3aSCb3 The Influence of Socioindexical Expectations on Speech Perception in Noise

Kevin B. McGowan <kmcgowan@rice.edu>
Department of Linguistics, Rice University

Abstract

Previous research has shown that listener percepts can be altered by the manipulation of listener expectations of speaker identity [e.g., Niedzielski (1999), Hay et al. (2006)]. This study examines the extent to which socioindexical information can be useful to listeners. English speakers with little or no Chinese experience and Heritage Mandarin English speakers transcribed Chinese-accented speech in noise. In a between-subjects design, listener expectations about speaker identity were manipulated by presenting an Asian face, a socioindexically uninformative silhouette, or a Caucasian face as the purported speaker of high and low predictability sentences. Consistent with previous findings, listeners presented with an Asian face were significantly more accurate transcribers than those presented with a Caucasian face. Intriguingly, the silhouette condition patterned with the Asian face for experienced listeners but with the Caucasian face for inexperienced listeners. Unexpectedly, inexperienced listeners, while overall less accurate than experienced, Heritage listeners, saw the same magnitude improvement with the Asian face. Furthermore, transcription errors are inconsistent with suggestions (e.g., Staum Casasanto, 2009; Johnson, 2006) that listeners will alter base activations of prelexical or lexical forms to accommodate an expected accent. Implications for theories of speech perception and word recognition will be explored.

Background

'Socioindexical' refers to aspects of the speech stream or context that convey information about social identity: age, gender, sexual orientation, class, regional origin, race. A number of recent experiments have shown that manipulating socioindexical cues can shift listeners' mapping of acoustic input onto mental representation. A few examples:

Rubin (1992) Asian face versus Caucasian face alters perceived ethnicity, accent and listener performance on a word-recollection task. (reviewed in Lippi-Green 1997).

Niedzielski (1999) Detroiters label examples of an accent like their own more accurately when told the speaker was Canadian than when told the same voice was a fellow Detroiter.

Strand (2000) Listeners are able to repeat words more quickly when the speaker's voice is more stereotypically male or female.

Staum Casasanto (2009) Listeners choose sentence endings consistent with AAE features more often and more quickly if they believe the speaker to be African American.

Theoretical Goals

Johnson (2006) makes the prediction that socioindexical cues, like those in Strand (2000), will pre-activate relevant exemplars –raising their base activation levels and making these exemplars more prototypical, more available than they normally would be.

This suggests that socioindexical cues should be able to **enhance** perception of accented speech in noise. It also implicitly assumes that this effect will only hold for experienced listeners (listeners with relevant exemplars to pre-activate).

Predictions

1. Manipulating socioindexical beliefs will enhance perception of Chinese-accented speech in noise.
2. Inexperienced listeners will show less improvement than experienced listeners.

Method

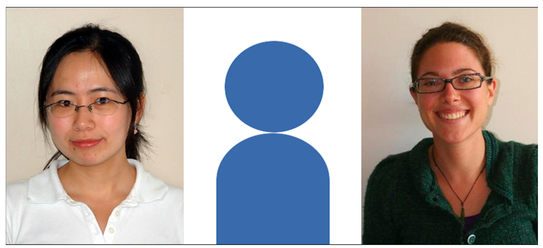
Matched Guise Task

Listen: to sentences embedded in -4dB SNR multi-talker babble. Each sentence was heard once and could not be repeated.

Transcribe: listeners were asked to type what they heard as accurately as possible.

Visual Stimulus

Listeners were randomly assigned to one of three face conditions prior to their arrival in the lab. Each participant saw only one face and was led to believe that this was the face of the talker.



Asian face Silhouette Caucasian face

Aural Stimuli

Listeners heard a series of high and low predictability sentences (from Bradlow and Alexander 2007). Examples:

High Predictability	Low Predictability
Elephants are big animals.	He pointed at the animals.
A pigeon is a kind of bird.	We pointed at the bird.
The war plane dropped a bomb.	Dad talked about the bomb.

etc...

Participants

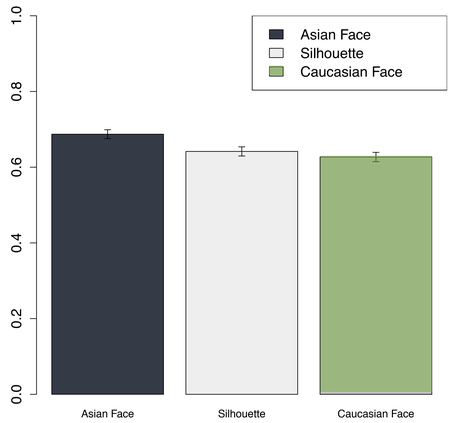
57 self-reported inexperienced listeners from the University of Michigan Psychology subject pool

30 self-reported Heritage Mandarin speakers with little or no speaking ability (*childhood overhearers*, Knightly et al. 2003) from the University of California, Berkeley

After participating in this task, all listeners performed a yes/no authentic Chinese identification task that, in general, supported listeners' self-reported experience levels.

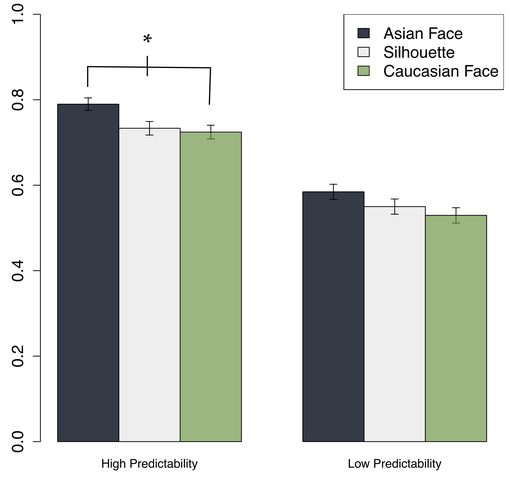
A response was coded as 'correct' if it contained the underlined target word.

Results

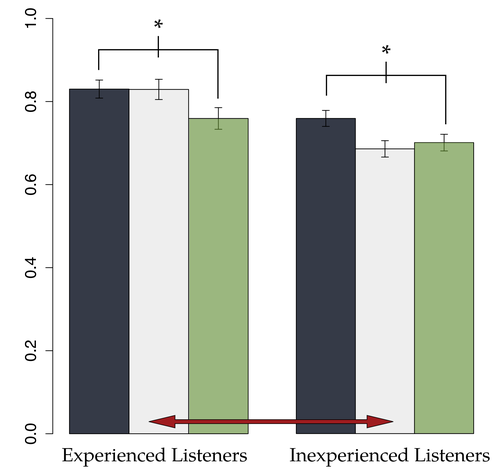


Proportion correct.

	Coef β	SE(β)	z	p
(Intercept)	2.06	0.28	7.3	**<.001**
Silhouette	−0.24	0.13	−1.9	0.0638
Caucasian Face	−0.35	0.13	−2.7	**.0064**
Low Predictability	−1.17	0.07	−15.7	**<.001**
experienced	−0.58	0.11	−5.2	**<.001**



Proportion correct by predictability



High Predictability: Experienced vs. Inexperienced Listeners

Discussion

Experienced and Inexperienced listeners alike transcribe Chinese-accented sentences (high predictability sentences) in noise **more accurately when shown an Asian face** than when shown a Caucasian face.

Experienced listeners are more accurate overall, but the magnitude difference between the Asian and Caucasian face conditions seems to be the same. Performance on the Silhouette Face condition may indicate listeners' default assumptions about speaker identity.

Conclusions

There is clear evidence that experienced listeners are, overall, more accurate than inexperienced listeners when transcribing Chinese-accented speech. This finding is consistent with an exemplar theoretic approach to phonology and the lexicon.

However, these results do not support a model in which social knowledge pre-activates appropriate exemplars. Experienced and inexperienced listeners were equally assisted by the presentation of an Asian face (though, of course, not necessarily via the same mechanism).

Attribution of other socioindexical perceptual effects in the literature to exemplar activation may be premature.

References

Johnson, K. Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. Journal of Phonetics, 34:485–499, 2006.

Knightly L.M., Jun, S., Oh, J.S., and Au, T.K. Production benefits of childhood overhearing. J. Acoust. Soc. Am. Volume 114, Issue 1, pp. 465-474 2003.

Niedzielski, N. The effect of social information on the perception of sociolinguistic variables. Journal of Language and Social Psychology, 18(1):62–85, 1999.

Rubin, D. L. Nonlanguage factors affecting undergraduates' judgments of nonnative english-speaking teaching assistants. Research in Higher Education, 33(4):511–531, 1992.

Staum Casasanto, L. Experimental Investigations of Sociolinguistic Knowledge. Ph.D. thesis, Stanford University Department of Linguistics, 2009.

Strand, E. A. Gender Stereotype Effects in Speech Processing. Ph.D. thesis, The Ohio State University, 2000.