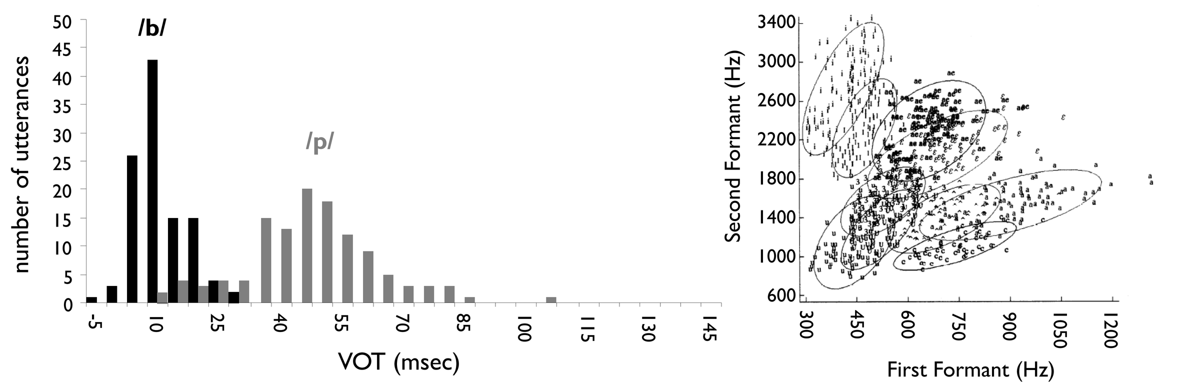
Sentence predictability modulates the auditory N1 event-related potential component

McCall E Sarrett, PhD* & Joseph C Toscano, PhD

[mccall.sarrett, joseph.toscano]@villanova.edu

Introduction

The acoustics of spoken language are highly variable:

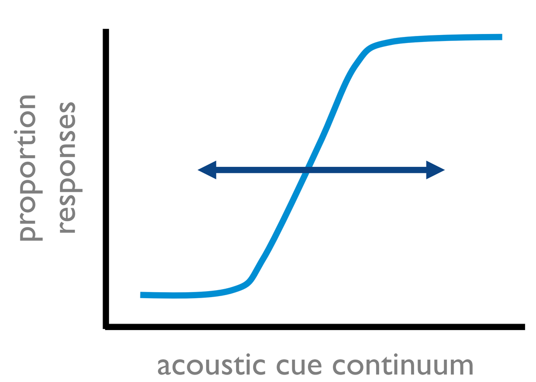


There is no one-to-one mapping between a given acoustic cue and the phoneme category it corresponds to

Psycholinguistic work has shown

(1) which acoustic dimensions are relevant to categorization
Allen, Miller, & DeSteno (2003); Hillenbrand, Clark, Getty, & Wheeler (1995)

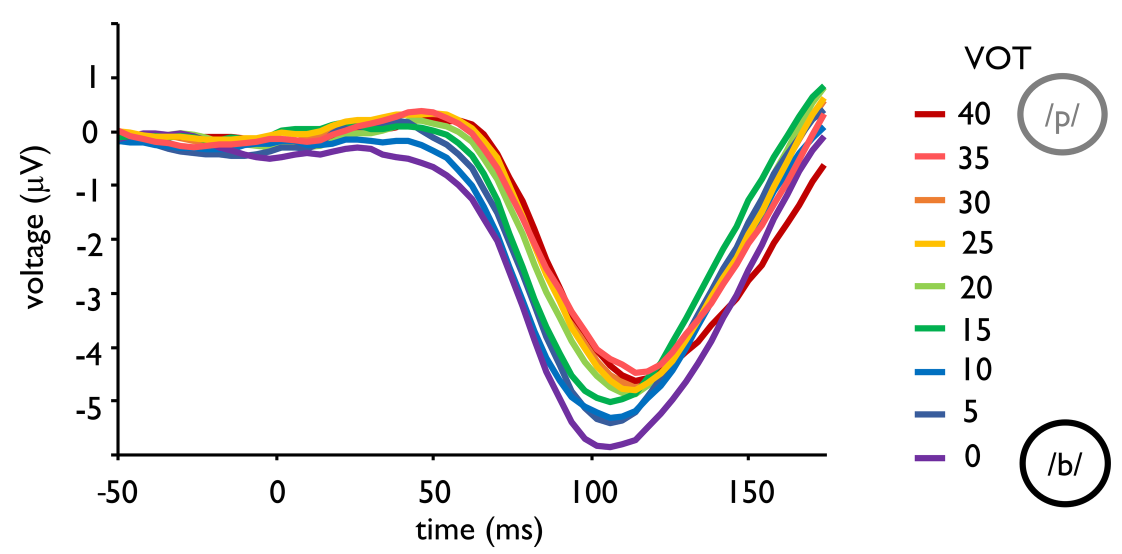
(2) that listeners use top-down expectations to shift categorization responses
Connine (1987), Ganong (1980), Miller, Green, & Schermer (1984)



Acoustic Encoding

More recent work has used the event-related potential (ERP) technique to address the neural mechanisms subserving these effects

This work suggests Voice Onset Time (VOT; the primary acoustic cue to voicing, which distinguishes sounds like /b/ and /p/) is encoded linearly at the N1 component

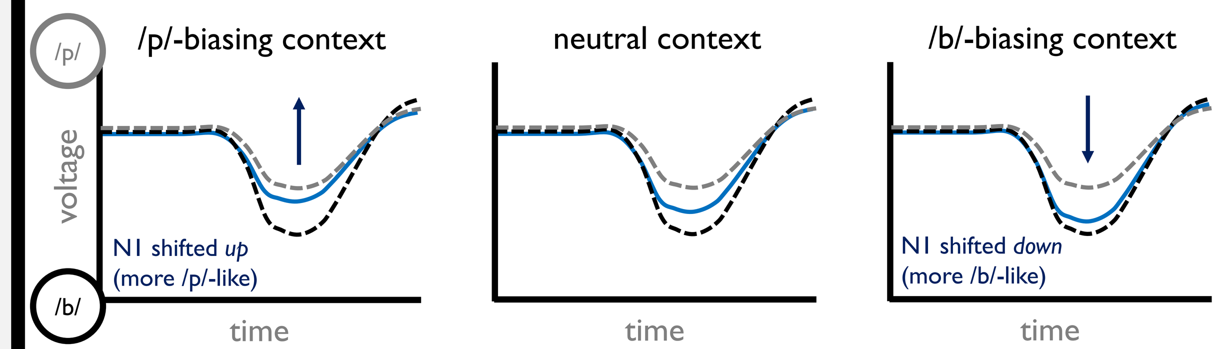


Shorter, more /b/-like VOTs yield a larger amplitude N1 component, and longer, more /p/-like VOTs yield a less negative N1 Toscano et al (2010)

→ How might top-down influence from a listener's contextual expectations affect acoustic encoding?

Contextual Influences

A series of studies have demonstrated that contextual expectations influence acoustic encoding when cues are ambiguous:



That is, listeners encode the same physical (acoustic) stimulus differentially, depending on their contextual expectations

This top-down influence of expectations has been shown with:

- orthographic primes — Getz & Toscano (2019)
- lexical status — Noe & Fischer-Baum (2019)
- picture primes — Sarrett & McMurray (in prep)
- sentence contexts — Sarrett, McMurray, & Kapnoula (2020)

It remains unknown exactly which factors may influence the strength and timing of feedback from higher-level information

If this influence is driven by activation of specific lexical items, then factors like cloze probability and entropy may modulate the strength and specificity of top-down activation

Current Study

The present study seeks to better characterize the nature of top-down influences during auditory sentence processing

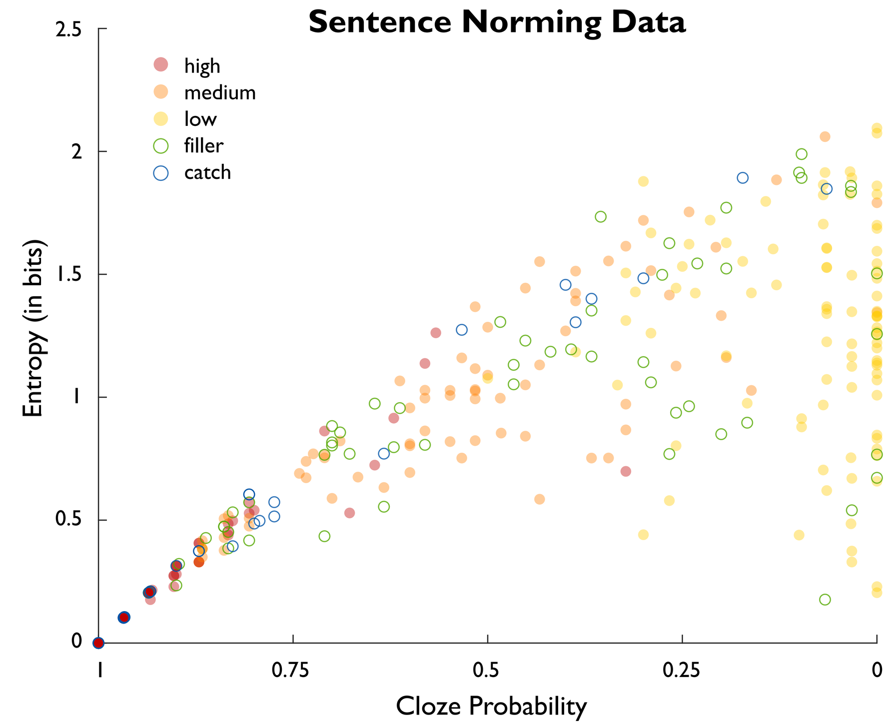
We manipulated how well carrier sentences predicted a sentence-final target word along two dimensions:

- cloze probability—how well the target word is predicted
- entropy—how many possible words could reasonably complete the sentence

High entropy
"The weather forecast called for extreme *cold*"
(alternatives: *heat, snow, storms, rain, wind, etc*)

Low entropy
"The child wouldn't swim because the water was too *cold*"
(alternatives: *deep, murky, etc*)

Sentence Norming Data



High cloze
"The number 3.14 is also known as *pi*"

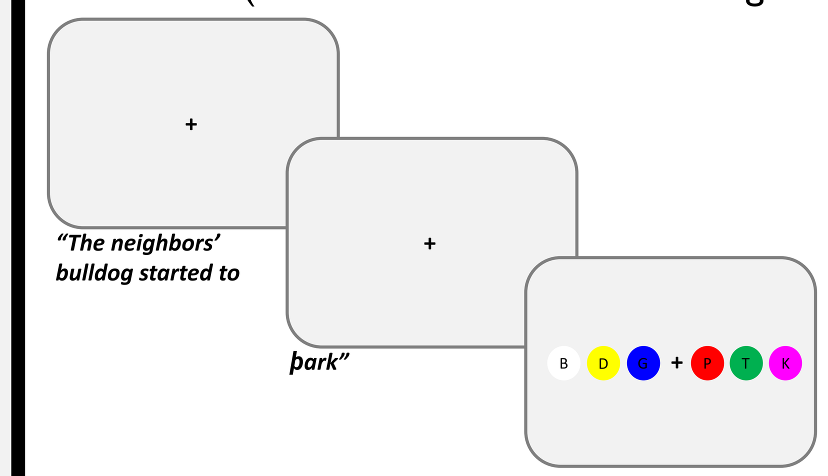
Medium cloze
"Grandma Betty told her she was as sweet as a *pie*"

Low cloze
"Wynton served himself some *pie*"

Method

We collected EEG data while participants heard sentences over earphones and responded via button press with which sound (/b,d,g,p,t,k/) a sentence-final target word started with

Target words were from minimal pairs (e.g. *beach/peach*), manipulated to be acoustically ambiguous, based on participant-specific phoneme boundaries (determined in an initial testing session)



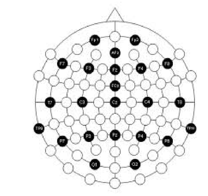
Experimental trials: presented with acoustically ambiguous target word
2 x contextual biases
(voiced or voiceless)
3 x cloze probability conditions
(high, medium, low)
35 x unique sentences per bias/cloze

Filler trials: presented with acoustically unambiguous target word
2 x contextual biases
30 x unique sentences per bias

Catch trials: no target word presented (fill-in-the-blank to best complete)
35 x unique sentences

EEG data preprocessing

Collected using 32-channel BrainVision actiChamp
Electrodes placed according to International 10-20 system
Recorded continuously & referenced online to left mastoid
Digitized to 500 Hz & re-referenced to average mastoid
Sequentially low- and high-pass filtered from 0.1 to 30 Hz
Epoched to the start of the sentence-final target word with a 200 ms baseline
N1 average taken at frontal electrodes (F3, Fz, F4) from 100-150 ms

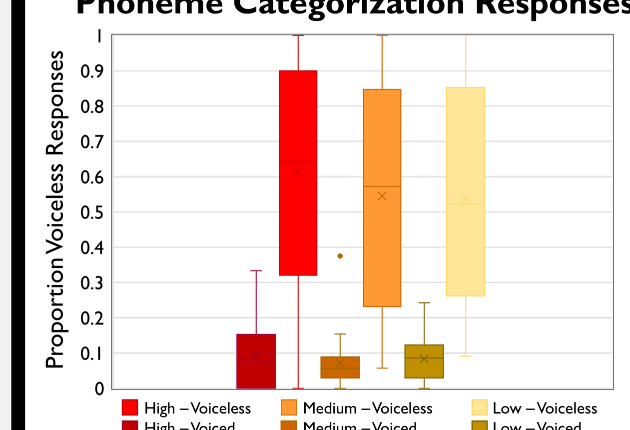


Results

Participants (N=22) showed sensitivity to both cloze probability and entropy in their behavioral responses

Statistical significance was evaluated using linear mixed effects modeling in R (R Core Team, 2022)

Phoneme Categorization Responses



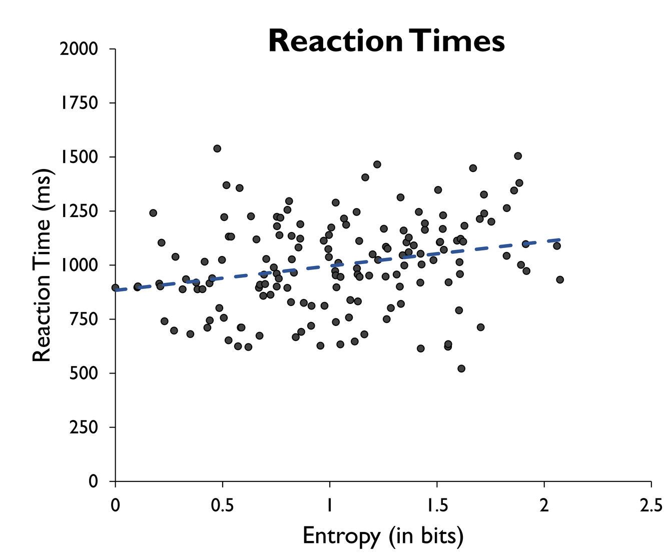
The bias of the preceding sentence significantly influenced how listeners categorized the ambiguous target word ($p < .001$)

The effect of bias appeared to be larger in sentences with a higher cloze probability, but this effect was only marginal (bias x cloze interaction, $p = .07$)

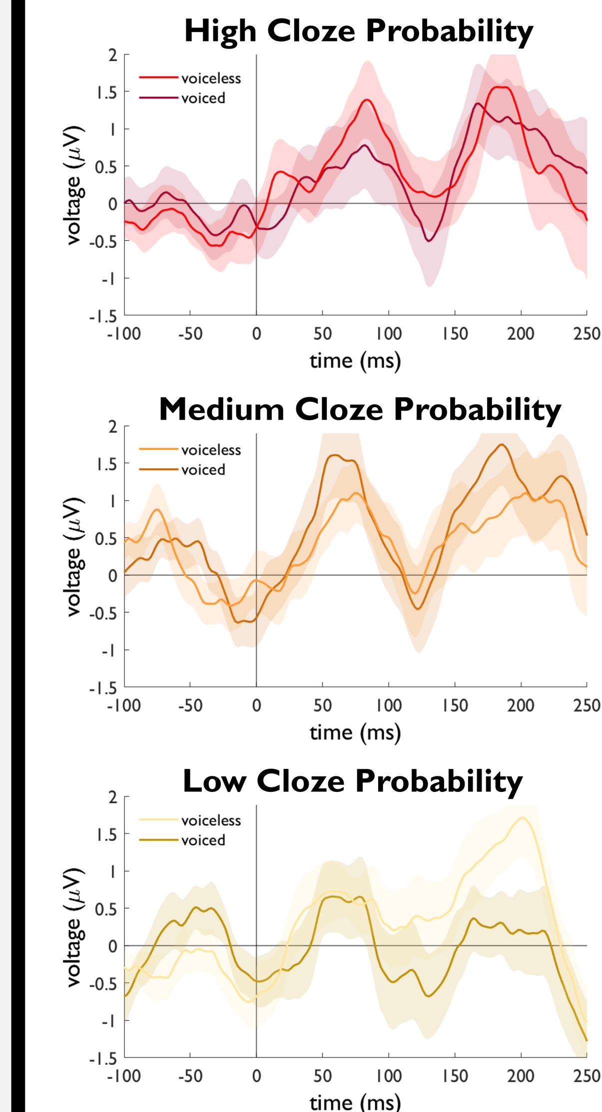
Sentence bias also significantly influenced participant reaction times (RTs; $p < .001$): Sentences that predicted a voiceless target word resulted in significantly longer RTs

Moreover, higher entropy sentences also resulted in significantly longer RTs ($p < .001$)

Reaction Times



EEG



The pattern of results in the ERPs remains somewhat unclear

We did not detect a significant main effect of bias on N1 mean amplitude, i.e. we did not find evidence that sentence context influences acoustic encoding, and we also did not find evidence of an interaction with cloze probability

It is possible (and likely) that the current data set is underpowered to detect these subtle effects

Another possibility is that sentence bias may influence perceptual representations later in the timecourse of processing, after the N1 peak (consistent with prior work) Sarrett et al (2020)

Future analyses will examine the full timecourse of processing

Discussion

First, we found that sentence bias significantly influences how listeners categorize an acoustically ambiguous target word: in a /b/-biasing context, participants are more likely to respond that they heard a /b/; in a /p/-biasing context, the converse is true

Moreover, the degree of this influence may interact with how predictive the preceding sentence is: More predictive sentences may result in a larger influence of sentence bias on participant responses

Second, entropy influences participant RTs: When a greater number of words could complete a sentence, participants respond more slowly

Finally, we may be underpowered to detect neurophysiological correlates, but the behavioral data validate our stimuli and our task

Further data collection is needed to interpret potential EEG effects

In conclusion, we hope this work will help to tease apart the influences of cloze probability, entropy, and sentence bias on acoustic encoding, and give insight into the neural mechanisms supporting dynamic interactions during speech perception

Acknowledgments

This material is based on upon work supported by the National Science Foundation under Grant No. 1945069.
We would like to thank Olivia Montañez for help with data pre-processing, and all the members of the WRAP Lab for help with data collection.