

# INTELLIGIBILITY OF BRITISH ENGLISH ACCENTS IN NOISE FOR SECOND-LANGUAGE LEARNERS

*Melanie Pinet, Yuanlin Gan, Bronwen Evans, and Paul Iverson*

UCL Department of Speech, Hearing and Phonetic Sciences

m.pinet@ucl.ac.uk

## ABSTRACT

Standard or prestige accents (e.g., southern British RP, General American) are highly intelligible to a range of native speakers, which may be caused by listener experience (e.g., the accents are frequently heard in the media). The present study investigated the role of experience by testing Mandarin speakers with intermediate English proficiency in terms of their speech-in-noise recognition for a wide range of British English accents. The results demonstrated that Mandarin listeners were highly correlated with native British English listeners in terms of the relative intelligibility of individual talkers, with both groups finding standard accents to be more intelligible. It is thus plausible that the intelligibility of standard accents might be relatively independent from experience; standard accents may tend to have acoustic-phonetic features that make them inherently more intelligible.

**Keywords:** accent, talker-listener interaction, speech perception, noise, acoustic similarity.

## 1. INTRODUCTION

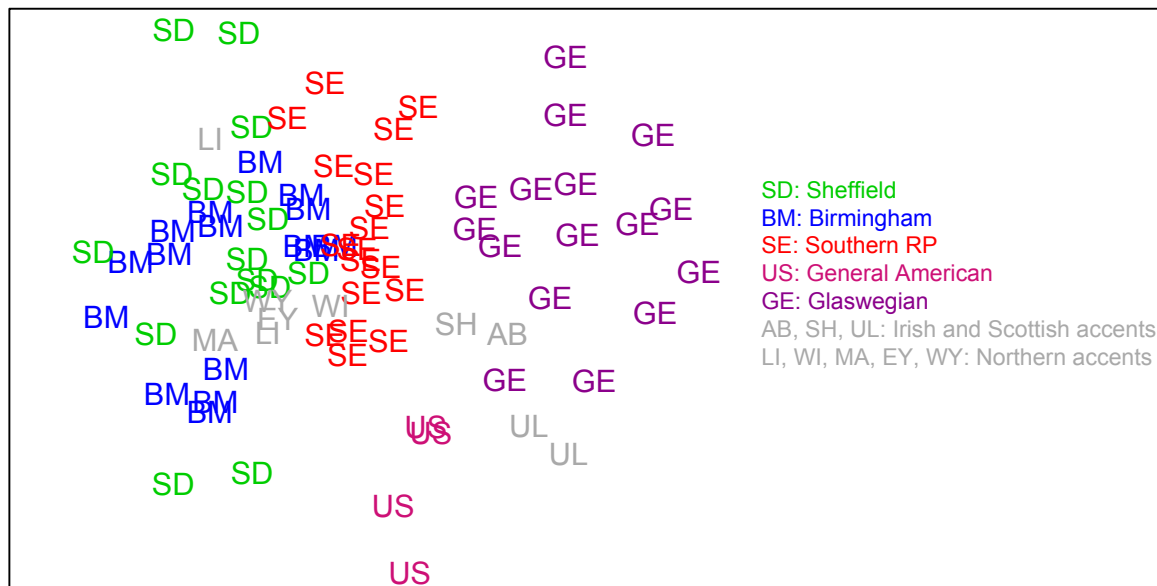
Talker-listener accent similarity affects speech intelligibility under noisy conditions. For example, native (L1) listeners are more accurate at recognizing L1 over non-native (L2) speech [e.g., 2, 9, 10]. For L2 listeners, their performance is affected by their L2 experience; low-proficiency L2 listeners can have higher intelligibility for L2-accented speech, but they begin to have advantages for L1 accents as they gain experience and their own degree of L2 accent decreases [1, 9, 10, 12]. Acoustical measurements of talker-listener accent similarity can thus be highly predictive of talker intelligibility [9].

This similarity effect can break down, however, for standard accents within an L1. For example, Adank et al. [1] found that Glaswegian English listeners have high intelligibility for southern British English speech, despite the fact that these

two accents are very different. Clopper and Bradlow [6] likewise found that standard American accents were the most intelligible to listeners who had a range of American accents. It is plausible that listeners learn to understand standard accents through familiarity and experience [1, 5], particularly through media exposure [11], such that they learn to understand an accent that is very different from their own.

Recent work has investigated this relationship between talker-listener similarity and the advantages of standard accents [7]. In this study, L1 British English listeners and speakers with a range of accents were tested on speech-in-noise recognition of sentences. Figure 1 displays a multidimensional scaling (MDS) map based on the acoustic differences between the speakers. The map separated accent groups along the horizontal dimension, with northern accents to the left, southern in the center, and Scottish/Irish accents to the right. The most highly intelligible talkers (mostly with southern British RP and US standard accents) were positioned in the center of the accent map for the listeners, with declining intelligibility for speakers further from the center. There were also talker-listener accent interaction effects (e.g., Glaswegian listeners performed better on their own accent than did southern listeners) but this wasn't observable for all groups (e.g., Sheffield listeners were more accurate on southern speech than their own accent). Even though most of the listeners had an advantage for their own accent to some extent, overall the intelligibility of individual talkers was highly correlated across listener groups, demonstrating that there were common factors influencing intelligibility.

These findings suggest that standard accents may be highly intelligible because of their position in the wider accent landscape; they might represent an average, or prototype, of all of the accent variation that a listener has heard, or their central position might reflect their influence on the speech of other varieties (i.e., other accents tending to



**Figure 1.** Multidimensional scaling map of acoustic differences among British and American talkers from a previous study [7]. Each point represents a single talker. Talkers are close together in the space when the acoustical differences are small in terms of vowel spectra and segment duration. Speakers were separated along the horizontal dimension in terms of accent (e.g., northern, southern, and Scottish/Irish left to right), whereas the vertical dimension represents more idiolectal differences.

adopt standard features). However, the consistency of the standard-accent advantage raises questions about the role of experience. That is, it is possible that these accents are inherently more intelligible, incorporating a set of robust acoustic-phonetic features that make speech perception easier for almost all listeners.

The present study investigated these possibilities by testing L2 Chinese Mandarin listeners with intermediate English proficiency on the English accent varieties of this previous study [7]. Although most groups of L2 listeners would have more familiarity with standard varieties, such as southern RP, than varieties like Glaswegian, it is also true that their L1 phonology would be expected to dominate their L2 speech perception [9]. Therefore, if the accent prototype effect found previously [7] is based on the listener's experience, then intermediate-proficiency L2 listeners ought to have a very different pattern of results than L1 listeners.

## 2. METHOD

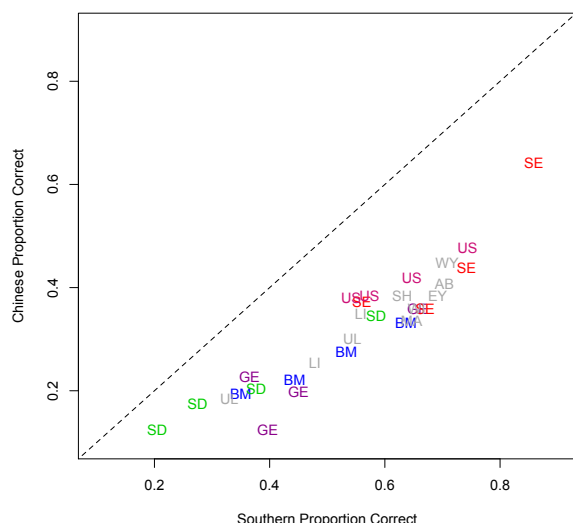
### 2.1. Subjects

The subjects were 12 native Chinese Mandarin listeners aged 22 to 24 (mean: 23 years). They

had learned English as a L2 from age 12 and all had been living in the UK for 9 months except for one subject who had spent a year in Sheffield. The Chinese listeners' results were compared to that of a group of 13 monolingual southern RP British English listeners (age range 19 to 32; mean: 25 years) tested in a previous study [7]. None of the subjects reported any speech, hearing or learning difficulties.

### 2.2. Materials

Thirty talkers of southern RP (SE), Glaswegian (GE), Sheffield (SD), Birmingham (BM), General American (US) and other miscellaneous accents (Scottish and Irish accents: Aberdeen 'AB', Scottish Highlands 'SH', Ulster 'UL'; UK Northern accents: Liverpool 'LI', Wigan 'WI', Manchester 'MA', East Yorkshire 'EY' and West Yorkshire 'WY') were recorded reading the full set of the BEL sentences (e.g., *The park opens in eleven months* [4]). The digitized recordings were embedded in speech-shaped noise with a signal-to-noise ratio of -3dB. The speech-shaped noise was generated for individual talkers such that it matched the smoothed long-term average spectrum of their speech.



**Figure 2.** Proportion correct words in sentences mixed with noise, for Chinese and southern RP British English listeners. Each point represents a single talker (see Figure 1).

### 2.3. 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 logged how many keywords were spoken correctly). They listened to all 450 sentences and were given regular breaks. The stimuli were presented in a random order and the pairings of sentences and accents were counterbalanced between subjects (i.e., each subject heard each sentence only once, but the sentences were spoken in all accents across the experiment).

## 3. RESULTS

Figure 2 displays a scatterplot of the mean proportion correct for each individual talker, comparing Chinese and English listeners. Unsurprisingly, the Chinese listeners had lower speech recognition levels than the English listeners (i.e., the points of all talkers are shifted to the right of the equal-intelligibility line). However, individual differences in talker intelligibility were highly correlated between the two listener groups ( $r = 0.93, p < 0.001$ ). Despite their worse recognition of English speech due to their intermediate L2 proficiency and their different L1 phonological system, the Chinese

listeners had high and low intelligibility for the same talkers as the English listeners.

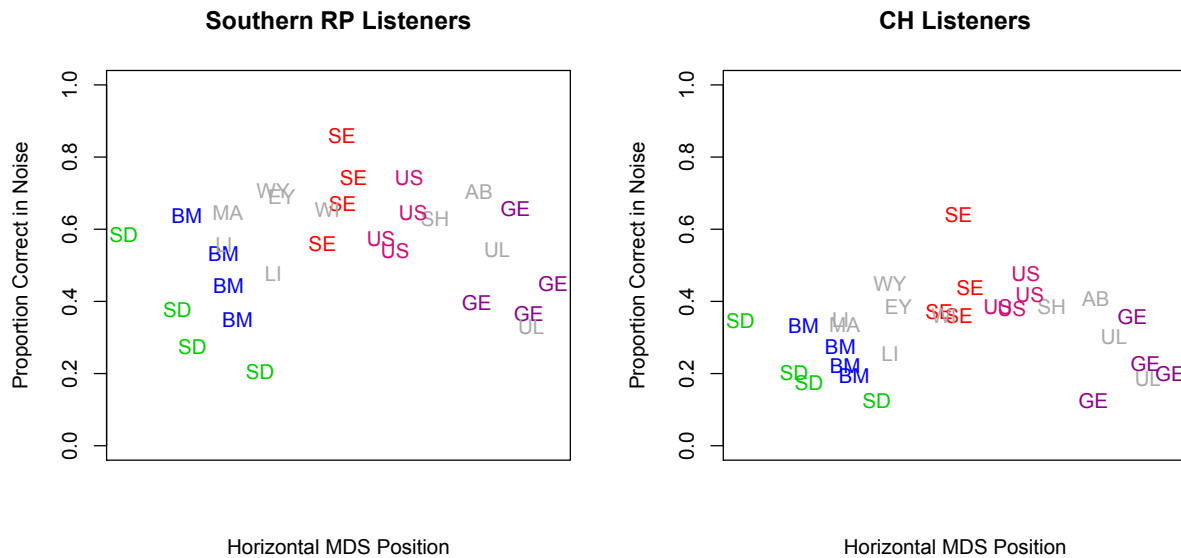
Figure 3 displays the mean proportion correct for each talker vs. the horizontal position of the talkers in the MDS solution (from Figure 1). For both listener groups, talkers who were closest to the middle on the horizontal axis of the MDS solution were most intelligible (English:  $r = -0.49, p < 0.01$ ; Chinese:  $r = -0.57, p < 0.001$ ). That is, both groups found the standard accents to be most intelligible (SE and US), and had declining intelligibility for talkers that were further from this central tendency in the accent landscape.

## 4. DISCUSSION

The results demonstrate that Mandarin listeners are surprisingly similar to L1 southern RP listeners in terms of which speakers and accents they find more intelligible, despite their very different phonological system and overall lower proficiency. Speakers with standard or prestige accents (SE and US) are more intelligible.

Why are standard accents highly intelligible to L1 and L2 listeners? It is unlikely that it just an effect of accent familiarity or experience. The Chinese listeners in the present study were indeed more familiar with standard English accents than they were with regional varieties, but their Mandarin phonological system dominated their use of English (i.e., they had low speech recognition scores and spoke with strong Chinese accents). Likewise, previous work has demonstrated that L1 speakers of regional varieties find standard accents to be more intelligible than their own [6, 7], and it is certain that these individuals hear accents through the media less frequently than they hear their own voice and the accents of people speaking around them.

It is more likely that standard accents have an inherent intelligibility for all listeners, but the exact phonological realizations that make standard accents so intelligible are yet to be examined. The MDS solution displayed here (Figure 1), was based on very broad measures of vowel formant differences and segmental duration. Different acoustic measures tend to be highly correlated (i.e., if two speakers tend to use similar vowel formant patterns, they also tend to have similar patterns of segmental duration variation), making



**Figure 3.** Proportion of keywords correct vs. horizontal position of the talkers in the MDS solution for southern and Chinese listeners.

it impossible to identify a well-defined set of acoustic features that cause a speaker to be located near or far from a central tendency in the MDS space. Further research with synthesized speech might be able to solve this issue by controlling these confounds, but it is difficult to fully capture natural variation with these carefully controlled materials.

If standard accents adopt inherently intelligible acoustic-phonetic features, then this provides a new view of sociophonetic variation. That is, it might be that certain points in the accent landscape are inherently “high ground” in terms of intelligibility, and that standard accents tend to develop to occupy this territory. That is, socially dominant accents might specifically acquire, over time, acoustic features that promote intelligibility, rather than these accents becoming more intelligible by remapping the perception of individual listeners through exposure.

## 5. ACKNOWLEDGEMENTS

The project research was supported by the ESRC.

## 6. REFERENCES

- [1] Adank, P., Evans, B., G., Stuart-Smith, J. and Scott, S., K. (2009). “Comprehension of familiar and unfamiliar native accents under adverse listening conditions”, *Journal of Experimental Psychology: Human Perception and Performance* 35, 520-529.
- [2] Bent, T., and Bradlow, A. R. (2003). “The interlanguage speech intelligibility benefit”, *Journal of the Acoustical Society of America* 114, 1600-1610.
- [3] Bradlow, A.R. and Bent, T. (2007). “Perceptual adaptation to non-native speech”, *Cognition* 106, 707-729.
- [4] Calandruccio, L. and Smiljanic, R. (2012). “New sentence-recognition materials developed using a basic nonnative English lexicon”, *Journal of Speech, Language and Hearing Research*, 55, 1-34.
- [5] Clarke, C. M. and Garrett, M. F. (2004). “Rapid adaptation to foreign-accented English”, *Journal of the Acoustical Society of America* 116, 3647-3658.
- [6] Clopper, C. G. and Bradlow, A. R. (2008). “Perception of Dialect Variation in Noise: Intelligibility and Classification”, *Language and Speech* 51, 175-198.
- [7] Iverson, P., Pinet, M., and Evans, B. G. (2014). “Mapping accent similarity and speech in noise intelligibility for British English accents”, *Acoustical Society of America* 135, 2422.
- [8] Pinet, M. and Iverson, P. (2010). “Talker-listener accent interactions in speech-in-noise recognition: Effects of prosodic manipulation as a function of language experience”, *Journal of the Acoustical Society of America* 128, 1357-1365.
- [9] Pinet, M., Iverson, P., and Huckvale, M. (2011). “Second language experience and speech-in-noise recognition: Effects of talker-listener accent similarity”, *Journal of the Acoustical Society of America*, 130, 1653-1662.
- [10] Stibbard, R. M. and Lee, J. I. (2006). “Evidence against the mismatched interlanguage speech intelligibility benefit hypothesis”, *Journal of the Acoustical Society of America* 120, 433-442.

[11] Stuart-Smith, J. (2007). "The influence of the media", in: Llamas, C., Mullany, L. and Stockwell, P. (eds.) *The Routledge Companion to Sociolinguistics*. Routledge, London, UK, 140-148.

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