# Can our eyes trick our ears? An electrophysiological study of the effects of facial cues on the processing foreign-accented and native-accented speech

Carla B. Fernandez & Janet G. van Hell
The Pennsylvania State University
Cbf140@psu.edu

In the face of growing ethnic diversity, we often form an impression of people based on how they sound or look, although this tendency may be affected by experiences with foreign accents and with speakers of different ethnicities[1]. Reverse linguistic stereotyping (RLS) posits that listeners ascribe stereotype characteristics to speech based on information about the speaker's social identity, even before hearing the speech[1]. This study examined the RLS account, and extends it to online sentence processing. Specifically, we examined how faces cuing the speaker's ethnicity (e.g., Asian face) create language expectations (here, Chinese-accented English), and how these biases impact the neural and cognitive mechanisms associated with the comprehension of American-or Chinese-accented English using Event-Related Potentials (ERPs).

Previous studies have found that foreign-accented speech leads to an increased difficulty in sentence comprehension than can affect semantic processing, but has particularly strong effects on grammatical processing[3,4,5,6,7]. However, these studies presented only sentences, which limits the generalizability to more naturalistic situations. Building on the RLS account, we hypothesized that adding a face that cues the speaker's identity and expectations about the speaker's accent impacts sentence processing. Indeed, related sociolinguistic studies have found that language expectations based on a speaker's ethnicity can lead to negative social bias and impair global language comprehension[8,9]. However, there are no studies examining online language processing focusing on the processing of semantic and syntactic information during sentence comprehension.

Building on Grey and Van Hell (2017), 20 monolingual listeners were presented with 480 English sentences: 120 sentences contained a semantic, 120 sentences were the correct equivalents, 120 sentences contained a pronoun error (pronoun errors are common in Chinese-accented English speakers), and 120 were the correct equivalents. 240 Sentences were paired with a congruent picture matching the accent (Chinese face-Chinese accent and Caucasian face-American accent) and 240 were paired with an incongruent picture not matching the accent (Chinese face-American accent and Caucasian face-Chinese accent). The pictures were of the same 4 speakers (two Caucasian and two Chinese speakers) who recorded the sentences and the matching remained constant throughout the experiment. ERP analyses focused on time windows related to the N400 (300-500ms) indexing semantic processing, and the Nref (270–900ms) indexing referential processing and searching for an antecedent based on the pronoun.

Analyses were conducted on mean amplitude values using two repeated measures ANOVAs: a midline ANOVA and a lateral one that included a factor of anteriority and laterality. Results showed that semantic violations elicited typical N400 effects in the congruent Caucasian face/American-accented English condition and an attenuated N400 in Asian face/Chinese-accented English (see Fig 1). In the incongruent conditions, there was an N400 effect in the Asian face/American-accented English condition, but no effect in the incongruent Caucasian face/Chinese-accented English condition (see Fig 2). Pronoun violations elicited an Nref in the two congruent conditions (Caucasian face/American-accented English and Asian face/Chinese-accented English, see Fig 3). In the incongruent conditions, there was a negativity in the Asian face/American-accented English condition (270-900ms), and no effect in the American-face-Chinese accented English condition (see Fig 4). These findings indicate that facial cues play an important role in activating certain stereotypes regarding a speaker (in line with the reverse linguistic stereotyping account[1]) and can affect the processing of both native and foreign-accented speech, depending on whether or not facial cues match listeners' expectations of a speaker's accent. Taken together, these findings provide a more in-depth understanding of the role that visual cues play in language processing.
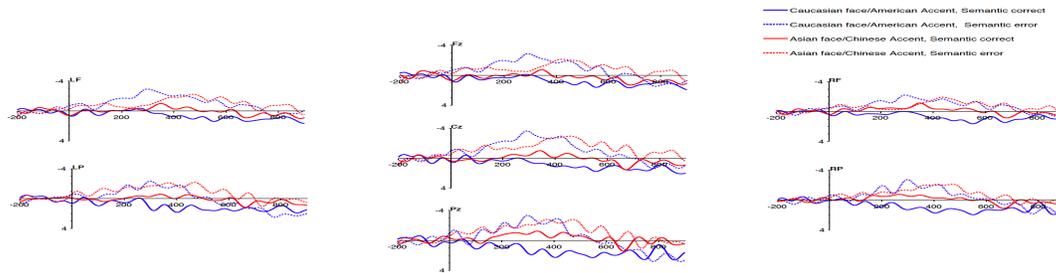
*Figure 1. Grand mean waveforms for semantic incorrect critical words (dashed lines) versus correct equivalents (solid lines) in both congruent conditions (Caucasian face/American accent in blue and Asian face/Chinese accent in red). Onset of the critical word is indicated by the vertical bar (negative is plotted up). Time is plotted on the x-axis; each tick mark indicates 200 ms. LF=left frontal; RF=Right frontal ;LP=left posterior; RP=right posterior.*
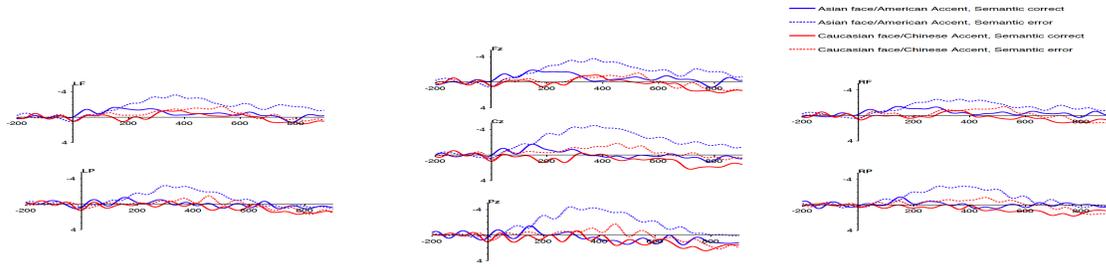


*Figure 2. Grand mean waveforms for semantic incorrect critical words (dashed lines) versus correct equivalents (solid lines) in both incongruent conditions (Caucasian face/Chinese accent in blue and Asian face/American accent in red). Onset of the critical word is indicated by the vertical bar (negative is plotted up). Time is plotted on the x-axis; each tick mark indicates 200 ms. LF=left frontal; RF=Right frontal; LP=left posterior; RP=right posterior.*
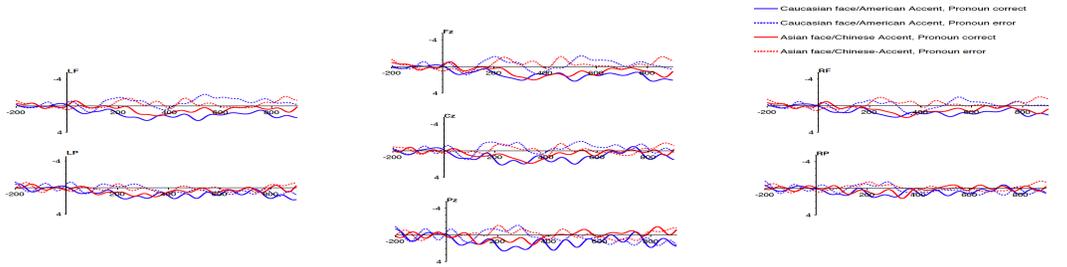


*Figure 3. Grand mean waveforms for pronoun incorrect critical words (solid lines) versus correct equivalents (dashed lines) in both congruent conditions (Caucasian face/American accent in blue and Asian face/Chinese accent in red) Onset of the critical word is indicated by the vertical bar (negative is plotted up).  Time is plotted on the x-axis; each tick mark indicates 200 ms. LF=left frontal; RF=Right frontal; LP=left posterior; RP=right posterior.*
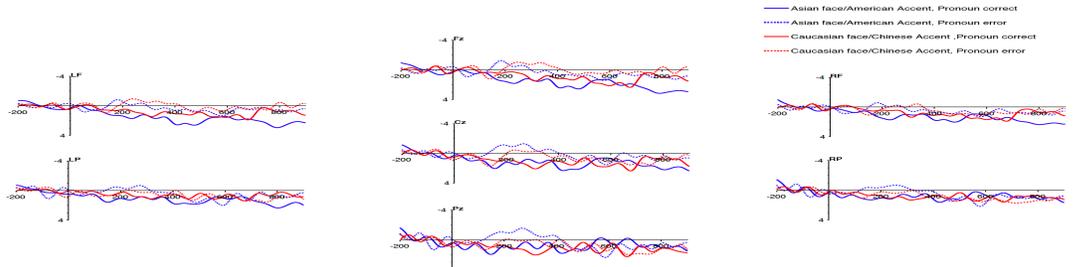


*Figure 4. Grand mean waveforms for pronoun incorrect critical words (dashed lines) versus correct equivalents (solid lines) in both incongruent conditions (Caucasian face/Chinese accent in blue and Asian face/American accent in red). Onset of the critical word is indicated by the vertical bar (negative is plotted up). Time is plotted on the x-axis; each tick mark indicates 200 ms. LF=left frontal; RF=Right frontal; LP=left posterior; RP=right posterior.*

## References

**1**.Kang, O., & Rubin, D. L. (2009). Reverse linguistic stereotyping: Measuring the effect of listener expectations on speech evaluation. **2**. Lippi-Green, R. (2012). *English with an accent: Language, ideology and discrimination in the United States.* **3.**Cristia, A., Seidl, A., Vaughn, C., Schmale, R., Bradlow, A., & Floccia, C. (2012). Linguistic Processing of Accented Speech Across the Lifespan. **4.** Goslin, J., Duffy, H., & Floccia, C. (2012). An ERP investigation of regional and foreign accent processing. **5**. Grey, S., & Van Hell, J. G. (2017). Foreign-accented speaker identity affects neural correlates of language comprehension. **6**. Hanulíková, A., Van Alphen, P. M., Goch, M. M., & Weber, A. (2012). When one person's mistake is another's standard usage: The effect of foreign accent on syntactic processing. **7**.Romero-Rivas, C., Martin, C. D., & Costa, A. (2015). Processing changes when listening to foreign-accented speech **8**. Babel, M., & Russell, J. (2015). Expectations and speech intelligibility. **9**.Banks, B., Gowen, E., Munro, K. J., & Adank, P. (2015). Audiovisual cues benefit recognition of accented speech in noise but not perceptual adaptation.