

Why [spa] not [psa]? On the perceptual salience of initial /s/-stop and stop-/s/ sequences

ASIMINA SYRIKA^{*}, JAN EDWARDS^{*}, MARIOS FOURAKIS^{*}, EUN JONG KONG^{*}, BENJAMIN MUNSON^{**}, MARY E. BECKMAN^{***}

University of Wisconsin-Madison^{*}, University of Minnesota^{**}, Ohio State University^{***}

INTRODUCTION

- Initial /s/-stop clusters (e.g., /sp/) occur frequently in the world's languages, but initial stop-/s/ clusters (e.g., /ps/) are relatively infrequent.
- In a large typological study, Morelli (1999; 2003) found that languages that allow stop-/s/ clusters in initial position also allow /s/-stop clusters, while the reverse is not true.
- Moreover, in languages that contain both sequence types (e.g., Greek) children produce /s/-stop clusters correctly before they produce stop-/s/ clusters (Syrika et al., 2007).
- Some researchers (e.g., Wright, 2001; 2004) have proposed a perceptual explanation for the above phenomena:
 - a stop is easier to perceive when it is adjacent to a vowel (as in /s/-stop clusters), since there are cues in the fricative spectrum, stop burst, and vowel transitions
 - a stop released into a fricative may be more easily masked by the high-frequency noise
- However, few studies have attempted to quantify this 'perceptual advantage' of initial /s/-stop clusters experimentally.

MAIN GOAL

- To examine experimentally whether /s/-stop clusters are perceptually more salient than stop-/s/ clusters in initial position in Greek:
 - to look at the effect of signal-to-noise ratio (snr)
 - to look at the effect of vowel context
- To examine how the acoustic cues in the fricative can signal the stop place of articulation in initial /s/-stop and stop-/s/ sequences

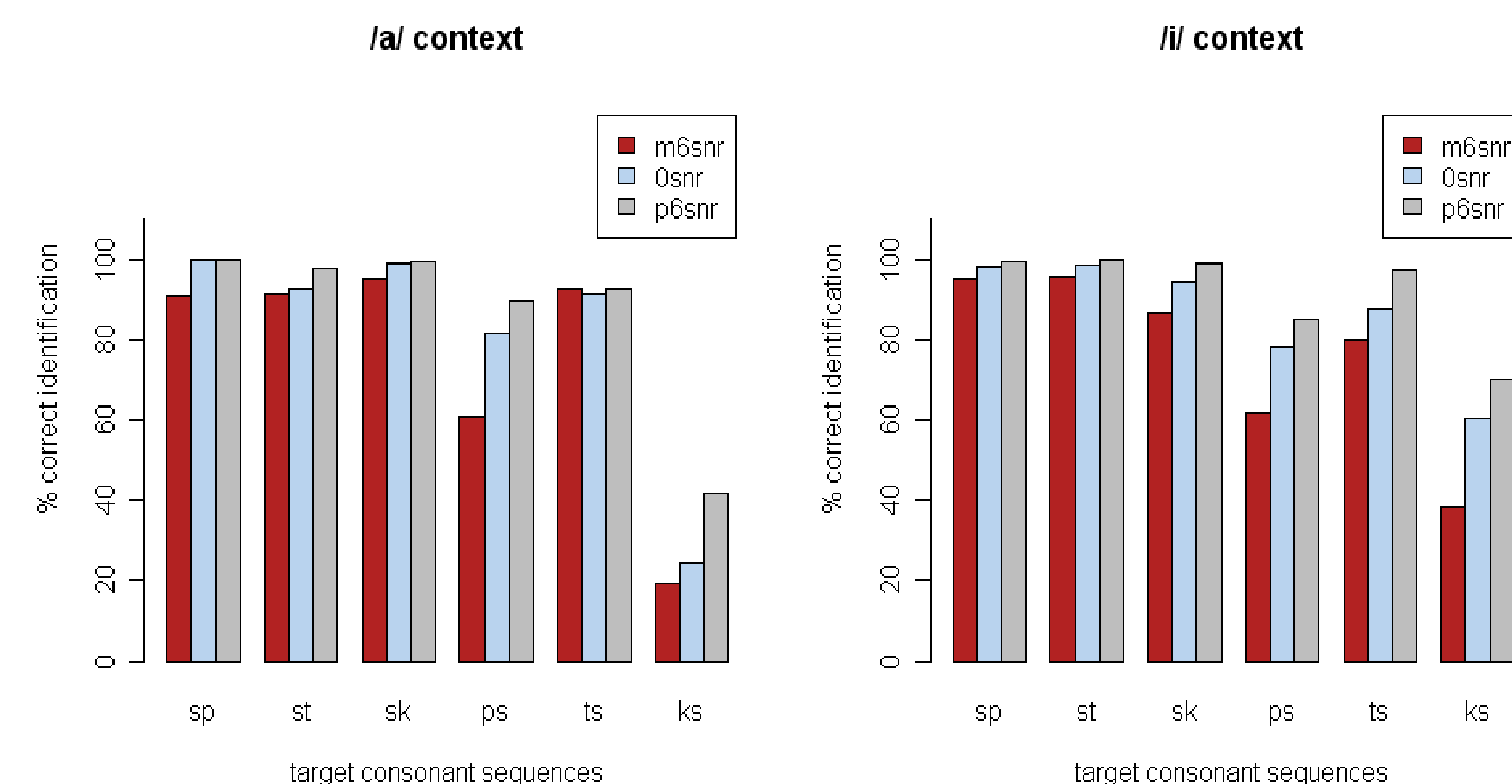
METHOD

- Participants:** 20 Greek young adults with normal hearing
- Stimuli:** extracted CCV sequences beginning with /s/-stop and stop-/s/ clusters
 - extracted sequences were from real word productions of 20 Greek-speaking young adults
 - two vowel contexts: /a/ and /i/
 - syllables were mixed with Greek multitalker babble (noise)
 - three signal-to-noise ratios: -6, 0, +6 dB
- Procedure:** forced identification task (3-choice)
 - two different blocks: /s/-stop and stop-/s/

TASK



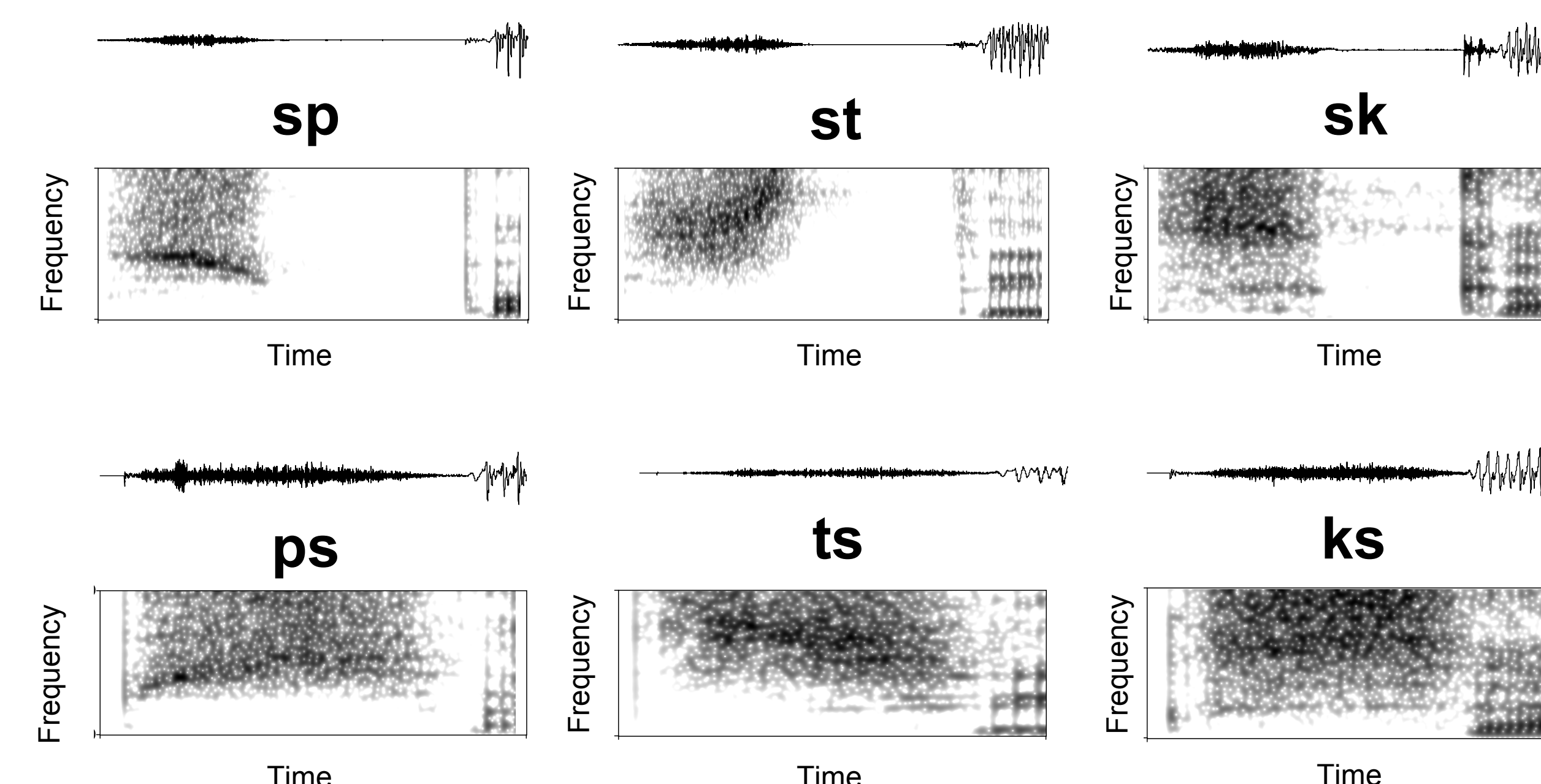
FIGURE 1. IDENTIFICATION IN NOISE BY VOWEL AND SNR



RESULTS: IDENTIFICATION ACCURACY

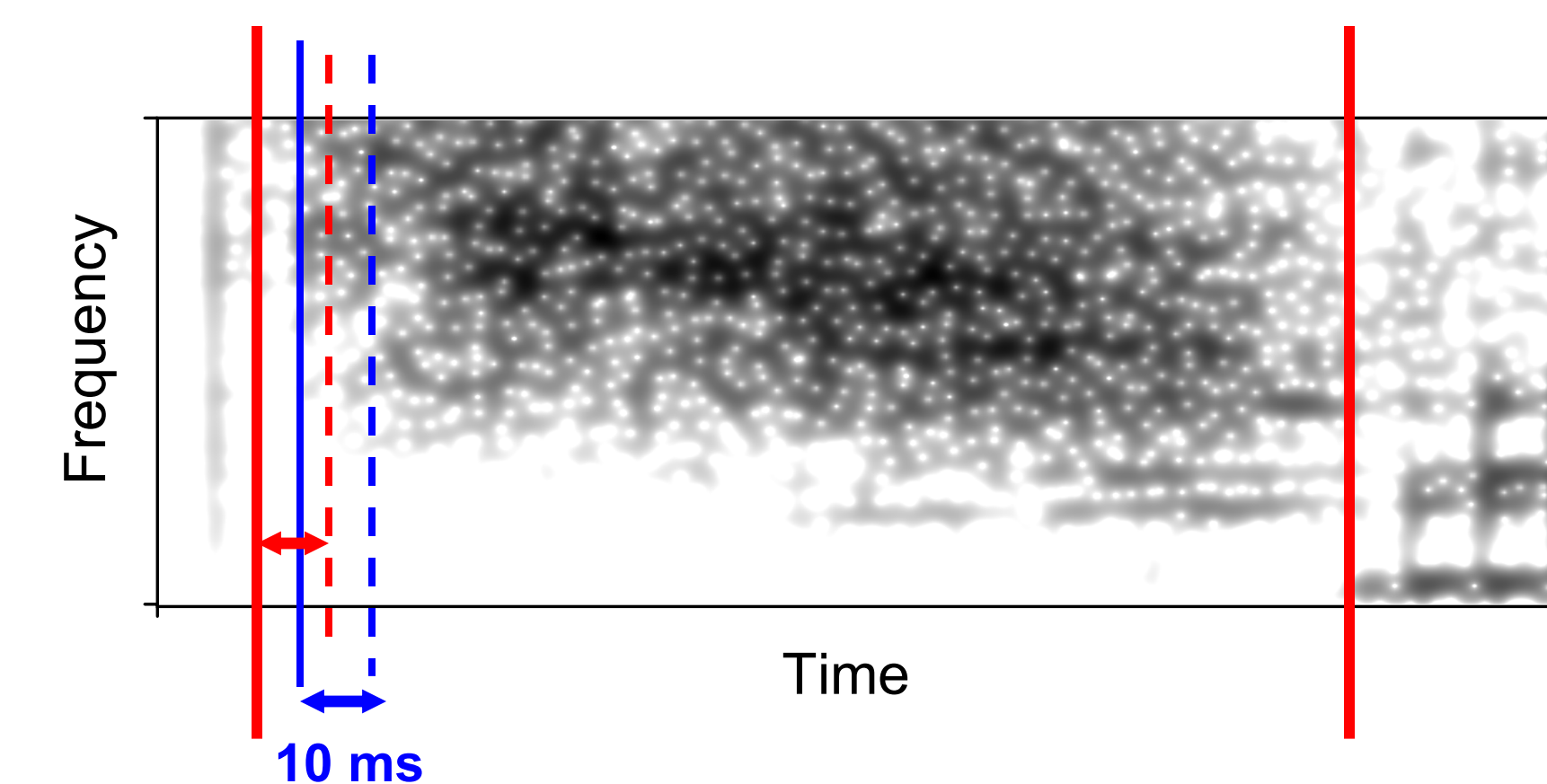
- /s/-stop clusters:
 - correctly identified even in the noisiest condition
- Stop-/s/ sequences:
 - /ps/ and /ks/ less accurately identified than affricate /ts/
 - /ks/ was the least correctly identified
- Effect of snr in the predicted direction
- Effect of vowel context:
 - /ts/ more accurately identified before /a/ (for -6, and 0 dB condition)
 - /sk/ more accurately identified before /a/ (for -6 dB condition)
 - /ks/ more accurately identified before /i/ (all conditions)

ACOUSTIC CUES TO STOP PLACE OF ARTICULATION



- The energy concentration for /s/ changes depending on adjacent stop consonant:
 - downward direction before /p/
 - upward direction before /t/
 - straight direction before /k/
- Mirror image of /s/-stop clusters:
 - upward direction after /p/
 - downward direction after /t/
 - straight direction after /k/

QUANTIFYING PLACE INFORMATION IN FRICATIVE NOISE



- Peak ERB:** frequency of the loudest peak in the sones by ERB spectrum (Moore et al., 1997)
 - psychoacoustic analog to the peak frequency
- Direction of change in Peak ERB:**
 - calculated peak ERB at onset and offset of frication for each stimulus item

FIGURE 2. PEAK ERB START BY PEAK ERB END FOR /S/-STOP (LEFT) AND STOP-/S/ CLUSTERS (RIGHT) IN THE NOISIEST CONDITION

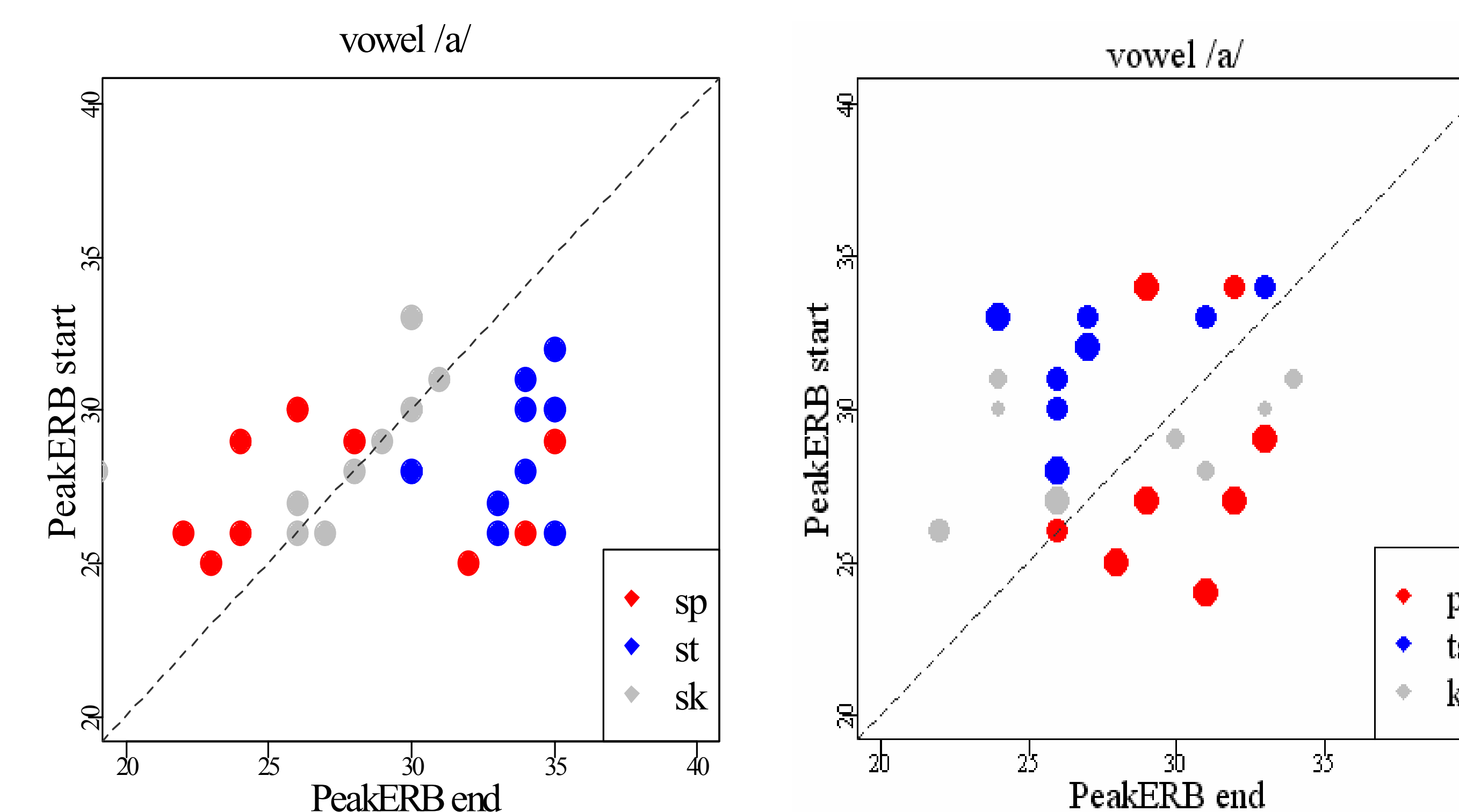


FIGURE 3. RELATING PERCEPTION TO ACOUSTICS

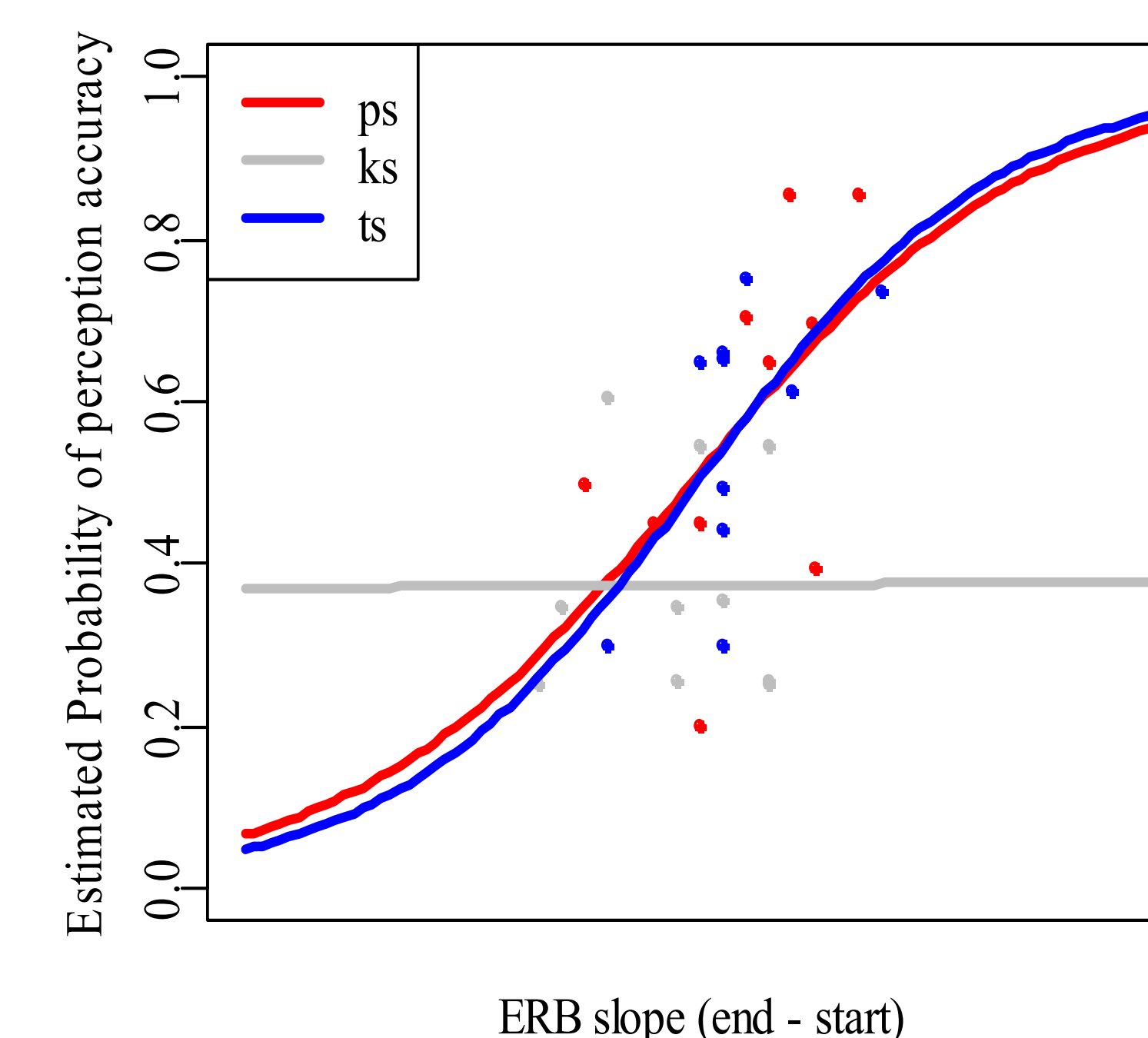


Figure 3 shows the fitted regression line from a Hierarchical Logistic Regression Model predicting accuracy (correct or incorrect) as a function of ERB slope (end-start) for the different stop-/s/ sequences in the noisiest condition (-6 dB)

FIGURE 2 RESULTS: /s/-stop

- /sp/ tokens generally had high start values and low end values of peak ERB
- /st/ tokens generally had low start values and high end values of peak ERB
- /sk/ tokens had similar start values and end values of peak ERB

FIGURE 2 RESULTS: stop-/s/

- /ps/ tokens generally had low start values and high end values of peak ERB
- /ts/ tokens generally had high start values and low end values of peak ERB
- /ks/ tokens did not show a consistent pattern

FIGURE 3 RESULTS

- Changes in peak ERB slope significantly affected identification accuracy for /ps/ and /ts/
- No relationship between ERB slope and identification accuracy in the case of /ks/

DISCUSSION AND CONCLUSION

- /s/-stop clusters more accurately identified in noise than stop-/s/ clusters (except affricate /ts/):
 - suggests that /s/-stop clusters are more perceptually robust than stop-/s/ clusters
 - supports perceptual salience explanation of Greek cluster types in acquisition and across languages
- /ps/ and /ks/ clusters less accurately identified in noise than affricate /ts/
- /ks/ clusters had the lowest identification scores
- Low performance of /ks/ clusters because they might be least well-differentiated acoustically
- Importance of perceptual salience:
 - time-course of acquisition
 - error patterns
 - distribution of sounds across languages
- Perceptual salience can be quantified in terms of:
 - resistance to noise
 - acoustic differentiation

ACKNOWLEDGMENTS

- Work supported by NIDCD Grant R01DC02932 and NSF Grant BCS-0729140 to Jan Edwards.
- Thanks to the people who participated in the production and perception experiments reported in the current study.
- Special thanks to Ioannis Syrikas for assistance with data collection, and Vivian Iliadou for letting us use the Greek multitalker babble.
- Thanks also to the members of the *Cross-Language Phonology and Word Learning Lab* at the Waisman Center, and the members of the *paidologos* project for their valuable feedback at different parts of this project.

the παιδολογος project
cross-language investigation of phonological development