanation for this small learning effect is the period of exposure to the accent. The listeners were exposed to 56 BKB sentences per block, representing 6 to 7 minutes of exposure to the accented talker. Clark and Garrett [4] have shown very rapid adaptation to a novel accent (within 1 minute of exposure), but it could be that adaptation continues beyond this initial rapid learning effect. Also, Clark and Garrett [4] measured reaction times rather than changes in intelligibility. The accent-specific improvements in intelligibility found by Bradlow and Bent [3] included training that was split across two days, although the number of sentences was similar to that used here. It is possible that learning effects are better consolidated when listeners have more time to process the accent exposure. It is also plausible, however, that rapid adaptation effects are generally quite small, and that it may take much longer-term exposure (e.g., living in a community that speaks that accent) in order to perceive a novel accent as well as ones own [5, 6].

Blocking was also shown to affect the listeners' adaptation processes. Indeed, the listeners performed significantly better on the accents when they were presented in single-accent block compared to when they were mixed with other accents. However, there is no clear indication that the blocking design changed the SE listeners' accent selectivity, because the difference between mixed and single accents did not reliably vary depending on whether the accent was familiar or unfamiliar. There was less of a difference for the FE accent, but this talker was also less intelligible overall under the selected noise levels, and the intelligibility level may affect the degree of learning [3].

To conclude, the present study investigated L1 listeners' ability to reverse their selective tuning to their own accent through the process of adaptation. The results showed a moderate overall learning effect, indicating a persistent selective tuning for the listeners' own accent that held across adaptation conditions.

5. ACKNOWLEDGEMENTS

The project research was supported by the MRC and ESRC research councils of the UK.

6. REFERENCES

[1] Bench, J., Kowal, A., Bamford, J. 1979. The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *British Journal of Audiology* 13, 108-112.

[2] Bent, T., Bradlow, A.R. 2003. The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America* 114, 1600-1610.

[3] Bradlow, A.R., Bent, T. 2007. Perceptual adaptation to non-native speech. *Cognition* 106, 707-729.

[4] Clarke, C.M., Garrett, M.F. 2004. Rapid adaptation to foreign-accented English. *Journal of the Acoustical Society of America* 116, 3647-3658.

[5] Evans, B.G., Iverson, P. 2004. Vowel normalization for accent: an investigation of best exemplar locations in northern and southern British English sentences. *Journal of the Acoustical Society of America* 115(1), 352-361.

[6] Evans, B.G., Iverson, P. 2007. Plasticity in vowel perception and production: A study of accent change in young adults. *Journal of the Acoustical Society of America* 121(6), 3814-3826.

[7] Evans, B.G., Taylor, E. 2010. Investigating the time course of perceptual adaptation to unfamiliar accented speech. *BAAP 2010* London.

[8] Floccia, C., Butler, J., Goslin, J., Ellis, L. 2009. Regional and foreign accent processing in English: Can listeners adapt? *Journal of Psycholinguistic Research* 38, 378-412.

[9] Pinet, M., Iverson, P., Huckvale, M. 2011. Second language experience and speech-in-noise recognition: Effects of talker-listener accent similarity. *Journal of the Acoustical Society of America* under revisions.

[10] van Wijngaarden, S.J., Steeneken, H.J.M., Houtgast, T. 2002. Quantifying the intelligibility of speech in noise for non-native listeners. *Journal of the Acoustical Society of America* 111, 1906-1916.