

Unpredictable speakers: Evidence that social category labels mediate the role of experience

Abby Walker, Carla Fernandez and Janet van Hell

According to usage-based models, listeners track when a speech signal and a context typically co-occur, and linguistic processing is easier when the incoming signal is congruent with contextual cues, and more difficult with the incoming signal is incongruent with such cues (e.g., Babel & Russell 2015). For signals that don't have a strong contextual bias because they occur equally often in multiple contexts, such models would predict that processing ease should fall between a congruent and incongruent context-signal pairing, which is borne out when looking at word and social-group pairings (Walker & Hay 2011). In this paper we test the theory by looking at dialect and speaker pairings: is lexical access facilitated for known speakers with predictable dialects when using the predicted dialect, inhibited for the same speakers when using a different dialect, and do responses to known but dialectally-unpredictable speakers fall in between?

Stimuli were produced by six actresses, who were audio- and video-recorded producing two monologues about their character's lives as well as 300 real and 120 nonsense monosyllabic words. They recorded the materials in two guises: with a Standardized US accent, and with a Southern US accent. 146 US-English speaking participants did a speeded lexical decision task (64 had lived in a Southern region prior to college). In the *Video-Monologue* condition, they first watched and listened to the monologues (familiarization), where two of the actors were presented as Southern, two were presented as Standard, and two were presented as Unpredictable, switching between Southern and Standard (Fig. 1). For the lexical decision task, 50% of words from Unpredictable speakers were from their Southern guise, and 50% were from their Standard guise. For Southern and Standard speakers, 25% of real words were incongruent with the speaker's predictable dialect, allowing us to investigate the effect of dialect incongruencies while maintaining the association between a speaker and a given dialect.

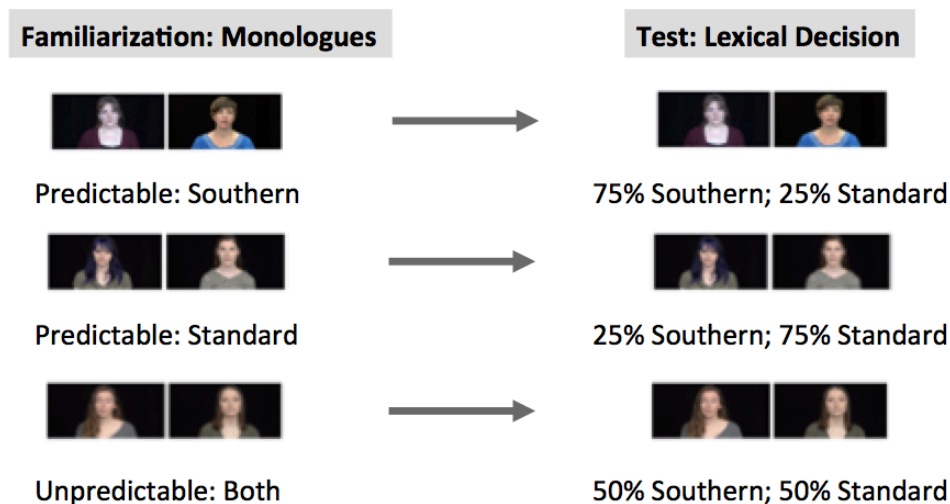


Fig 1. Experimental design for the Video-Monologue condition. Other conditions differed in the absence of video and/or the familiarization stage. Speaker-dialect combos differed across lists.

There were three additional conditions, added to assess the importance of the familiarization stage and of video: a *Video* condition without familiarization, an *Audio-Monologue* condition with familiarization but no video, and an *Audio* condition, with neither familiarization nor video. At the end of all conditions, participants completed the same post-task

questionnaires about dialect attitudes, identities and exposure. Based on responses, all participants were given a Southern Accent Score (among other scores), reflecting their self-reported accentedness.

In both *Audio* conditions, listeners made more mistakes and were slower with Southern compared to Standard tokens regardless of their regional background or Southern Accent Score (matching earlier work using authentic Southern stimuli (Walker et al. 2018)). The *Video-Monologue* condition (Fig 2) yielded an effect of speaker identity for listeners with high Southern Accent scores: they were significantly slower for Southern and especially Standard tokens when these were incongruent with a predictable speaker’s accent, but they did worse still with tokens from Unpredictable speakers. Listeners with lower Southern Accent scores were faster with Standard tokens regardless of congruency, but even these listeners were significantly slower with tokens from Unpredictable speakers. In the *Video* condition, there was no effect of (in)congruency, but overall listeners were significantly slower with Unpredictable speakers.

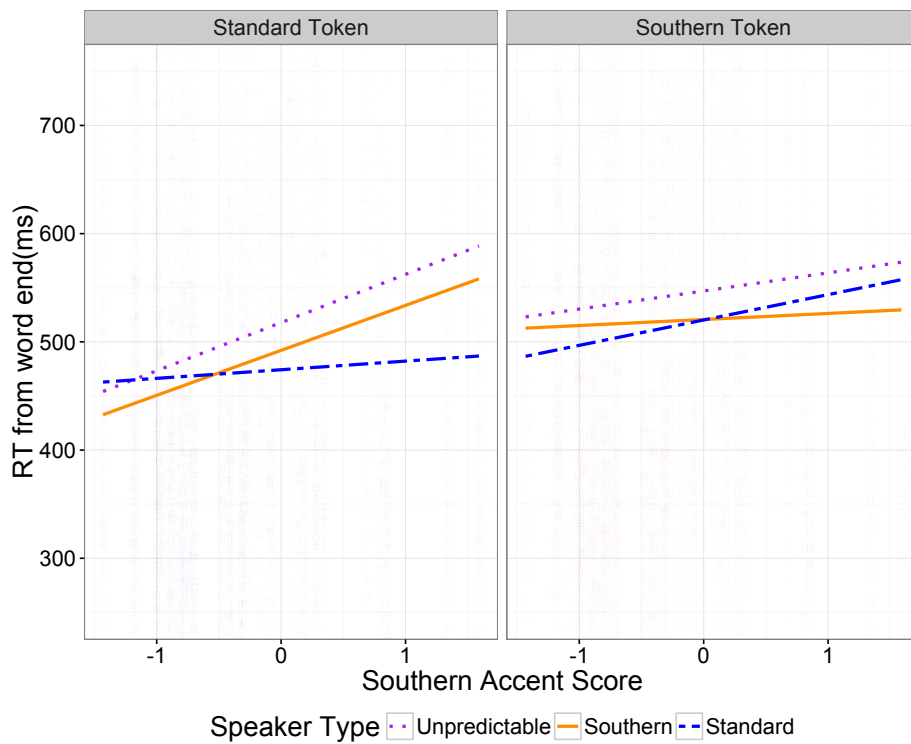


Fig 2. Response times (ms) by token dialect, speaker type, and listener’s Southern Accent Score, in the Video-Monologue condition.

The finding that listeners with higher Southern Accent scores make stronger use of speaker identity in dialect processing mirrors findings that early bilinguals use speaker identity more than monolinguals/late bilinguals (Molnar et al. 2015; Fecher & Johnson 2018). Critically, even though listeners get equal exposure to both dialects from Unpredictable speakers, and very little exposure to the second dialect from Predictable speakers, they do worse with Unpredictable speakers than they do with incongruencies from a Predictable listener. This suggests that the ability to categorize a speaker’s dialect is more useful than hearing variation from that speaker (cf. Martin et al. 2016), and highlights how the impact of experience on speech perception is mediated by (social) category labels.