Using nonsense word detection to investigate sentence processing

SHORT ABSTRACT

In this study, participants read sentences and decided whether they contained a nonsense word. Performance was significantly affected by syntactic structure which, in a task where processing was not required, suggests syntactic parsing is an automatic reflex. Performance was also affected by nonsense word position (participants were slower for later positions). There was no interaction between nonsense word position and sentence type, which hints at a possible divide between word-recognition and sentence-comprehension systems.

LONG ABSTRACT

Introduction. Human sentence processing is robust in the face of noise. Not only can we read the sentence, "*John ate the blargmur,*" we can also very easily infer that "*blargmur*" likely refers to something that can be eaten. In previous work, participants were presented with whole sentences and asked to identify whether it contained a nonsense word (Stromswold, et al., 1996). It was found that performance on nonsense word detection was affected by the syntactic structure of the sentences. The significant structure effect in a task where sentence processing is not required suggests that syntactic parsing is an automatic reflex.

In this study, we attempt to replicate the findings of Stromswold, et al. (1996) using a wider range of sentence structures. We further attempt to examine the relationship between the cost of detecting a nonsense word and its syntactic position.

Method. We constructed 60 quadruplets of relative clause (RC) sentences, and each of the 4 sentences in a group contained the same pair of verbs and triplet of nouns. (To ensure plausibility, all nouns were animate.) An example quadruplet is presented below:

1) The actor who impressed the critic humiliated the director.
2) The actor who the critic impressed humiliated the director.
3) The director humiliated the actor who impressed the critic.
4) The director humiliated the actor who the critic impressed.

Each participant saw only one sentence from each quadruplet. In 48 of the RC sentences, a single lexical word (noun or verb) was replaced with an orthographically and phonologically plausible nonsense word (with lexical words 2, 3, 4, and 5 each being replaced in 12 RCs respectively), and 12 RC sentences contained only real words. Furthermore, within participants, RC type was orthogonal to position of nonsense word (i.e. 4 RC types X 5 positions). In addition to the 60 RCs, we had had 180 filler sentences, half of which contained a nonsense word replacing one of the lexical words. Thus, overall 57.5% sentences contained nonsense words.

Twenty-six participants were presented with whole sentences one at a time and asked to decide if they contained a nonsense word. They responded by pressing one of two keys to signal a presence or absence of a nonsense word and their response times were recorded.

Results. We analyzed the target sentences and only those trials which the participant got correct (95%). A 4 (Type) X 5 (Nonsense word position) ANOVA of response times revealed a significant main effect of sentence type ($F(3,25)=5.12$; $p<.005$) with participants being slowest on object-extracted center-embedded sentences (see example 2 above). There was also a significant main effect of nonsense word position ($F(4,25)=8.19$; $p<.001$). We found that participants were slower for later nonsense word positions (fourth or fifth lexical word) than for early nonsense word positions (second or third lexical word). There was no difference between the time for later nonsense word positions and sentences not containing a nonsense word. The interaction was not significant ($F(12,25)=1.22$; $p=.27$).

We further analyzed only those target sentences that contained nonsense words using a 4 X 4 ANOVA. Once again, there was a significant main effect of sentence type ($F(3,25)=3.77$; $p<.05$) and nonsense word position ($F(3,25)=19.56$; $p<.001$), and the interaction was again not significant ($F(9,25)=1.22$; $p=.28$).

Discussion. The main effect of sentence type rules out a simple model where linear scanning of words is performed to detect the nonsense word. A more likely model is one where the sentence is processed till the nonsense word. That also explains why participants were faster for the early positions than for the later positions. In other words, participants maybe did scan word-by-word starting from the left, but also processed the sentence along the way, perhaps incidentally.

The fact that there is no interaction between the sentence type and the nonsense word position suggests that morphological cost (i.e. the cost of deciding if the word is a nonsense word) is unaffected by the position of the word in the syntactic structure. This result is especially interesting in light of modern expectation-based models of sentence processing (e.g. Hale, 2003; Levy, 2008). Those models combine morphological processing cost with syntactic processing cost while computing the processing cost for a word in a sentence. Thus, they make the inherent prediction that morphological processing cost is affected by syntactic position, which contradicts our findings.

In conclusion, the findings of our study suggest sentence processing is an automatic reflex as proposed by Fodor (1983). Our findings also align with Fodor's (1983) hypothesis that word recognition systems are independent of sentence comprehension systems.

References.

1) Stromswold, K., Caplan, D., Alpert, N., and Rauch, S. (1996) Localization of syntactic comprehension by positron emission tomography. *Brain and Language*. 52:452-473.
2) Hale, J. (2003). The information conveyed by words in sentences. *Journal of Psycholinguistic Research,* 32:101-123.
3) Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106:1126-1177.
4) Fodor, J. (1983). Modularity of mind. Cambridge, MA: MIT Press.