

Using Eye Gaze to Explore Decision-Making during Binaural Processing: Comparing Temporal Fine Structure and Envelope Cues

Yuhang (Will) Li, Kumari Anshu, Lukas Suveg, Shelly P. Godar, G. Nike Gnanateja, Ruth Y. Litovsky
University of Wisconsin-Madison, USA
e-mail: li2746@wisc.edu

Introduction

Spatial hearing is supported by **interaural time differences (ITDs)** and **interaural level differences (ILDs)**¹.

ITDs are conveyed at low frequencies in the temporal fine structure (TFS) and can also be transmitted via slow envelope modulation (ENV)² of high frequencies. ILDs are most effective at high frequencies.

Bilateral cochlear implant users show reduced ITD sensitivity due to limited temporal precision and reduced access to TFS³.

This study aimed to understand **how constraining temporal cues affected binaural decision-making for simulated CI stimuli in listeners with typical hearing**.

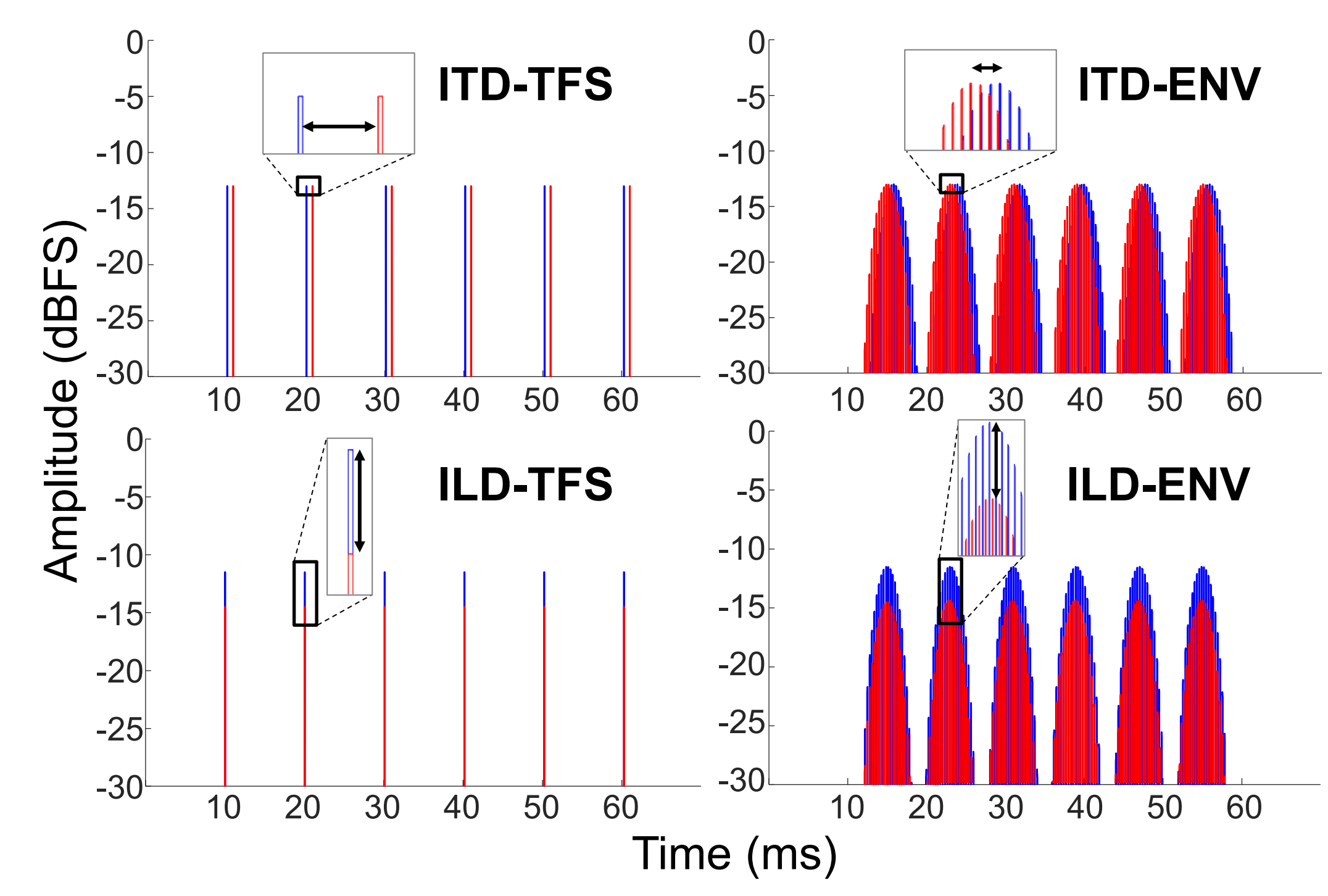
Eye-tracking paradigms, beyond just-noticeable difference (JND) thresholds, provide access to the **time course of decisions** and the **strategies underlying perceptual judgements of binaural stimuli**.

How do ITD and ILD sensitivity differ in JND thresholds and temporal decision dynamics for temporal fine structure versus envelope cues in CI-like stimuli?

Cochlear Implant-like Stimuli

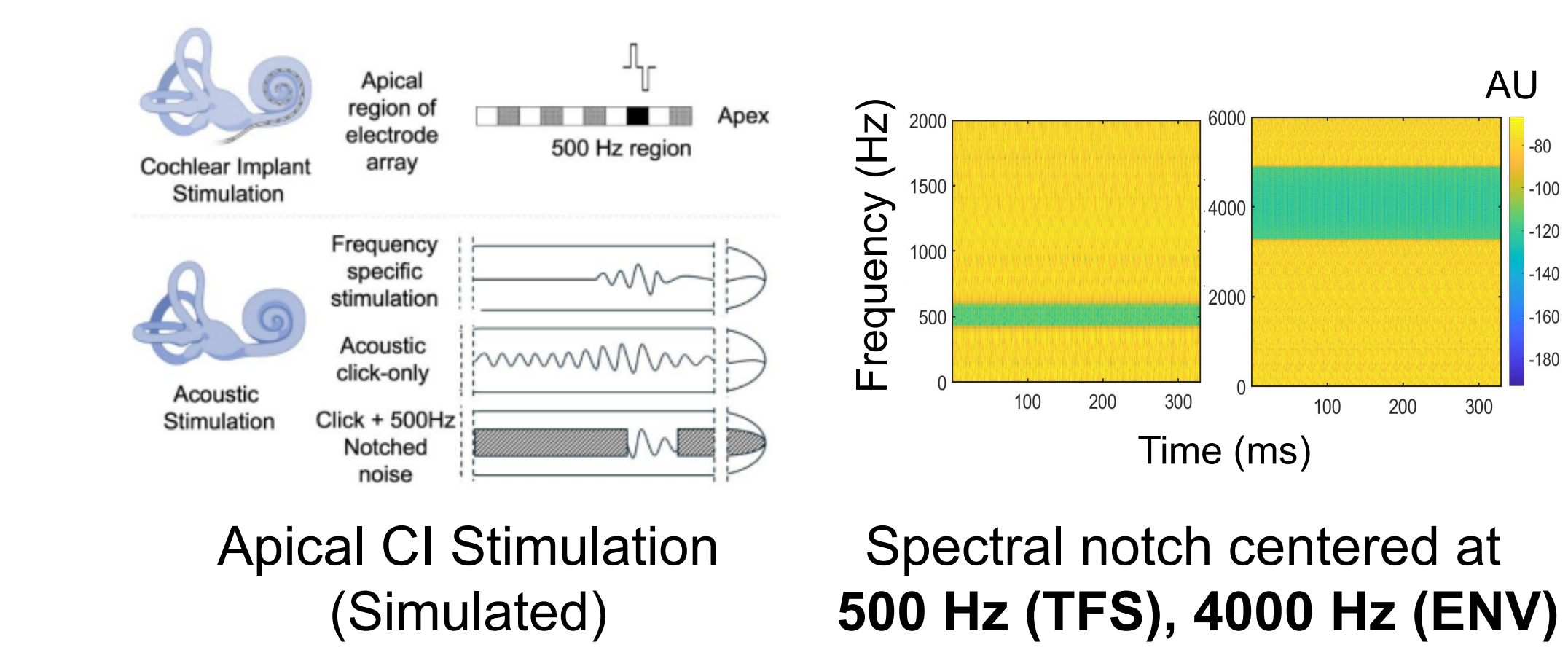
Short duration click trains⁴ with ITD & ILD Cues

- TFS - 100 pulses per second (pps)^{5,6}
- ENV - 4000 pps, 125 Hz amplitude modulation⁷



Stimuli were **embedded in notched noise** to limit **excitation spread**, reducing the broader excitation associated with Gaussian Envelope Tones (GETs)⁷.

Resulting excitation patterns approximate **single-electrode cochlear implant stimulation**.



Methods

A within-subject experiment probed binaural cue processing across **cue type (ITD, ILD)** and **temporal cue (TFS, ENV)**.

Typical-hearing adults (N = 14, ages 19 - 28) completed testing with both cue types and both temporal cues in separate blocks.

Cue magnitudes:

- ITD (μ s)

TFS: 10, 20, 40, 80, 140, 200, 400, 750

ENV: 10, 20, 80, 140, 200, 300, 400, 750

- ILD (dB)

TFS & ENV: 0.5, 1, 2, 3, 6, 10, 16

For each condition, participants completed 30 trials per cue magnitude.

Looking while-listening paradigm^{8,9}

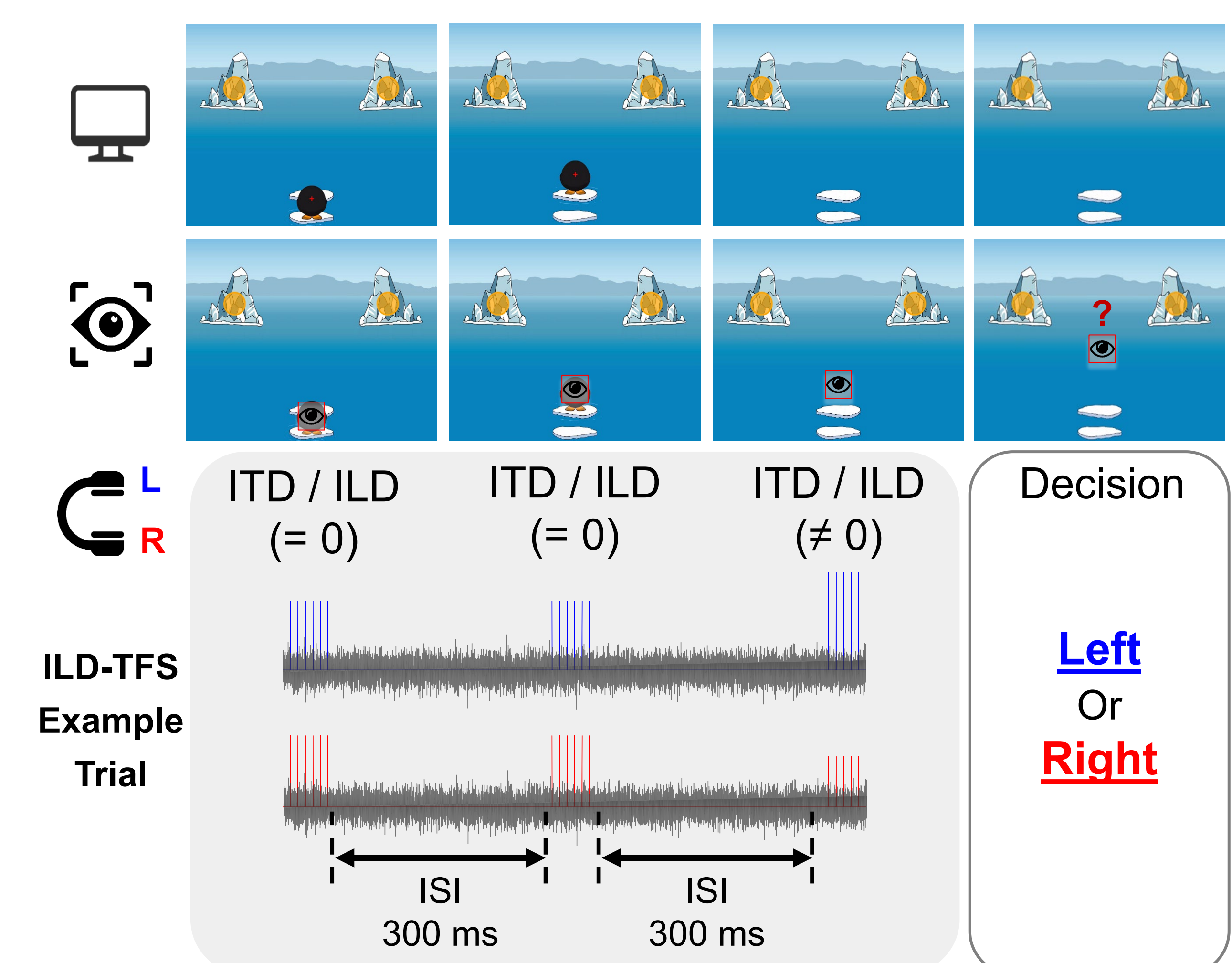


Fig. 1: Left-right judgment in a 3-interval, 2-alternative forced-choice task. Target interval is always the 3rd (50% left, 50% right).

Eye Tracking

Eye gaze was recorded continuously using a **Tobii Pro Fusion** (Tobii AB, Sweden) eye tracker.

Sampling rate: 120 Hz

Results

TFS yields lower JNDs than ENV for both ITD and ILD.

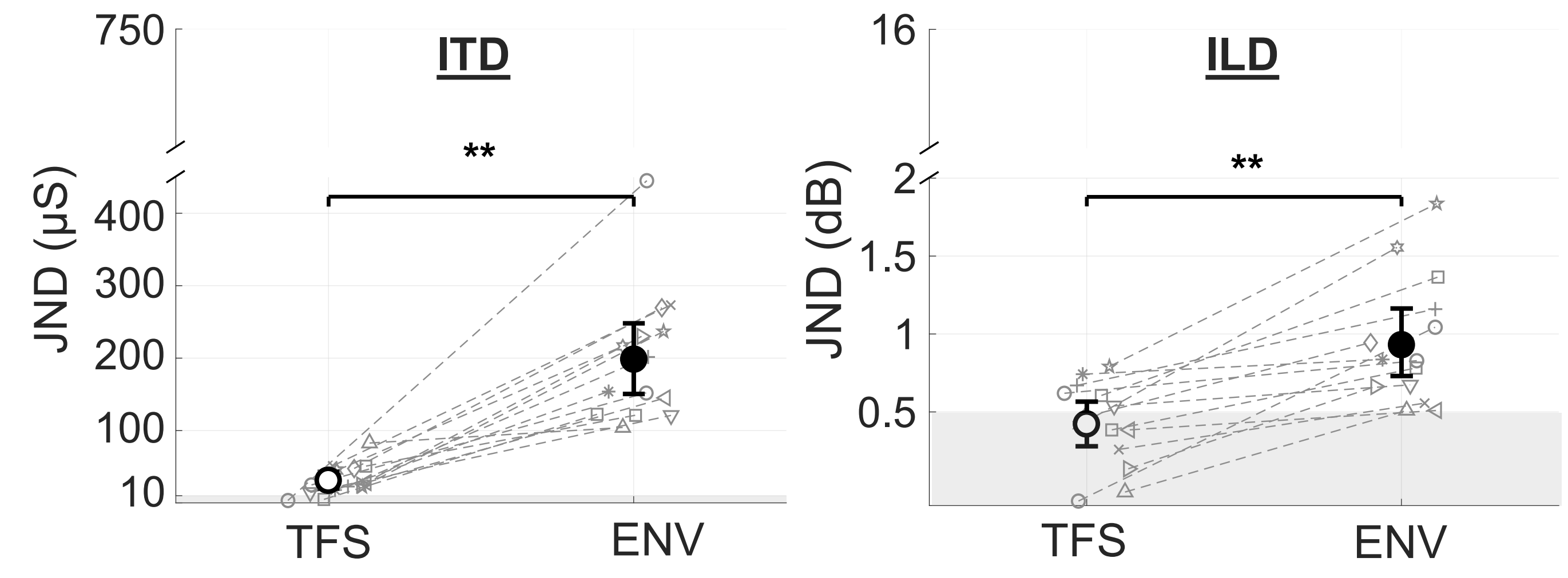


Fig. 2: JND thresholds were derived from logistic fits (*psignifit*^{10,11}) at 70.7% correct. Gray region denotes JNDs below the tested cue range (estimated).

Response times decreased with increasing cue magnitude, with faster responses for TFS than ENV.

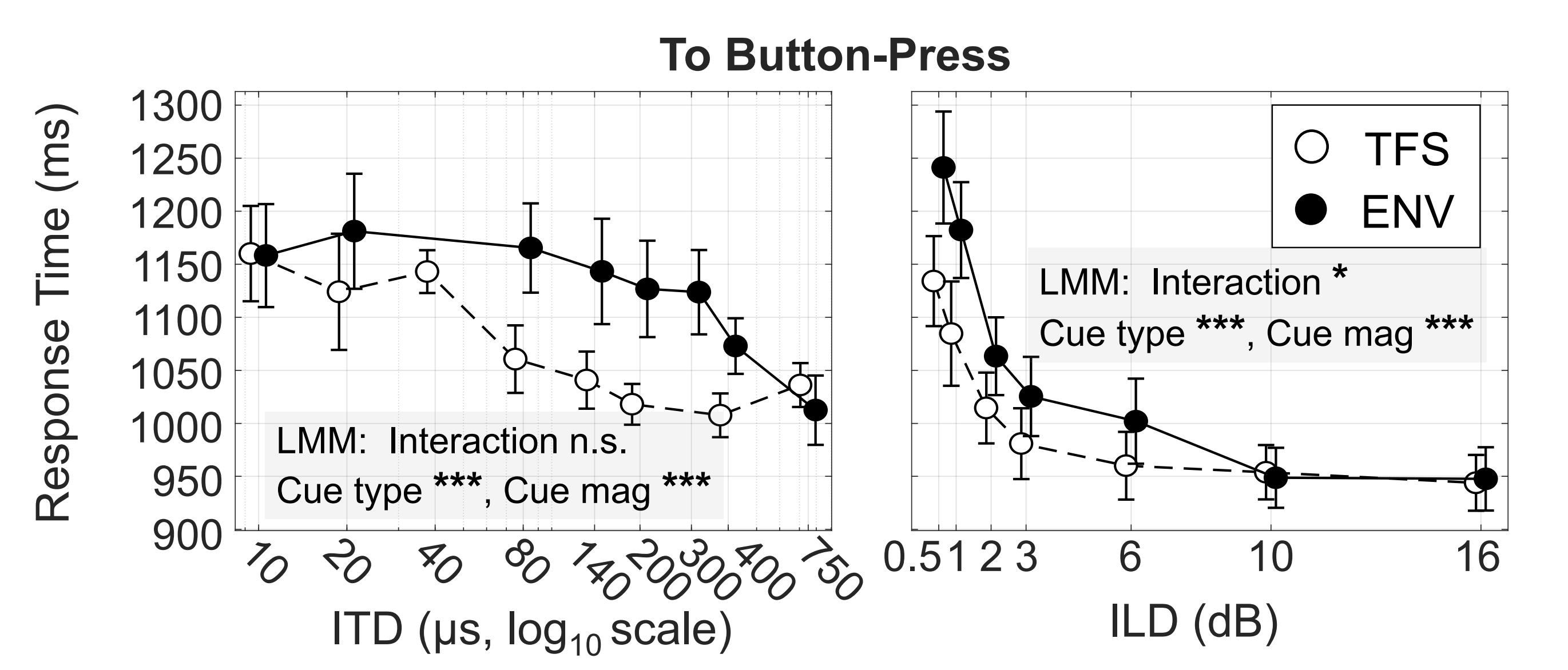


Fig. 3: Group means (± 1 SE) of response time from target stimulus onset to button press, plotted as a function of ITD magnitude ($\log_{10}$ scale) and ILD magnitude for TFS and ENV conditions.

Gaze to the correct target rises earlier and more steeply as cue magnitude increases.

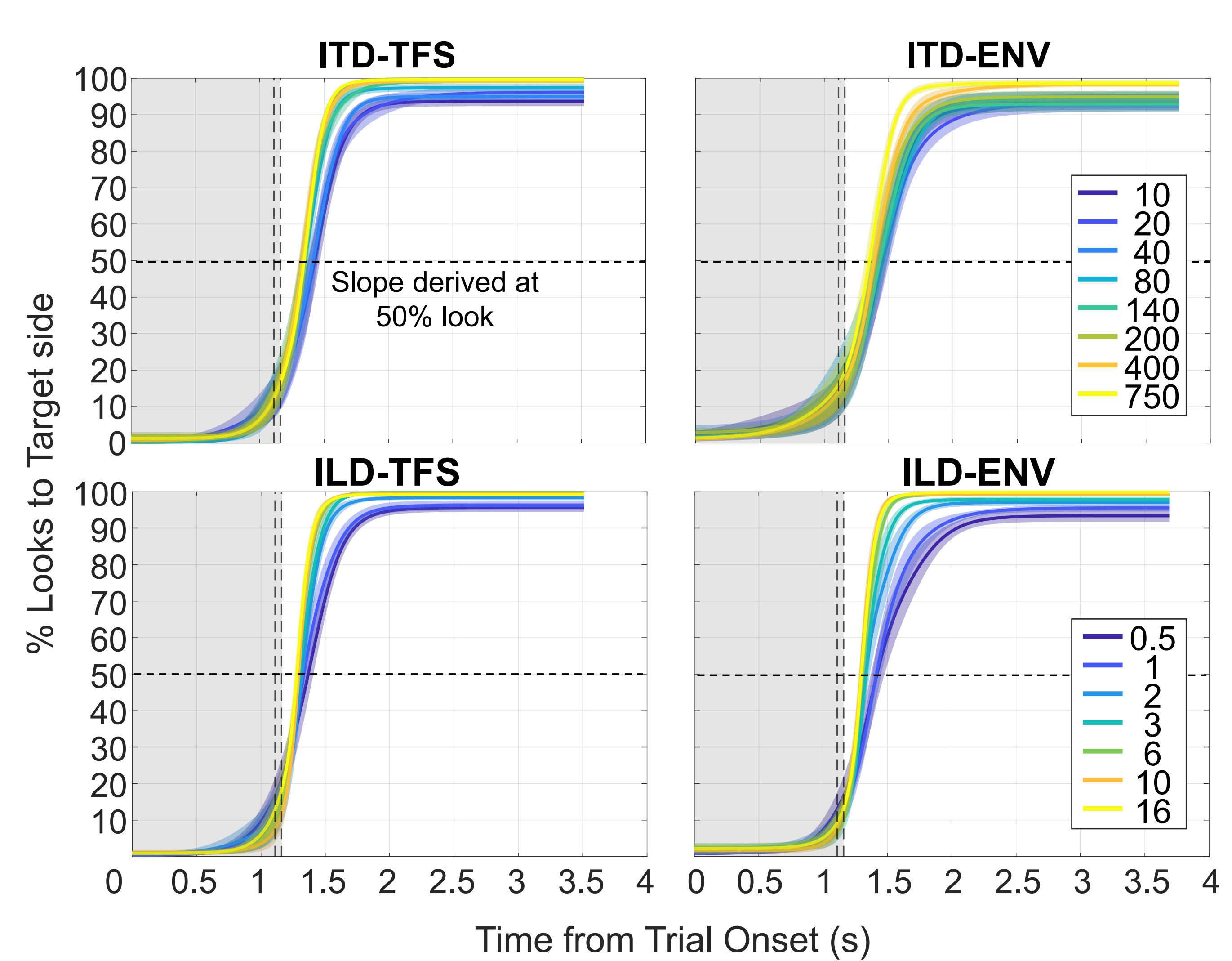


Fig. 4: Percent of looks to the target side over time for correct trials only, in which the final eye-gaze selection matched the target side. Eye-gaze trajectories were fit at the participant $\times$ cue-magnitude level using logistic regression (*psignifit*^{10,11}). Vertical dashed lines mark target onset and offset; shaded regions indicate stimulus timing.

Rate of evidence accumulation becomes faster with increasing cue magnitude, with TFS supporting faster commitment than ENV for ITD, while ILD slopes increase similarly across carriers.

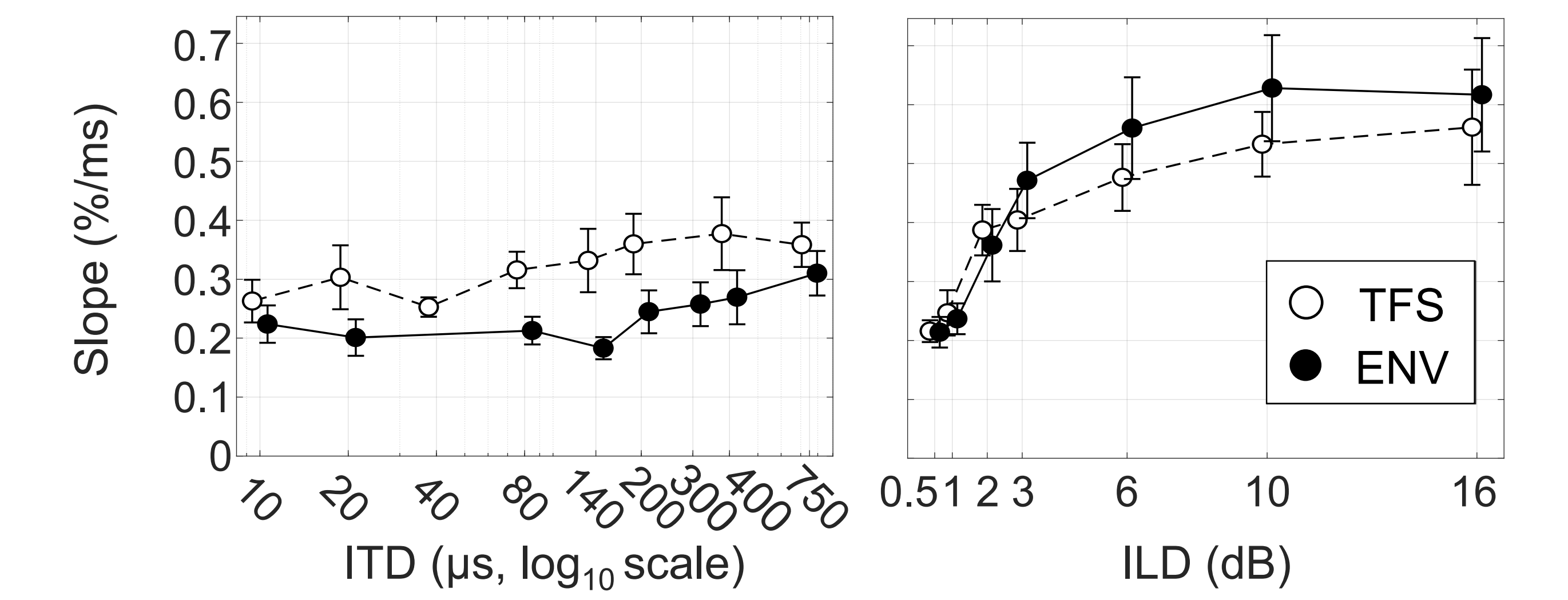


Fig. 5: Slope provides an approximation of the rate at which gaze commits to the target side, with **steeper slopes indicating earlier commitment to the perceptual decision**. Slope was defined as the derivative of the participant-level logistic fit at 50% looks to the target.

Number of saccades decreased with increasing cue magnitude, with more gaze shifts for ENV than TFS.

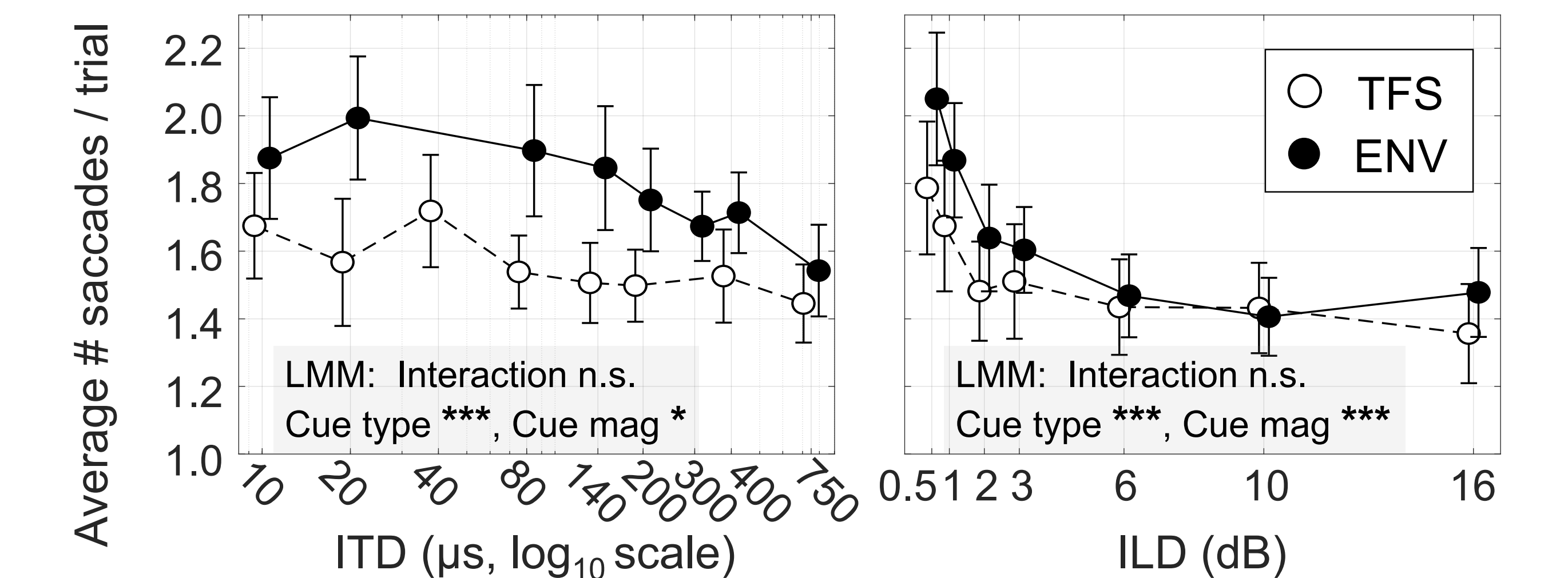


Fig. 6: Group means (± 1 SE) of saccade count per trial plotted as a function of ITD ($\log_{10}$ scale) and ILD magnitude for TFS and ENV conditions.

Time to first saccade decreased with increasing cue magnitude, with earlier first saccades for TFS than ENV.

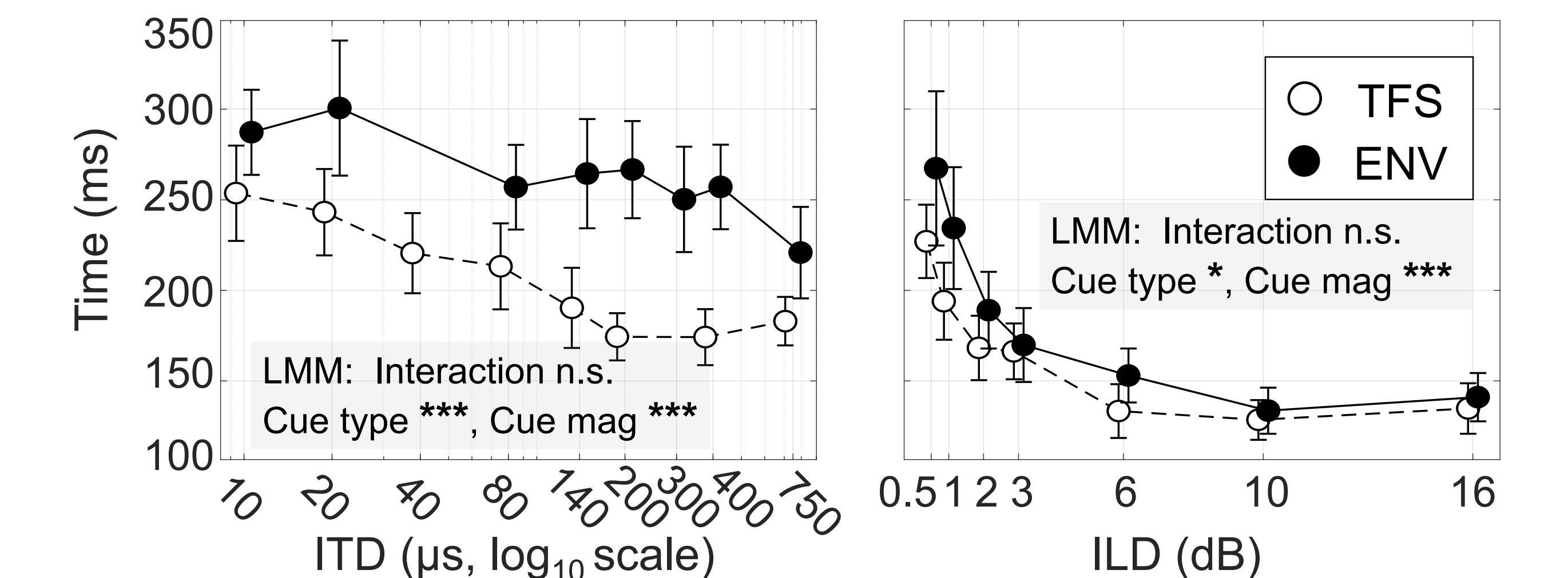


Fig. 7: Group means (± 1 SE) of response time from target stimulus to onset of the first saccade as a function of ITD ($\log_{10}$ scale) and ILD magnitude.

Conclusions

Binaural hearing ability are better for temporal fine structure (TFS) than envelope (ENV) cues for both ITDs and ILDs.

- For ITDs, TFS supports lower JNDs, steeper accumulation slopes, and faster saccadic responses than ENV, showing more sensitive and efficient binaural processing.
- For ILDs, decision dynamics are less differentiated between TFS and ENV, suggesting that cue salience for ENV and TFS is processed more similarly by the auditory system.

Eye-tracking provides perceptual information beyond JND and reveals cue-specific decision strategies. Results establish a **translational baseline** for interpreting how CI users access ITD and ILD cues under constrained binaural encoding.

Acknowledgement
Thank you to Amy K. Bramley for the assistance with participant testing!
This work was supported by NIH-NIDCD (R01DC020355 to RYL) and in part by a Core grant to the Waisman Center from the National Institute of Child Health and Human Development (P50HD105353).