Phonological learning based on interactive and mediated speech

Jane Stuart-Smith, Rachel Smith and Sophie Holmes

Department of English Language, University of Glasgow; j.stuart-smith@englangarts.gla.ac.uk; r.smith@englangarts.gla.ac.uk; s.holmes@englangarts.gla.ac.uk

Introduction

Current models of language change assume that face-to-face interaction is essential for the kind of phonological learning which leads to change, and reject the possibility that such learning occurs from non-interactive speech sources such as radio and television. This assumption is challenged by recent findings showing that the broadcast media are involved in certain phonological changes. But almost no research directly investigates whether these different experiences of language cause differences in phonological learning. We report the first steps of a research programme to compare how listeners learn about an accent other than their own through face-to-face interaction versus watching a video of the same.

Background

Listening to speech from speakers in the broadcast media, who may have accents different from our own, is now an everyday occurrence for most individuals. Yet very little is known about how processes of speech perception and phonological learning differ when listening non–interactively to speech, as compared to when we participate in live interaction with others. Indeed, there are differing assumptions about the potential impact of the broadcast media on speech perception. Hay et al. (2006), working on a vowel merger in NZ English, argue for the availability of exemplars from previous media exposure, whilst Evans and Iverson (2004), who looked at the perceptual categorization of vowels in speakers of British English, with different opportunities for face-to-face interaction as a result of where they lived, seem to come to the opposite conclusion.

There are good reasons to think that interaction might give rise to special kinds of learning about the accents of others. Interactive speech may engage attention more fully, and be more memorable, than mediated speech, because it is often perceptually richer (e.g. providing three-dimensional visual as well as auditory information), and because it typically involves the collaborative pursuit of goals. Finally, interactional partners appear to align with or entrain to each other’s lexical and syntactic structures, phonetic patterns, breathing patterns and speech rhythms (e.g. Pickering and Garrod, 2004).

At the same time, there is evidence from sociolinguistics that people can learn about mediated speech. Work with adolescents in Glasgow, Scotland (e.g. Stuart-Smith, 2005) shows correlations between engagement with a London-based television soap, and production of phonetic features characteristic of London. Moreover, experiments on speech perception that use non-interactive recorded speech (as most do) show that speakers can carry out many kinds of perceptual learning about such speech (e.g. Norris et al., 2003). Intriguingly, recent work testing infants’ perception of a foreign-language phonological contrast suggests that live interaction with a speaker of the language gives rise to learning where equivalent exposure via video fails to (Kuhl et al., 2003). But there is no comparable research with adolescents or adults.

Experimental design and methodology

The experiment assessed how interactive vs. mediated exposure affects the way speakers of Scottish English learn about a distinction that is non-phonemic for them, but phonemic in Standard Southern British English (SSBE). The SSBE distinction investigated is between /a/ and /ɑː/, which occur in four lexical sets (Wells, 1982): TRAP (/a/), BATH, PALM and START (all /ɑː/): certain codas occur only with one vowel (e.g. /ŋ/ with /a/: hang, bank), while other codas can occur with either vowel (e.g. /nt/: ant, aunt). In Scottish English, words in all four sets contain /a/ [a]. The experiment had 40 participants, and three phases:

1) a rhyme-judgement test, to establish baseline knowledge. It used 48 monosyllables (selected from the TRAP, BATH and PALM lexical sets) and 52 distractors. Participants judged which vowel an SSBE speaker would use in each word, categorizing it as having the vowel of cat, bath, or a
different vowel. The design grouped the experimental words equally into “more predictable”/ “less predictable,” according to the extent to which /a/ or /ɑː/ might be predicted from their coda.

2) a period of either interactive or mediated exposure to SSBE. In the interactive condition, 20 participants played a word game with the SSBE-speaking experimenter, which required the experimenter to mention half of the experimental words. Each interaction was video-recorded and formed the exposure material for one of 20 participants in the mediated condition, who watched it and kept a tally of the scores as the players played the game.

3) a different randomization of the rhyme-judgement test, to examine change in participants’ knowledge of the SSBE system.

The data were analysed in R using generalized linear mixed-effects modeling. Models were fitted for two dependent variables, correct responses, and consistency of responses across the rhyme-judgement tests.

**Results**

We expected all participants to improve at the task, and those in the interactive condition to improve more. Neither prediction was straightforwardly supported by the results, but a range of differences between the interactive and mediated conditions emerged when lexical set was taken into account. Overall, the proportion of correct responses did not increase from the first to second rhyme-judgement test, and decreased for BATH words. There was no difference in the extent to which interactive and mediated participants improved at the task; but interactive participants did make more correct responses to BATH and PALM words than mediated participants, and marginally fewer correct responses to TRAP words; there was a trend for these patterns to become more pronounced from the first rhyme-judgement test to the second.

Interactive and mediated participants also differed marginally significantly in terms of the consistency of their responses, with interactive participants more consistent than mediated participants. Responses to TRAP words were much more consistent overall than those to BATH, and responses to words not heard in the game were marginally more consistent than those to words actually heard.

**Discussion**

For our Scottish speakers, the hardest words to categorise were those containing /ɑː/ in SSBE (i.e. BATH, PALM). Participants who received mediated exposure performed particularly poorly on these words, and also showed less consistency in responding across the two tests than those in the interactive condition. A possible interpretation of these findings is that mediated participants underwent more disruption to their existing knowledge about the distribution of the sounds than interactive participants did, and that this disruption represents the initial stages of learning, which is made possible by the reduced cognitive load of watching a game as opposed to participating in it. Broader issues arising from the results include why participants failed to improve at the task overall, and why there was no difference in the interactive and mediated groups’ improvement at the task. These will be addressed in ongoing refinements to our developing methodology.

**References**


