# EXPECTED UTILITY OF LATER CONTEXT MEDIATES MAINTENANCE OF SUBCATE-GORICAL INFORMATION

Spoken language understanding involves the integration of auditory cues with lexical context [1], including *later* context beyond the word boundary. Optimal integration with later context requires maintaining relevant information about auditory cues. Indeed, such integration has been found in previous work [2,3]. But, memory limits prevent listeners from maintaining such information indefinitely [4]. We hypothesize that the degree to which listeners maintain subcategorical information is guided by its *expected utility*: the more listeners expect the interpretation of subcategorical information to benefit from later context, the more likely they are to maintain that information in memory. Here, we test i) whether listeners can maintain subcategorical information, ii) if maintenance is the default, and iii) the informativity of later context in recent experience changes the degree to which listeners maintain subcategorical information.

Methods. 4 web-based studies (MTurk, N=39-117 subjects/study) have participants listen to sentences and respond whether they heard *tent* or *dent* (Table 1). We manipulate voice-onset time (VOT) to cover a continuum of /t/-/d/, and later context to bias towards "tent" (e.g., "campground"), "dent" (e.g., "fender"), or neither. Studies 1-3 always use the first two context conditions (i.e., the informative ones). **Studies 1 and 2** reduce the massive amount of repetition and unnaturally high levels of cue conflict common in previous work [1,2]. These changes avoid inducing experiment-specific strategies.

Figure 1

Informative Context — Uninformative Context

Context
- Dent-bias (red)
- Tent-bias (blue)

/t/ Response (log-odds) vs VOT (ms)

With these problems removed, we actually find significantly *larger* effects of subcategorical maintenance than previous work. This is expected if listeners typically maintain subcategorical information in memory during everyday language processing. **Study 3** tests this interpretation further. Using data from Studies 1-2 and an additional replication, we analyze whether the maintenance effect is present from the start of the experiment or only emerges over time. In all cases, we see a strong effect on the very first trial (e.g. Fig 2). This further suggests that subcategorical maintenance is typical of language use.

Figure 2 (Study 3): Context effect across trials

/t/ Response (log-odds) vs Trial
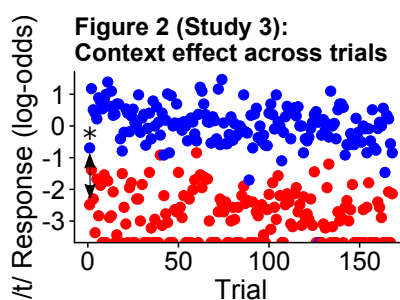
We hypothesize that this is due to typically high informativity of later context in natural language. **Study 4** tests this, asking whether maintenance of subcategorical information can be reduced if listeners expect upcoming context to be uninformative about the interpretation of earlier speech input. Subjects are divided into two exposure groups (Table 1). For the *Informative Context Group*, later context is *always* informative. For the *Uninformative Context Group*, it *never* is. In the later test phase, all subjects hear sentences with informative later context. We find that both groups maintained information in the test phase(ps< 0.001). Critically, the effect is reduced in the Uninformative Context group ($p = 0.008$; Fig 1). **Conclusions.** We find that listeners maintain subcategorical information about the speech input in memory for integration with later context. This seems to be a default strategy, in line with the typically high informativity of later context in natural language use. When this informativity is removed, listeners reduce information maintenance. Together, these results suggest that listeners use the statistics of the input to guide strategies for memory allocation during real-time processing. [1] McMurray et al 2009 *JML* [2] Brown-Schmidt & Toscano (2017) *LCN* [3] Connine et al (1991) *JML* [4] Christiansen & Chater 2015 *BBS*
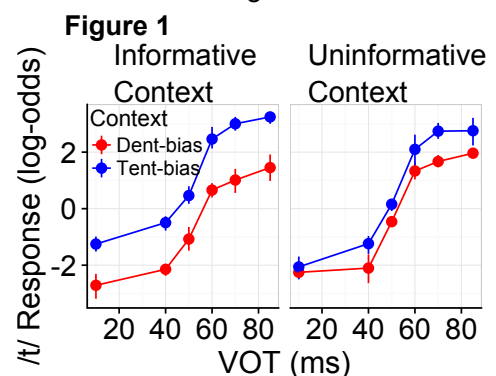
Group	Exposure Phase	Test Phase
Informative	...[t/d]ent was noticed in the campground/fender...	...[t/d]ent in the forest...
Uninformative	...[t/d]ent was noticed...	...[t/d]ent in the forest...