Effects of Different Types of Noise on Foreign Accent Adaptation

María Gavino, Niki Katherine García-Holmes, Elisa Ferracane, Cynthia P. Blanco, Rajka Smiljanic

Department of Linguistics, The University of Texas at Austin

I. Introduction

- **Foreign-accented (FA)** speech can negatively affect speech intelligibility, efficiency, accuracy of linguistic processing, and the ability to recall what has been heard [4, 7, 8, 11].
  - Listeners can adapt quickly to FA speech. [1, 2, 3].
  - Familiarity with accent outside the lab or before test facilitates processing of accented words [12]

- Real-world listening conditions often occur in noisy environments:
  - Similar to FA, noise reduces perceptual accuracy, attentional capacity, and memory capacity [10, 9, 5].

- Two kinds of noise:
  1. **Speech-shaped noise (SSN):** a white noise signal shaped to match the long-term spectrum of speech
     - SSN compromises audibility of the speech signal
     - when perceiving speech in SSN, listeners rely more on acoustic cues [6]
  2. **2-talker babble (2T):** competing speech with 2 background talkers
     - 2T presents linguistic/lexical competition with the target speech
     - when perceiving speech in competing speech, listeners rely more on lexical-semantic knowledge [6]

II. Research Questions

1. How does noise affect FA adaptation?

→ Expect lower accuracy and slower response times in noise vs. quiet

2. Does competing speech affect FA differently from white noise?

→ Expect FA to be more successful in 2T than in SSN because target speech stream will be more audible through competing speech

3. Does a listener's familiarity with the FA affect adaptation?

→ Expect listeners (primarily from Central Texas) to better adapt to the more familiar Spanish-accented speech compared to less familiar Korean-accented speech

III. Method

TALKERS
4 female talkers :
- 1st language **American English** (mean age: 19)
- 1st language **Korean** (mean age: 29)
  - Began learning English at 13, in U.S. for 3 years
- 1st language **Spanish** (mean age: 31)
  - Began learning English at 9, in U.S. for 5 years

STIMULI
- 48 meaningful English sentences (e.g. *The mean bear ate the fruit.*) [5]
- Recorded in conversational speech
- 4 target words per sentence, 48 unique final words
- Experiment 1: Presented in quiet,
- Experiment 2: Mixed with SSN (at -2dB SNR)
- Experiment 3: Mixed with 2T (-2SNR)

Quiet
SSN
2T

| The | woman | met | the | rich | actor |
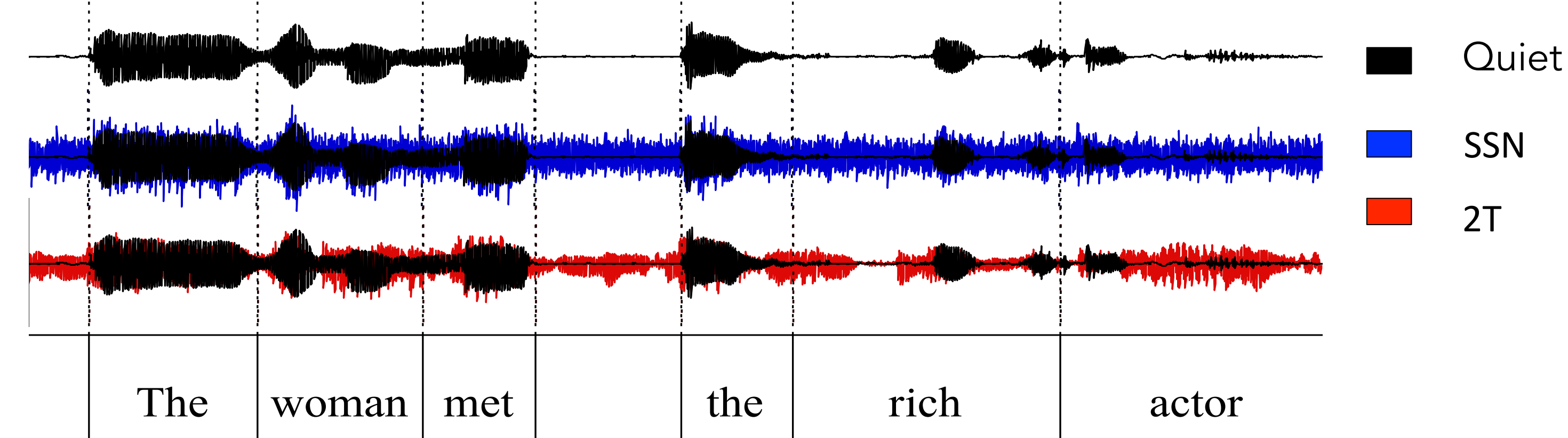
LISTENERS
- Monolingual AE listeners (age range 18-35); 65 in Ex1, 65 in EX 2, and 50 in Ex 3
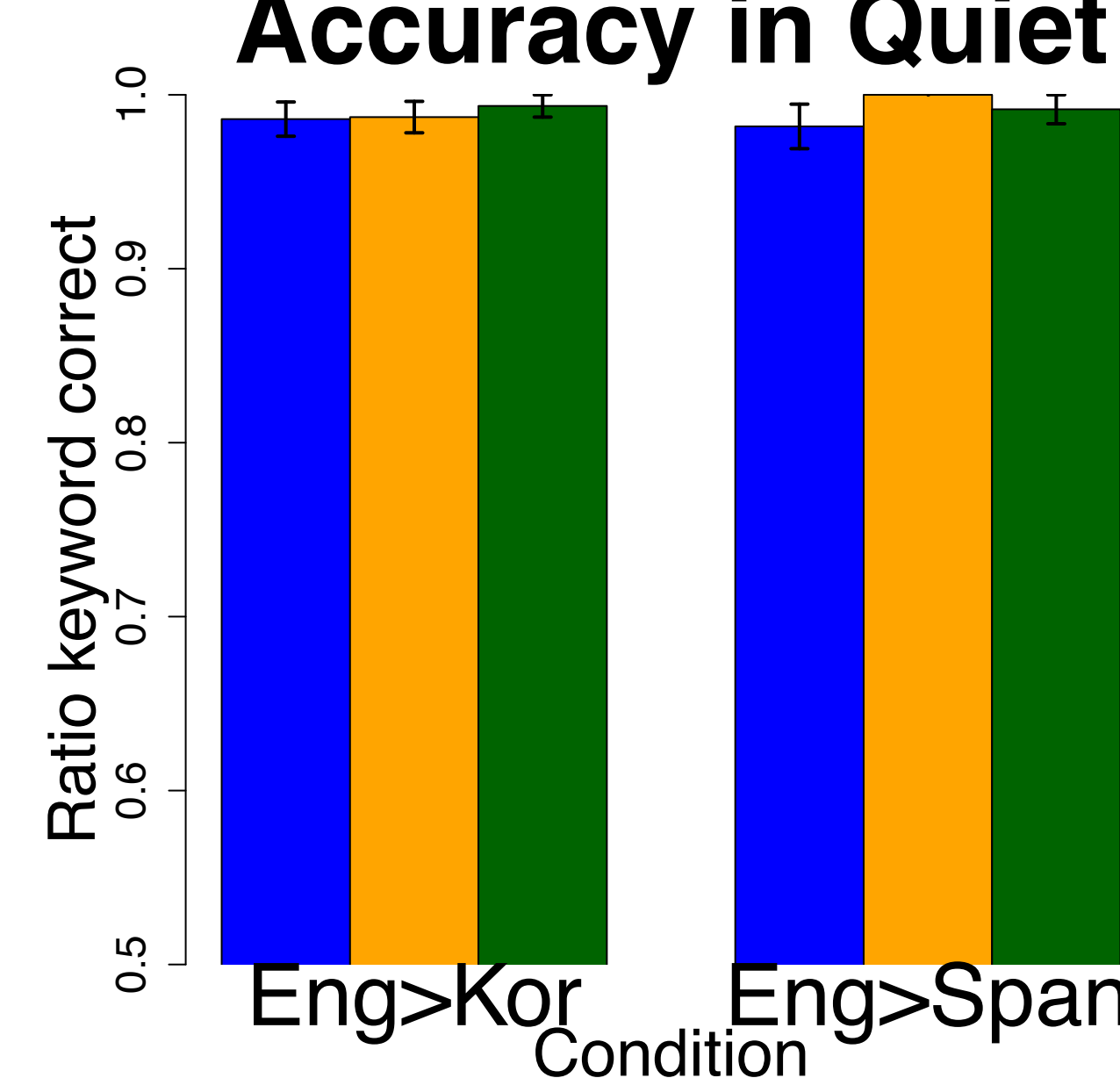
TASK: Visual probe decision
- Participants hear sentence and decide if the word written onscreen is the final word of the sentence they heard
  - 4 blocks of 12 sentences, each block produced by a different talker
  - 4 sentences in each block had mismatched probes (total of 16 mismatched across all blocks)

  - **Control conditions (3):** Talker changes (all blocks in same accent: English, Spanish or Korean)
  - **Switch conditions (2):** FA change (Blocks 2 and 3 in the same FA): Eng > Span or Eng > Kor

Block 1 — Block 2 — Block 3 — Block 4
Foreign-accented productions
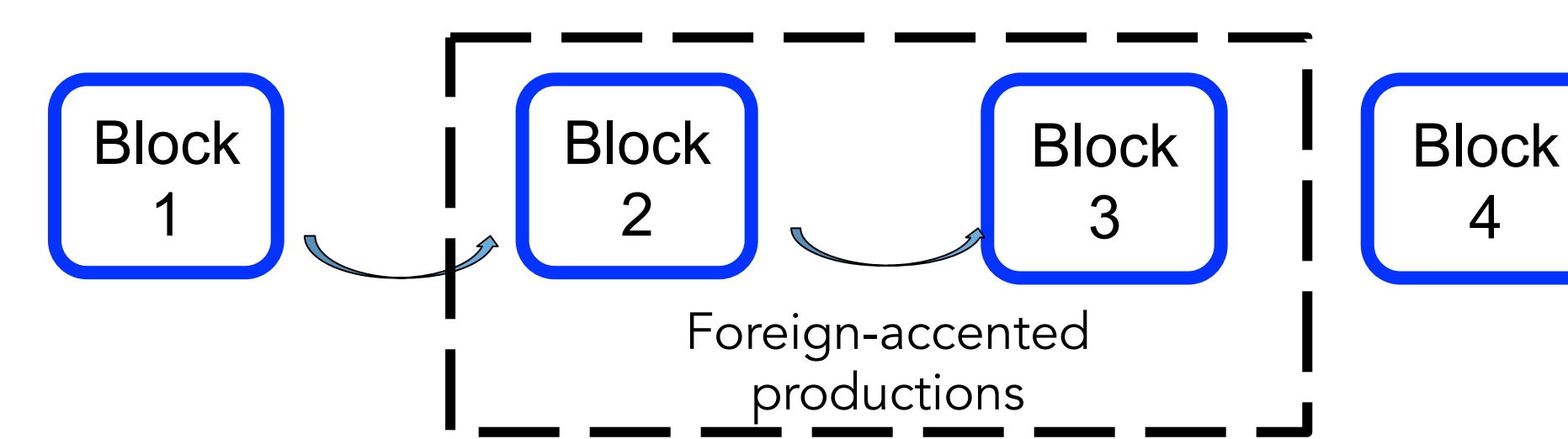
- **Analysis:** Accuracy (probe word decision) and reaction time (RT)

IV. Results

Accuracy in Quiet — Block1, Block2, Block3
Ratio keyword correct / Condition — Eng>Kor, Eng>Span

Quiet:
- High accuracy for all accents across all blocks (performing at ceiling)

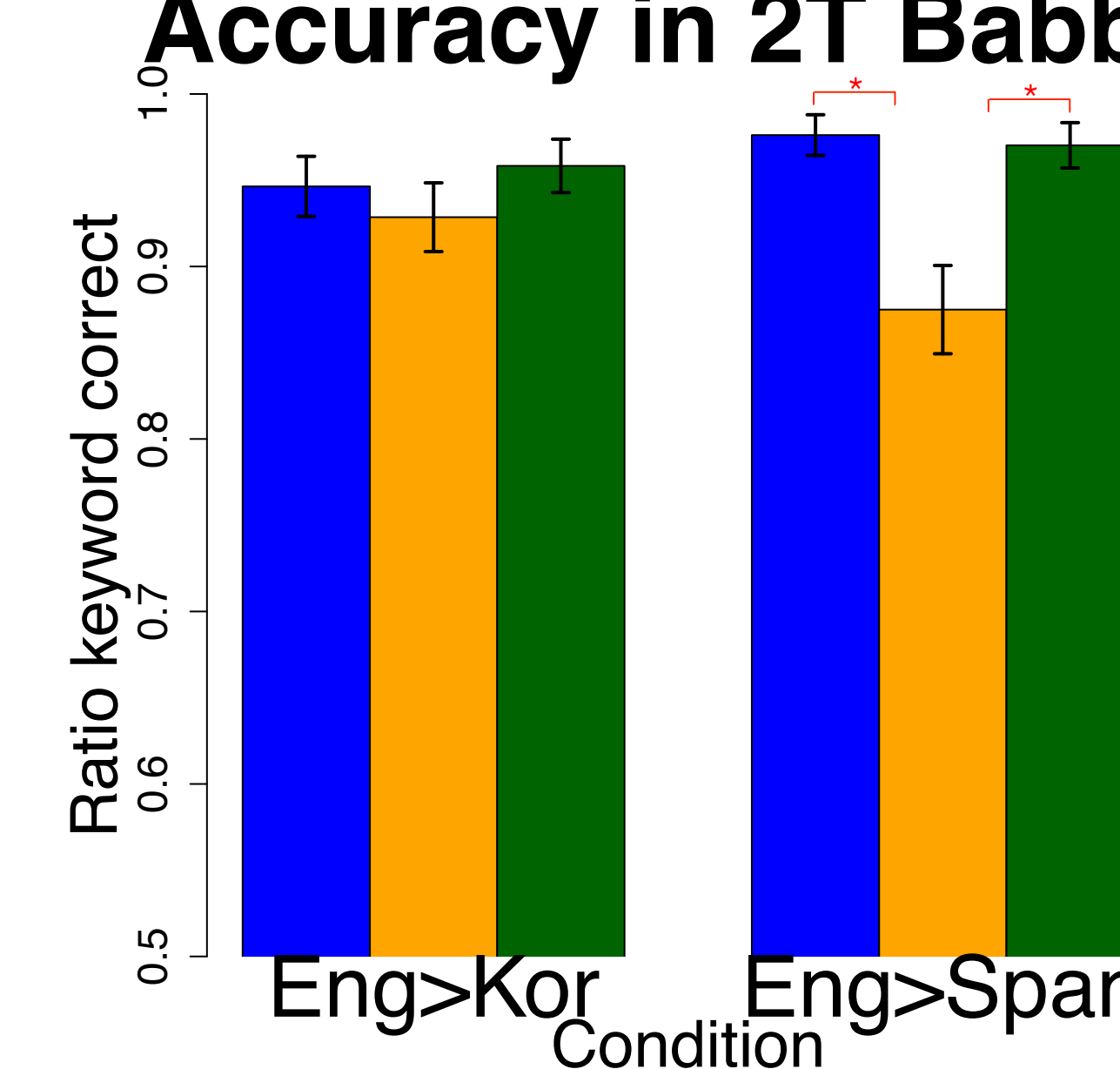**Accuracy in SSN**
Ratio keyword correct / Condition — Eng>Kor, Eng>Span
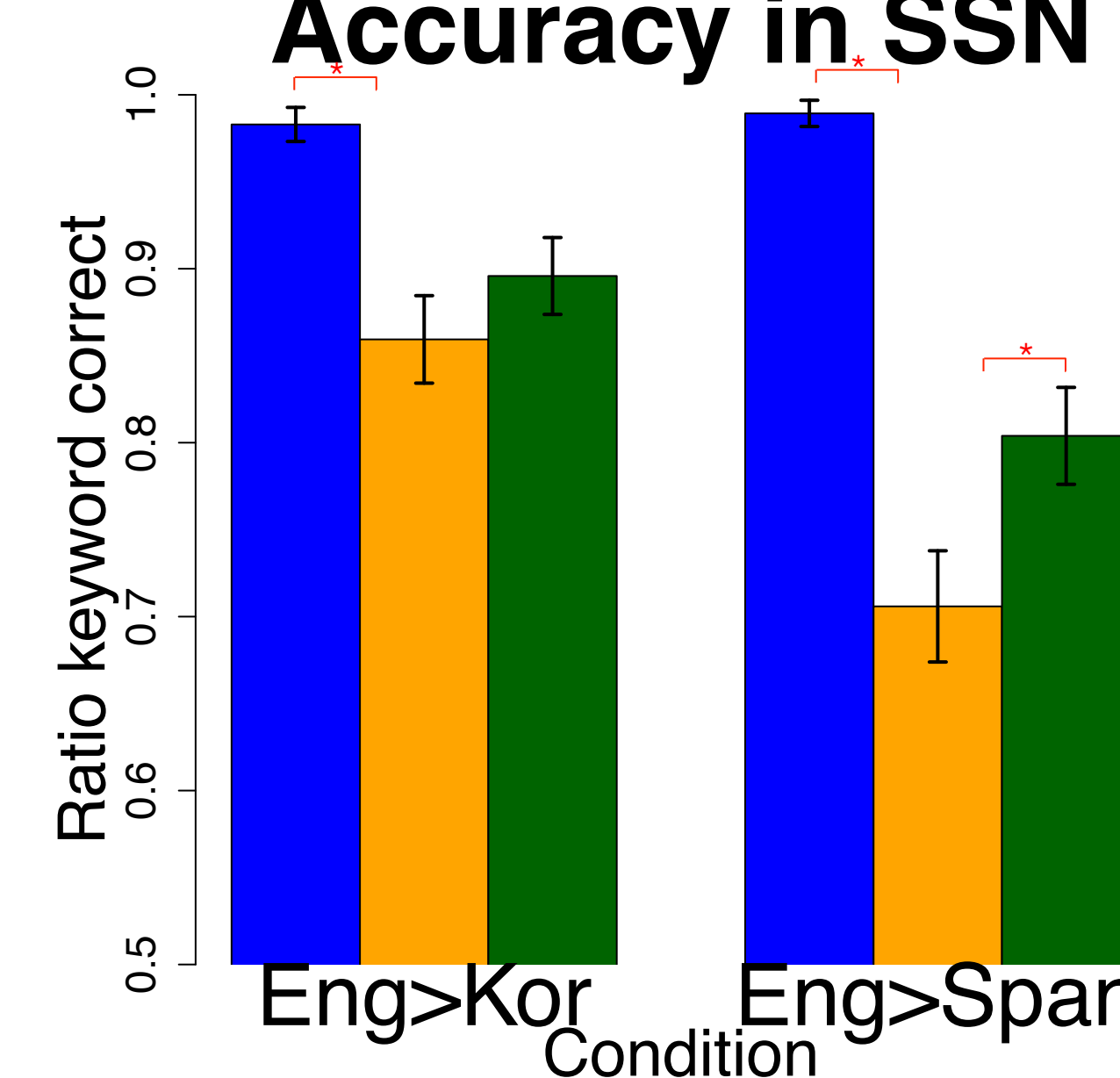
SSN:
- Lower accuracy for Span. vs. Kor.
- Some evidence of adaptation:
  - Block3 more accurate than Block2 in Span. switch

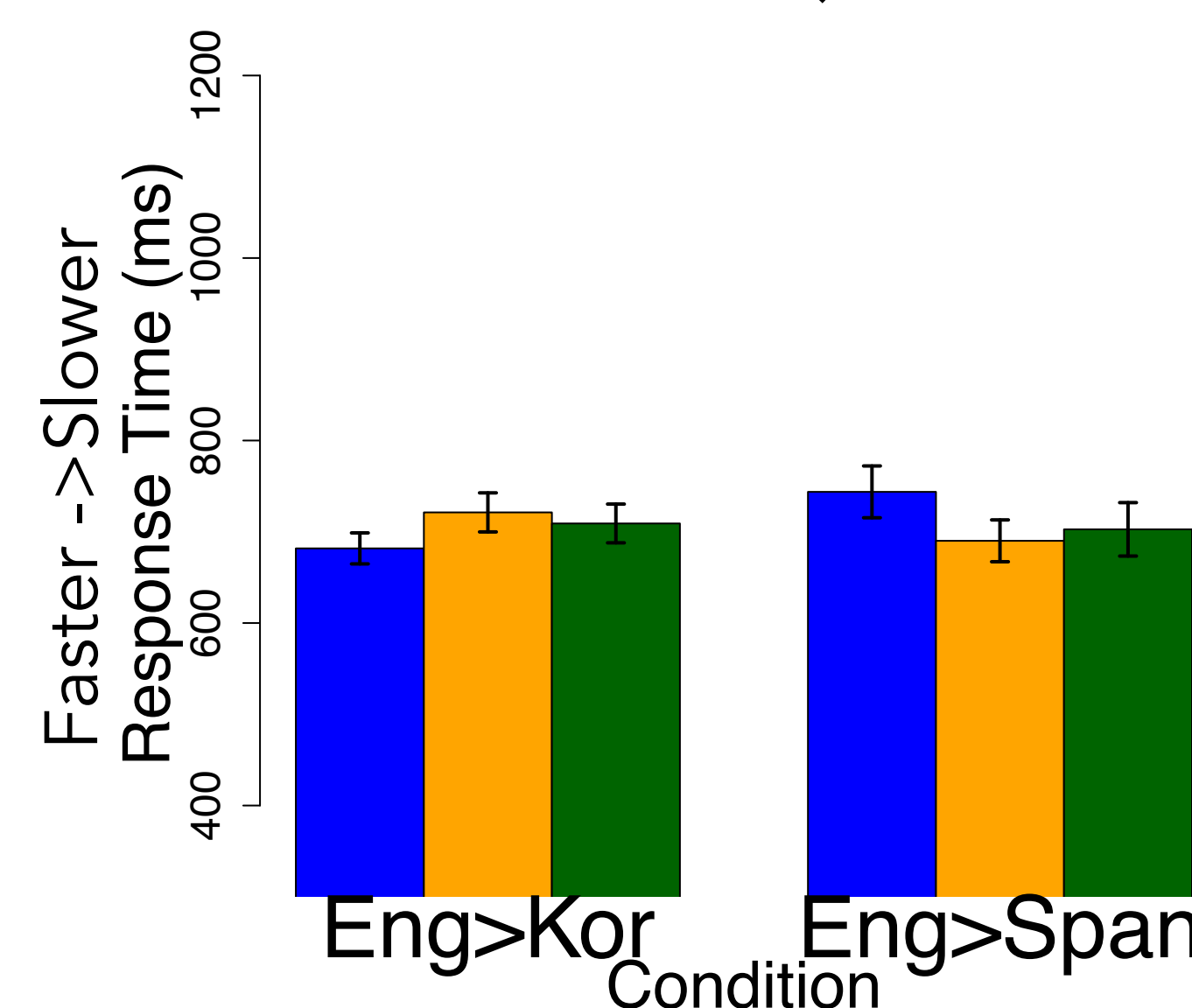**Accuracy in 2T Babble**
Ratio keyword correct / Condition — Eng>Kor, Eng>Span

2T babble:
- Overall higher accuracy than SSN
- No difference between Kor. and Span.
- Some evidence of adaptation:
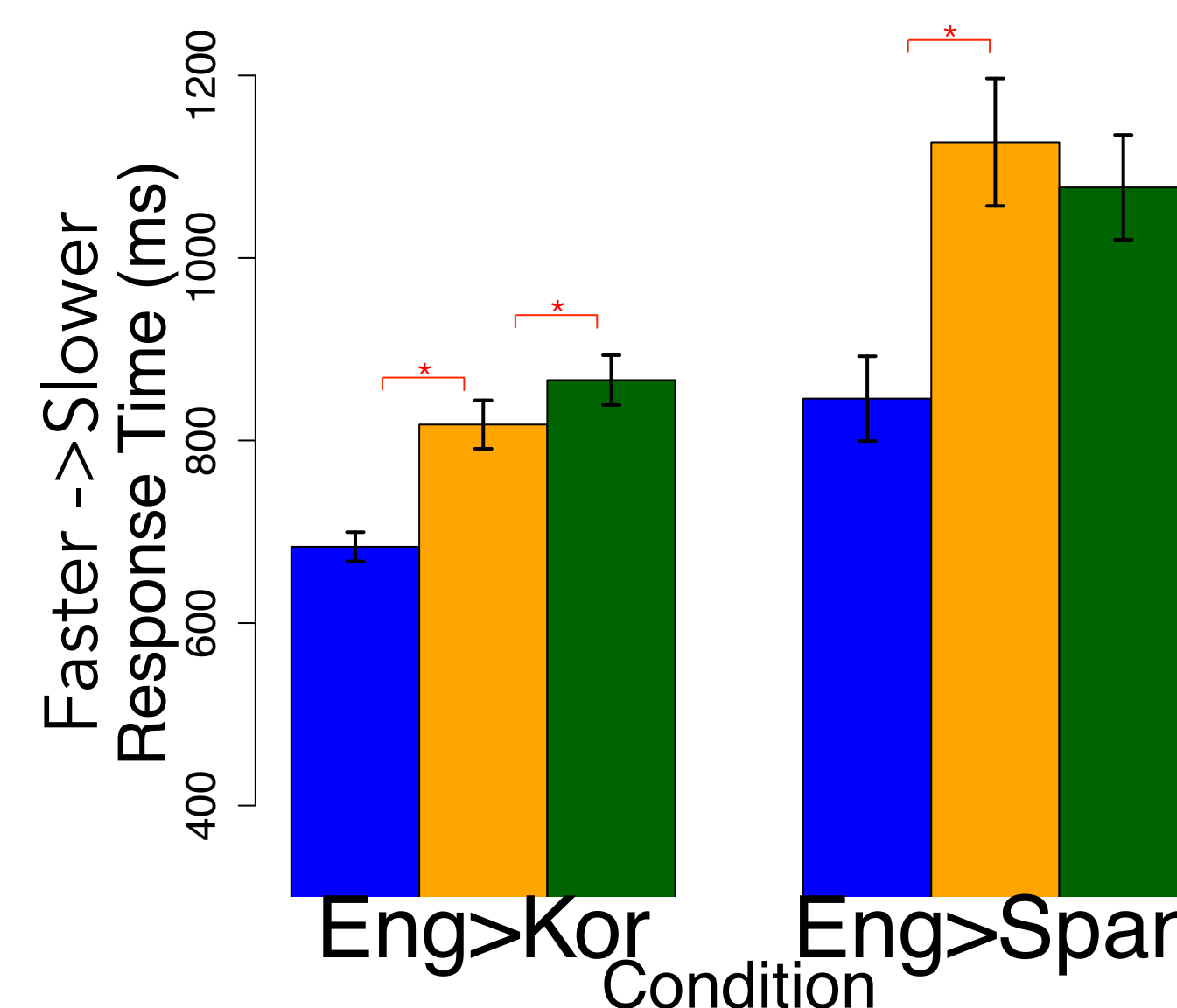 - Block3 more accurate than Block 2 in Span. switch

RT in Quiet
Faster ->Slower Response Time (ms) / Condition — Eng>Kor, Eng>Span

Quiet:
- No significant effect of accent on RT
- No evidence of adaptation (performing at ceiling)

RT in SSN
Faster ->Slower Response Time (ms) / Condition — Eng>Kor, Eng>Span

SSN:
- Listeners are overall slower in SSN than in quiet
- Marginally significant effect of accent; Kor. faster than Span
- No evidence of adaptation:
  - RTs do not improve from Block2 to Block3 in Span.
  - Adaptation rate worsens from Block2 to Block3 in Kor.

RT in 2T Babble
Faster ->Slower Response Time (ms) / Condition — Eng>Kor, Eng>Span
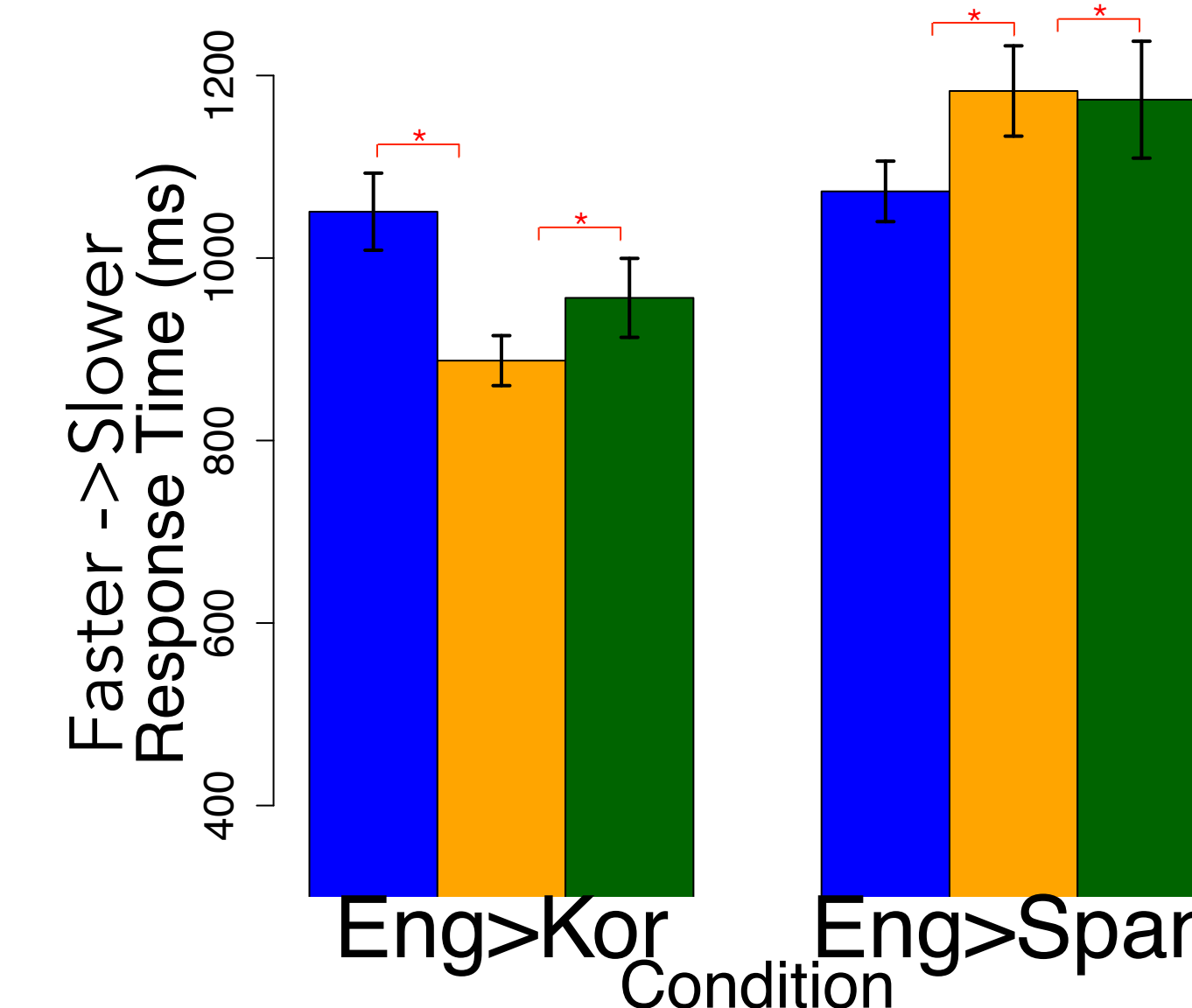
2T babble:
- Listeners are overall slower in 2T than in SSN
- Significant effect of accent; Kor. faster than Span.
- Some evidence of adaptation:
  - RTs improve from Block1 to Block2 in Kor. But adaptation rate worsens from Block2 to Block3 in Kor.
  - RTs marginally improve from Block2 to Block3 in Span.

V. Conclusion

- **RQ1:**
  - Compared to Quiet, listeners were overall slower and less accurate in noise
  - FA adaptation was disrupted by both SSN and 2T babble
  - No evidence of increased accuracy since there was faster RTs across blocks 2 and 3

- **RQ2:**
  - Listeners were more accurate, but slower in 2T babble compared to SSN
  - More evidence of adaptation in 2T babble where listeners are able to marginally adapt to the second Spanish speaker
  - Processing *FA speech in speech* may be more effortful (slower RTs)
  - More successful FA adaptation in 2T babble may be related to listeners' ability to successfully separate FA target from the native-accented background speech
  - 2T babble may be less disruptive than SSN for FA adaptation even though listeners were processing FA speech with an increased cognitive load (lexical competition)

- **RQ3:**
  - Spanish-accented speech was more difficult than Korean-accented speech
  - But longer RTs in both SSN and 2T for second Korean talker could be attributed to difficulty with unfamiliar accent

FUTURE WORK
- Examine intelligibility and accentedness of the L2 talkers
- Analyze the acoustic properties (i.e. location of pauses, speaking rate) and their relation to FA adaptation
- Investigate the impact of foreign-accented 2T babble on FA adaptation

REFERENCES
1. Bradlow, A. R., and Bent, T. (2003). Listener adaptation to foreign-accented English. Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona, Spain, 2003, edited by M. J. Sole, D. Recasens, and J. Romero, pp. 2881–2884.
2. Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. Journal of the Acoustical Society of America, 116 (6), 3647–3658.
3. Cristià, A., Seidl, A., Vaughn, C., Schmale, R., Bradlow, A., & Floccia, C. (2012). Linguistic processing of accented speech across the lifespan. Frontiers in Psychology, 3, 479.
4. Gilbert, R., Chandrasekaran, R., and Smiljanic, R. (2014). Recognition memory in noise for speech of varying intelligibility. Journal of the Acoustical Society of America,135 (1), 389-399.
5. Leek, M. R., Watson, C. S. (1984). Learning to detect auditorypattern components. Journal of the Acoustical Society of America, 76, 1037-1044.
6. Mattys, S. L., Brooks, J., & Cooke, M. (2009). Recognizing speech under a processing load: Dissociating energetic from informational factors. Cognitive Psychology, 59, 203 243.
7. Munro, M. J. and Derwing, T. M. (1995), Foreign Accent, Comprehensibility, and Intelligibility in the Speech of Second Language Learners. Language Learning, 45: 73–97.
8. Schmid, Peggy M., Yeni-Komshian, Grace H. (1999). The effects of speaker accent and target predictability on perception of mispronunciations. Journal of speech, language, and hearing research. 42, 56-64.
9. Vaden, K.I. Jr., Kuchinsky S.E., Cute, S.L., Ahlstrom, J.B., Dubno, J.R., Eckert, M.A. (2013). The cingulo-opercular network provides word-recognition benefit. J Neurosci, 33, 18979 –18986.
10. Van Engen, K. J., Chandrasekaran, B., & Smiljanic, R. (2012). Effects of speech clarity on recognition memory for spoken sentences. PloS one, 7 (9), e43753.
11. Wilson, E. O., & Spaulding, T. J. (2010). Effects of noise and speech intelligibility on listener comprehension and processing time of Korean-accented English. Journal of Speech, Language, and Hearing Research, 53, 1543–1554.
12. Witteman, M. J., Weber, A., & McQueen, J. M. (2013). Foreign accent strength and listener familiarity with an accent codetermine speed of perceptual adaptation. Attention, Perception, and Psychophysics, 75, 537-556.
The authors would like to thank members of the UT SoundLab for help with data collection, and Erika Hale and Sally Amen for guidance with the statistical analyses.

QUESTIONS? EMAIL DR. RAJKA SMILJANIC: rajka@mail.utexas.edu