Searching for importance: focus facilitates memory for words in English

Heather Kember[1,2] *, Jiyoun Choi*[2,3]*, Jenny Yu*[1,2]

[1]The MARCS Institute, Western Sydney University, Australia
[2]ARC Centre of Excellence for the Dynamics of Language, Australia
[3]Hanyang Phonetics and Psycholinguistics Lab, Hanyang University, Korea
h.kember@westernsydney.edu.au, jiychoi@hanyang.ac.kr, jenny.yu@westernsydney.edu.au

Abstract

Focus, both prosodic and syntactic, conveys processing advantages in English. We compared recognition memory for words that received syntactic, prosodic, syntactic + prosodic, or no focus. English speakers listened to blocks of sentences, then were presented with words and asked if they remembered them. Words with focus were recognised faster than words without focus, and there was an additive effect, such that words with both focus types were responded to even faster than words that had only prosodic, or syntactic focus. These findings suggest a common mechanism for drawing attention towards focused constituents within an utterance.
Index Terms: prosody; focus; memory; speech processing

1. Introduction

Listeners will focus attention on crucial components within an utterance that constitutes newly introduced information [1]. This is considered to be a language universal. In English, marking of focus can be achieved through both syntactic and prosodic means: prosodic focus is marked with a pitch accent [2], and syntactic focus can be achieved via cleft structures such as It-cleft (e.g. it was the…), or There-insertion (e.g. There was this…). Both of these focus-producing devices have been shown to provide processing advantages. For example, words marked by syntactic focus are remembered better compared to words without syntactic focus [3, 4]. As for prosodic focus, sounds within salient words are recognised more rapidly in comparison to sounds within non-salient words [5]. This processing benefit conveyed by prosody is not solely due to the increased acoustic saliency – when the focused words themselves are removed and replaced with neutrally produced words, participants are still faster at detecting phonemes at the prosodically salient position [6, 7]. This suggests that the listener entrains to the utterance prosody, predicting where the focused word will be located, and directs attention to that position.

While these processing benefits for focused constituents within utterances have been documented in English and other languages [8, 9], it is an open question what the mechanism is that is driving this facilitatory processing of focused information, and indeed whether the mechanisms are the same for the different types of focus. [10] suggests a central mechanism, which is a search for the most semantically central part of a speaker's message, which can be provided by any of the focus types. Their paper used semantic focus only however, and not other focus types. Another study [4] presented sentences visually, and tested recognition memory for words that were either identical, phonologically related, or semantically related to target words within the sentence. Overall, words that had syntactic focus were consistently remembered better than words without syntactic focus. Syntactic focus facilitated remembering the identical targets and phonological targets (only with a memory delay) but memory for semantically related targets was not improved. The authors suggested that the memory trace created by syntactic focus is specifically for the identity of the word itself and phonological information, rather than the general meaning of the word. Taken together, these previous studies suggest a central mechanism for seeking the semantically important parts of an utterance, and facilitating processing of the specific word identity.

It is also unclear how syntactic and prosodic focus interacts with each other, as all studies mentioned up to this point, only assessed a single type of focus in an utterance. In [6], the authors compared the effect of semantic and prosodic focus in a phoneme monitoring task. They found that both focus types facilitated phoneme detection over non-focused words, but that there was also an interaction between the two, arguing for a common cause. Given syntactic focus also serves this same purpose as semantic focus (i.e. to alert the listener to important information), we suggest that we will also see this additive effect of focus type in our study.

In this paper we present a pair of experiments investigating the role that syntactic and prosodic focus plays in sentence processing in English. In Experiment 1, we tested recognition memory for words that had been focused (either syntactically, prosodically, or both) in comparison to unfocused words. Experiment 2 replicated Experiment 1, but also manipulated the target words presented, such that targets were either identical or related semantically or phonologically.

2. Experiment 1

Both experiments tested recognition memory for words that had been presented in one of four conditions: no focus, prosodic focus, syntactic focus, or prosodic + syntactic focus. We recorded recognition accuracy and reaction time.

2.1. Method

2.1.1. Participants

32 native speakers of Australian English were recruited for participation (M_{age} = 29.63, SD_{age} = 13.32). Some participants learned other languages (no simultaneous bilinguals). Participants had no self-reported speech, reading, or hearing issues.

2.1.2. Stimuli

140 experimental sentences were constructed. Each sentence contained two possible high frequency target words, one early, and the other later in the sentence. One target had syntactic

focus and the other did not have syntactic focus. No target words were repeated across sentences. We used either the "It was the…" or "There was this…" syntactic cleft structures meaning all syntactically focused words were the first target.

A female, native speaker of Australian English recorded two versions of each sentence. In version one, prosodic focus was placed on the first target, and in version two it was placed on the second target. This generated four different experimental conditions (two per sentence). For example:

1) There was this **medal** (PS) displayed on the stand (NF) in the living room
2) There was this medal (ST) displayed on the **stand** (PR) in the living room

(NF = no focus, PR = prosodic focus, ST = syntactic focus, PS = prosodic + syntactic focus. Prosodic focus indicated by bold type, target word indicated by underline). To encourage the speaker to place focus on the correct target, two questions were written for each sentence and read to the speaker during recording. Participants only heard one version of each sentence to ensure any effect was not simply due to repeated presentations of the same words. To control for the confound of syntactically focused words appearing as target 1, we also created a set of 20 control sentences that also had early and late targets but without syntactic or prosodic focus.

2.1.3. Procedure

Sentences were pseudo-randomised so that all sentence types were evenly dispersed across the experiment. Sentences were then split into blocks of 10, with each block containing seven experimental sentences, two filler sentences, and one control sentence. Four different iterations of the experiment were created, varying the order of presentation of stimuli, as well as the sentence versions and targets. For example, in experiment version one the target word "stand" was tested in the no focus condition and in version two, it was tested in the prosodic focus condition. In versions three and four, the target word "medal" was tested in syntactic, and prosodic + syntactic conditions respectively. Participants were randomly assigned to one of the four experimental versions, with numbers equal in each of the groups.

Participants were tested individually in a sound attenuated room in a single session lasting approximately 45 minutes. E-Prime was used to present the stimuli and record participant responses. Noise-attenuating Sennheiser HD 280 Pro headphones were fitted and adjusted to achieve comfortably audible volume levels for each participant individually.

The experiment alternated between blocks of sentences and recognition memory tests. To encourage attentiveness, participants pressed the spacebar to play each sentence. After the 10 sentences were played, instructions appeared on screen telling participants to prepare for the recognition memory test. Participants were presented with 10 words sequentially on screen (one from each sentence) and asked to indicate whether or not they remembered hearing it: the letter "M" for "yes" and the letter "Z" for "no". This was reversed for left handed participants (N=1) such that the dominant hand always indicated a positive response. Target words remained on screen for five seconds, or until the participant responded. Participants were instructed to respond as quickly and accurately as possible. Order of target words in the recognition memory test was identical to the order of sentence presentation in to keep the time delay between presentation and their respective targets as consistent as possible.

Participants were permitted to take short breaks after each recognition memory test, before continuing with the next block of sentences. A single practice block was given prior to the commencement of the experiment to ensure participants were comfortable with the method. The experimenter was present for this time to answer questions.

2.2. Results

2.2.1. Acoustic analyses of stimuli

In order to validate the location of prosodic focus on the target words, for experimental and control sentences, target words were segmented and annotated based on inspection of the waveform and spectrogram in PRAAT [11]. For each target, duration, pitch (F0 mean, min, and max), and intensity (mean) values were extracted. Table 1 shows mean values for each of these measures as a function of condition.

Table 1. *Mean acoustic values for target words in each focus condition.*

Condition	NF	PR	ST	PS
Duration	305.971	433.85	297.121	408.743
f0_min	157.336	168.707	193.971	183.429
f0_mean	211.264	233.014	236.257	249.507
f0_max	326.979	420.386	355.807	434.229
int_mean	51.736	55.343	54.836	56.886

2.2.2. Recognition accuracy scores

Recognition accuracy scores were aggregated within condition, to yield a proportion of correct responses for each participant. Figure 1 displays the mean response accuracy as a function of condition. We used a within-subjects ANOVA with focus conditions as the independent variable for analysis.

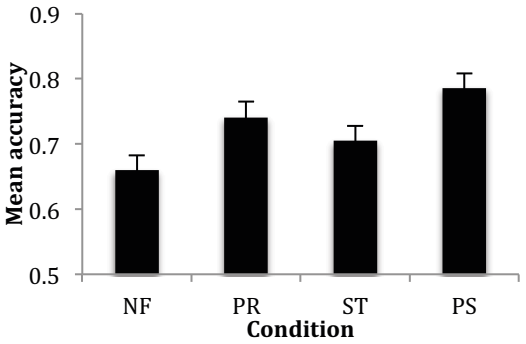

Figure 1. *Mean accuracy as a function of focus condition. Error bars represent standard error of the mean.*

The overall model was significant, $F(3,31) = 16.69$, $p<.001$. Words with focus (any type) were recognised more accurately than words without focus, $F(1,31) = 31.53$, $p<.001$. Words with prosodic focus were more likely to be recognised than words with no focus, $F(1,31) = 16.39$, $p<.001$. Words with both prosodic and syntactic focus were recognised more accurately than words in the syntactic focus alone condition, $F(1,30) = 18.21$, $p<.001$.

2.2.3. Reaction time results

Reaction times were aggregated to create a mean reaction time per condition dependent variable. We used a within subjects ANOVA with custom contrasts with condition as an independent variable. Figure 2 displays mean reaction time as

a function of condition. The overall model was significant, $F(3,31) = 11.36$, $p<.001$. Words with focus (any type) were recognised significantly faster than words without focus, $F(1,31) = 12.12$, $p=.001$. Words with prosodic focus were recognised significantly faster than words in the no focus condition, $F(1,31) = 5.12$, $p=.03$. Words with some kind of syntactic focus were recognised significantly faster than words with prosodic focus alone, $F(1,31) = 6.09$, $p=.03$. Words with both prosodic and syntactic focus were recognised significantly faster than words with only syntactic focus, $F(1,31) = 14.40$, $p=.001$.

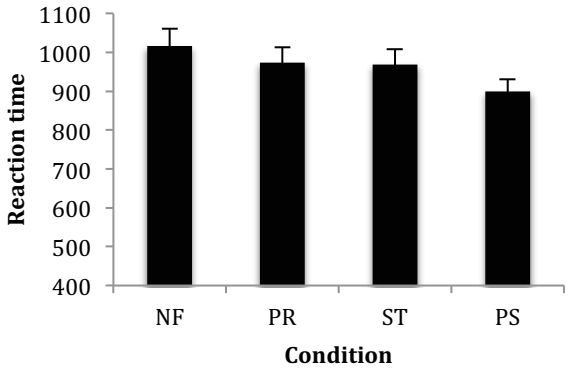

Figure 2. *Mean reaction times as a function of focus condition. Error bars represent standard error of the mean.*

2.2.4. Comparing experimental and control sentences

We further compared reaction time for words from experimental sentences to the control sentences. First we compared words that appeared as early targets (syntactic focus, prosodic + syntactic) with the early targets in the control sentences. Using a within subjects ANOVA with follow up Bonferroni adjusted pairwise comparisons, we showed the overall model was significant, $F(2,31) = 18.19$, $p<.001$, but crucially, both the experimental conditions were responded to significantly faster than the control targets, (M difference = 111.70, $p=.002$ for ST, and M difference = 178.99, $p<.001$ for PS). We then used a similar within subjects ANOVA to compare words that appeared as late targets (no focus, prosodic focus) with early control targets. The overall ANOVA was non significant, $F(2,31) = 2.32$, $p=.11$, however crucially, prosodically focused words were responded to significantly faster than late control targets (M difference = 52.41, $p=.05$).

2.3. Discussion

Our results showed that words with focus in English (whether conveyed by syntax or prosody) are more likely to be remembered, and are recognised faster than words without focus. Within the focus types, it seemed that words marked with some kind of prosodic focus received more of a facilitatory effect than words with syntactic focus. We also found an additive effect, such that words with both prosodic and syntactic focus were even more likely to be recognised, than words with either prosodic or syntactic focus alone. The reaction time results tell a somewhat different story, in that words with syntactic focus were recognised significantly faster than words with prosodic focus, but as with the recognition accuracy scores, there was an additive effect, such that words with both syntactic and prosodic focus were reacted to faster than words with a single focus type.

By comparing reaction times for target words from the control sentences with target words from the experimental sentences, we were able to determine that our findings were not simply an effect of regency; rather, word processing was facilitated by our manipulations of focus.

Our findings support previous work showing an additive effect for focus types. We also supported the hypothesis of a central search mechanism that seeks important information within an utterance to facilitate understanding [10].

3. Experiment 2

In Experiment 2, we investigated the underlying mechanisms for these processing advantages. We replicated Experiment 1, but changed the target words presented to participants, such that some were identical targets, and others were related either semantically, or phonologically.

3.1. Method

3.1.1. Participants

32 native speakers of Australian English were recruited for participation ($M_{age} = 21.19$, $SD_{age} = 3.75$).

3.1.2. Stimuli and procedure

The same sentences were used as Experiment 1. For 46 experimental sentences the same identical targets were used as above. For 47 sentences a phonologically related target word was selected (e.g. phonologically related target *shepherd* in place of the original *sheriff*). For the remaining 47 sentences a semantically related target was selected using the [12] database (e.g. semantically related target *tiger* in place of the original *lion*). The procedure was identical to Experiment 1.

3.2. Results

3.2.1. Recognition accuracy scores

As above, the recognition accuracy scores were aggregated to create a proportion correct dependent variable. We used a within subjects ANOVA with focus condition and target type as the independent variables.

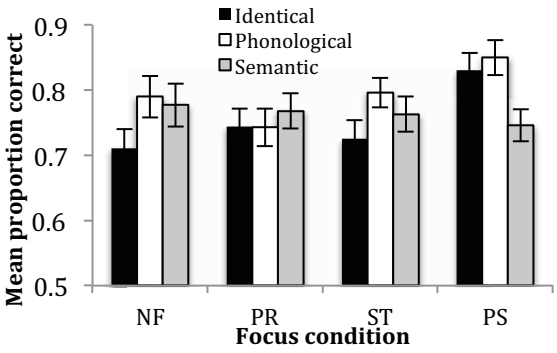

Figure 3. *The interaction between focus condition and target type. Error bars represent standard error of the mean.*

The main effect of focus condition was significant, $F(3,31) = 4.51$, $p=.005$, however the main effect of target type was not significant, $F(2,31) = .99$, $p=.38$. There was however, an interaction between focus condition and target type, $F(6,31) = 3.56$, $p=.002$ (displayed in Figure 3). The pattern of results for

the identical targets mirrored Experiment 1. For semantic targets, response accuracy was consistent, independent of the focus condition, and for phonological targets, they were recognised more accurately, when in the prosodic + syntactic condition versus the other focus conditions.

3.2.2. Reaction time results

We used a within subjects ANOVA with focus condition and target types as the independent variables. The overall effect of focus condition was significant, $F(3,31) = .014$, as was the main effect of target type, $F(2,31) = .022$. There was no significant interaction between these two variables, $F(6,31) = 1.70$, $p=.124$. The effects of focus condition replicated those described above, therefore here we only explain the effect of target type. Overall, participants responded fastest to identical targets, followed by phonological targets, followed by semantic targets. Figure 4 shows the different target types as a function of focus condition.

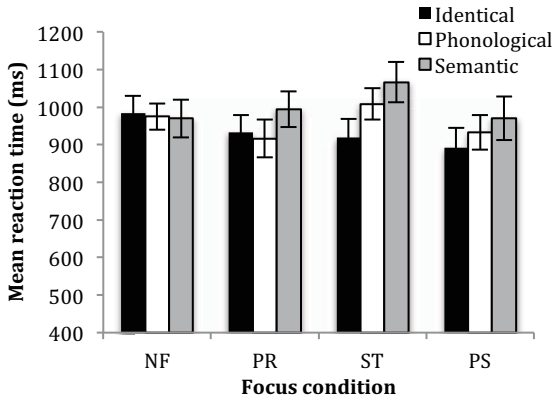

Figure 4. *Effect of focus condition and target type on reaction time. Error bars represent standard error of the mean.*

3.3. Discussion

The findings of Experiment 2, firstly replicate the effects of focus condition shown in Experiment 1. As for target type, when looking at the recognition memory results, there was no main effect of target type, but there was an interaction between target type and focus condition. For reaction time there was a main effect of target type, but no interaction. The findings together suggest that overall, identical targets are more likely to be recognised, and recognised faster, in a pattern consistent with Experiment 1. Phonological information was activated and responded to more accurately and faster in some focus conditions, but participants' ability to accurately respond to semantically related words did not seem to be related to which focus condition it was in. This suggests activation of a general semantic representation for the entire utterance, but that words with focus are attended to in greater detail via encoding of the specific word identity and phonological information.

4. General discussion

Taken together, these two experiments show that both focus types confer processing advantages in sentence processing in English and that this processing benefit is a specific memory trace of the target word itself, and its phonological information, rather than a general semantic representation.

This is the first study to our knowledge comparing the effect of syntactic and prosodic focus on processing in a single study in English. Consistent with [6], who showed an additive effect for semantic and prosodic focus processing, we also show an additive effect for syntactic and prosodic processing.

Our previous work in Korean also showed that while both focus types provided processing advantages, syntactic focus had more of an effect than prosodic focus on recognition memory [9]. Further cross-linguistic study is warranted.

5. Conclusions

These results contribute to a growing body of work assessing the role that focus plays in sentence processing. While both focus types do support efficient processing of words in English, it seems this effect is magnified when they work in concert. This lends weight to the hypothesis that it is a shared underlying mechanism that seeks new or important information within an utterance and directs attention to it.

6. Acknowledgements

This research was conducted with support from the ARC Centre of Excellence for the Dynamics of Language (project ID: CE140100041). The second author was also supported by a National Research Foundation of Korea Grant (2016S1A5B5A01025371).

7. References

[1] M. Krifka, "Basic notions of information structure," *Acta Linguistica Hungarica,* vol. 55, pp. 243-276, 2008.

[2] D. Bolinger, "Intonation across languages," *Universals of human language,* vol. 2, pp. 471-524, 1978.

[3] S. L. Birch, J. E. Albrecht, and J. L. Myers, "Syntactic focusing structures influence discourse processing," *Discourse Processes,* vol. 30, pp. 285-304, 2000.

[4] S. L. Birch and S. M. Garnsey, "The Effect of Focus on Memory for Words in Sentences," *Journal of Memory and Language,* vol. 34, pp. 232-267, 4// 1995.

[5] A. Cutler and D. J. Foss, "On the Role of Sentence Stress in Sentence Processing," *Language and Speech,* vol. 20, pp. 1-10, January 1, 1977 1977.

[6] E. Akker and A. Cutler, "Prosodic cues to semantic structure in native and nonnative listening," *Bilingualism: Language and Cognition,* vol. 6, pp. 81-96, 2003.

[7] A. Cutler, "Phoneme-monitoring reaction time as a function of preceding intonation contour," *Perception & Psychophysics,* vol. 20, pp. 55-60, 1976.

[8] Y.-C. Lee, B. Wang, S. Chen, M. Adda-Decker, A. Amelot, S. Nambu*, et al.*, "A crosslinguistic study of prosodic focus," *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing,* 2015.

[9] H. Kember, J. Choi, and A. Cutler, "Processing advantages for focused words in Korean," In J. Barnes, A. Brugos, S. Shattuck-Hufnagel, & N. Veilleux (Eds.), *Proceedings of Speech Prosody 2016* (pp. 702-705).

[10] A. Cutler and J. A. Fodor, "Semantic focus and sentence comprehension," *Cognition,* vol. 7, pp. 49-59, 1979.

[11] P. Boersma and D. Weenink. (2014, 2 January). *Praat: doing phonetics by computer (5.3.62 ed.)*.

[12] D. L. Nelson, McEvoy, C. L., & Schreiber, T. A., "The University of South Florida word association, rhyme, and word fragment norms," http://www.usf.edu/FreeAssociation/*,* 1998.