Individual differences in accent adaptation

Xin Xie & T. Florian Jaeger (University of Rochester)

Exposure to a talker with atypical pronunciations changes how this talker is perceived subsequently, often improving comprehension of that talker [1-2]. One line of research focuses on how listeners make adjustments when exposed to isolated words with atypical pronunciations of a particular contrast (e.g., /s/-/sh/; [3-4]). In these studies, as few as 10 words containing the sound can elicit significant adaptation for that contrast; post-exposure tests often involve a 2 Alternative Forced Choice (2AFC) task on minimal pairs and as such, listeners' attention is explicitly directed towards a particular sound/contrast. In contrast, research on the perception of globally accented L2 speech sometimes assumes that significantly more exposure is required for successful learning (but see [5-6]). These studies typically use transcription tasks to probe comprehension enhancements [1,7]. It is thus possible that adaptation proceeds equally rapidly in both paradigms but is left undetected in the transcription task. We report two experiments that directly compare listeners' adaptation to an unfamiliar L2 accent when assessed by a 2AFC or a transcription task. In addition, we ask whether the particular adaptation pattern is dependent on talker-specific properties by examining two test talkers (late L2 learners of English with intermediate intelligibility).

Methods. In two MTurk-based experiments (N = 47, 56), we assessed perception of word-final stop voicing in Mandarin-accented English, among native English listeners. Both experiments use the same stimuli and design (Fig.1), differing only in the task (2AFC vs. transcription). Between participants, both experiments manipulated whether exposure presented isolated spoken words from the L2-accented test talker (L2-accented exposure) or the same words from an L1 speaker (L1-accented exposure). This manipulation was crossed with the L2-accented test talker: half of the participants in each exposure condition heard test talker M4 and half heard M15. On each exposure trial, participants heard a word and had to choose which of two words on the screen they heard (2AFC) or had to transcribe the word (transcription). Either way, participants received immediate feedback about the correct response after each exposure trial. At test, participants completed the same task as during exposure but without feedback and on a new set of words. Critical trials (<50% in both exposure and test) involved minimal pairs with a word-final stop (e.g., 'seed' vs. 'seat'). Response accuracy was measured; transcriptions were considered as accurate if the voicing (voiced vs. voiceless) was correctly recognized.

Results and discussion. Data from each experiment was analyzed with logistic mixed-effects regression (accuracy ~ Test Talker * Condition * Voicing + maximal converging random effect structure; see Table 1). For the transcription task, our results showed a *Test Talker* effect (M15 > M4), a *Test Talker X Voicing* interaction, and critically, a three-way *Test Talker X Condition X Voicing* interaction. The three-way interaction was driven by a significant *Condition X Voicing* interaction for both talkers but in opposite directions: for M15, the experimental group had reduced accuracy for voiced tokens and increased accuracy for voiceless tokens, indicating a bias shifted towards voiceless tokens; for M4, there was an opposite bias shift towards voiced tokens. Further simple effect analyses were shown in Fig.2A. A similar pattern was observed for the 2AFC task, although the effects were overall smaller (Fig.2B). Pulling data from both experiments, there was an overall three-way interaction as observed separately for each task, with no *Test Talker X Condition X Voicing X Task* interaction. Taken together, the effects of accent exposure (~10 mins of exposure; replicating prior work using sentence stimuli [3-4]) were highly consistent across tasks but exhibited strikingly distinct patterns for two test talkers of the same L2 accent. Our finding offers two critical insights for future work on L2 speech perception. First, researchers should consider examining learning effects even within a short paradigm. Second and more importantly, the large by-talker differences in adaptation reveals an underestimated role of talker variability, even among talkers who are assumed to be extremely similar. Therefore, research on accent adaptation—in particular, work focusing on cross-talker generalization—should be cautious drawing conclusions from just one test talker (cf. [1,2]).

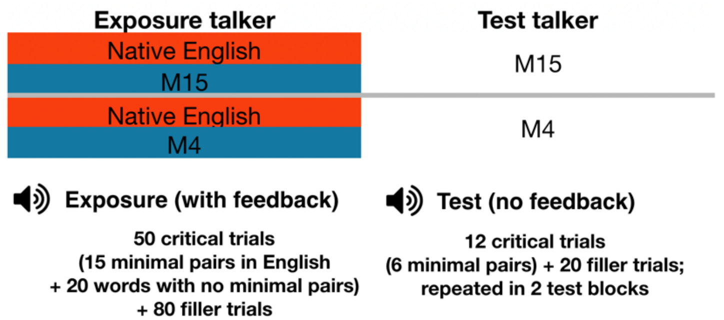

Figure 1.
Experimental design. In each experiment, two test talkers (M15 and M4) were employed. For each test talker, there were two exposure talker conditions (Native English. Vs. M15).

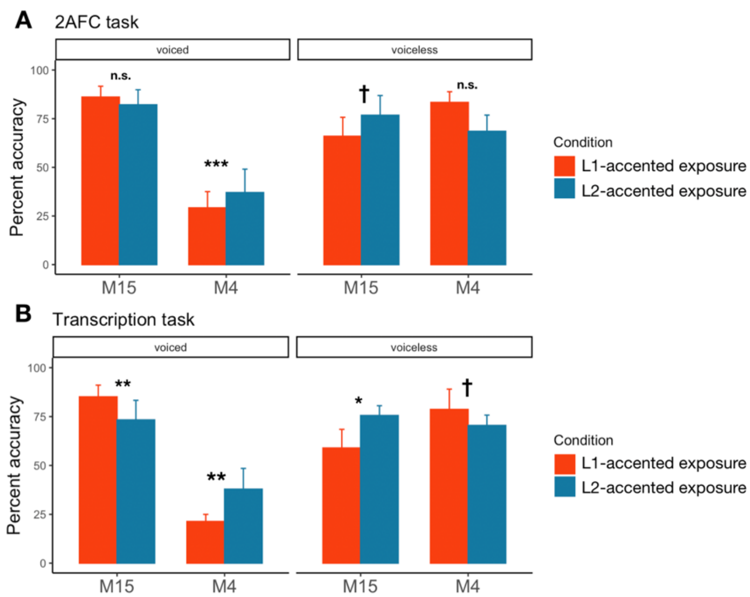

Figure 2.
Human responses. The x-axis shows performance for the two L2-accented test talkers.

** represents p < 0.05, ** represents* p < 0.01, *** represents p < .0001 and † represents p < 0.1.

Table 1. Mixed-effects model results.

** represents p < 0.05, ** represents* p < 0.01, *** represents p < .0001 and † represents p < 0.1.

2AFC Task

Levels	Coefficient $\hat{\beta}$	SE	z	p	
Intercept	-0.591	0.204	-2.897	0.004	**
Talker	1.116	0.145	7.695	0.000	***
Condition	0.376	0.153	2.451	0.014	*
Voicing	-0.795	0.181	-4.390	0.000	***
Talker X Condition	-0.351	0.145	-2.429	0.015	*
Talker X Voicing	1.122	0.110	10.168	< 2.00E-16	***
Condition X Voicing	0.247	0.121	2.043	0.041	*
Talker X Condition X Voicing	-0.525	0.110	-4.779	0.000	***

Transcription Task

Levels	Coefficient $\hat{\beta}$	SE	z	p	
Intercept	0.723	0.271	2.674	0.008	**
Talker	0.715	0.112	6.396	< 1.60E-10	***
Condition	0.011	0.125	0.091	0.928	
Voicing	-0.409	0.260	-1.575	0.115	
Talker X Condition	-0.164	0.111	-1.477	0.140	
Talker X Voicing	0.938	0.084	11.158	< 2.00E-16	***
Condition X Voicing	-0.013	0.100	-0.128	0.898	
Talker X Condition X Voicing	-0.539	0.083	-6.481	< 9.09E-11	***

[1] Bradlow & Bent, 2008 *Cognition*. [2] Xie et al., 2020 *JEPG*. [3] Norris et al., 2003 *Cognitive Psychology*. [4] Kraljic & Samuel, 2005 *Cognitive Psychology*. [5] Clarke & Garrett, 2004 *JASA*. [6] Xie et al., 2018 *JASA*. [7] Sidaras et al., 2009 *JASA*.