

ENGLISH LISTENERS' KNOWLEDGE OF THE BROAD VERSUS NARROW FOCUS CONTRAST

Jason B. Bishop

University of California, Los Angeles, USA

j.bishop@ucla.edu

ABSTRACT

The present study examined English listeners' knowledge of how the size of the focus constituent is expressed prosodically. Thirty English-speaking listeners participated in a prominence-rating experiment in which they heard the same production of a simple SVO answer sentence in different focus (i.e., different question) contexts. It was found that when the object was narrowly focused, it was heard as more prominent (and the preceding verb as less prominent) than when that same object was part of broader VP or Sentence foci. We interpret the findings as reflecting listeners' knowledge of speakers' productions of the information structure contrast.

Keywords: focus, information structure, prosodic prominence, intonation, "top-down" perception

1. INTRODUCTION

This study considers English listeners' knowledge of the prosodic realization of an information structural contrast: the size of the focus constituent. This contrast is shown below for two potentially different focus *types*, WH-focus (1) and what is often called 'corrective' focus (2):

- (1) a. *What happened?*
 b. *What did you do?*
 c. *What did you buy?*
 d. I bought a motorcycle.
- (2) a. *Why's your wife mad...
 because the roof's leaking again?*
 b. *Why's your wife mad...
 because you lost your job?*
 c. *Why's your wife mad...
 because you bought a car?*
 d. No...because I bought a motorcycle.

We will refer to cases as (1d) in the context of (1c) and (2d) in the context of (2c) as cases of 'narrow focus' on the object, compared with the other question contexts, which require an answer

with broader focus on either the entire verb phrase (1b and 2b) or the entire sentence (1a and 2a).

Prosodically, the realization of this contrast is said to be ambiguous, since the location of the answer's nuclear pitch accent does not vary along with the size of the focus in English SVO constructions; in each case, it falls on the final object. However, a number of phonetic studies in recent years have examined speakers' productions of sentences with different foci, and differences are often found [1, 2, 5, 9, 10, 11, 13]. These differences can be generalized as follows. Although the object always bears the nuclear pitch accent when it is narrowly focused, speakers tend to pronounce it as especially prominent, and prenuclear material as non-prominent. Conversely, when focus is on the entire sentence (broad Sentence focus) or the verb phrase (broad VP focus), the opposite pattern is found: although the nuclear accent is still on the object, the phonetic prominence of that accent is lower, and the prominence of prenuclear materials is higher. This seems to suggest that speakers manipulate the relative prominence of the nuclear accent to express focus size, and suppressing prenuclear prominence is one way to do this. Another generalization relates to across-speaker differences: some speakers primarily manipulate prenuclear prominence, some speakers primarily manipulate nuclear prominence, and other speakers do both. Thus speakers appear to have knowledge that the size of the focus relates to the prominence of the nuclear accent, and the role of prenuclear prominence is to modulate that nuclear accent's prominence. This has also been noted by [6], who suggests that the interpretation of focus size is inversely related to the amount that a given level of prominence exceeds the listener's baseline (i.e., non-focus) expectations, such that when a word is more prominent than expected, a narrow focus reading of a word is more likely.

However, it is unclear to what extent this pattern is a part the *listener's* knowledge of the contrast, as relatively little work has been done to

assess their use of prosody to disambiguate. One attempt at probing listeners' knowledge has involved presenting them with questions such as the those in (1), followed by answers that either have or do not have a prenuclear accent, and collecting appropriateness judgments. Investigation making use of that technique has generally found that listeners do not show a preference for the presence or absence of prenuclear accents [3, 8, 12] – suggesting the difference speakers produce might not be a salient part of their expectations.

Note, however, the possibility that these methods might not be optimal for eliciting the use of small phonetic differences like the ones reported. It may be, for example, that speakers (in the materials used in various experiments) do not encode robust phonetic cues to the contrast when the context is highly salient, especially when reading printed materials. Similarly, it is possible that listeners do not *attend* too closely to gradient differences in the prominence of the nuclear pitch accent to determine the focus size when the context is already salient. Some evidence that this might be the case comes from a recent experiment in which subjects were required to explicitly use prosody to disambiguate focus size. [5] showed that when speakers deliberately attempted to distinguish broad focus productions from narrow focus productions of a sentence, they produced the most dramatic version of the differences discussed above. Importantly, when they did this, listeners were quite successful in recovering the speaker's intended reading. Thus, there is evidence from at least one study that listeners do have expectations about how prosody can be used to express the size of a sentence's focus constituent in English. In the present study we explored how salient those expectations are. In the experiment described below, we collected English-speaking listeners' impressions of prosodic prominence for simple SVO sentences when those sentences appeared in pragmatic contexts that varied the focus structure of the sentence (i.e., following questions like those in (1) and (2), above). Crucially, we presented the same productions of a given sentence in each context, so that listeners' responses regarding a word's prominence across focus conditions could only be accounted for by its information structural meaning, rather than its actual acoustic realization alone.

2. EXPERIMENT

2.1. Methods

2.1.1. Materials

Experimental materials were mini-dialogues consisting of a question and a simple SVO answer. Because related work [4] suggested that highly contrastive contexts would serve as the strongest examples of focus, we used corrective focus rather than WH-focus. That is, short exchanges such as the answer in (2d) above, in reply to each of the questions in (2a-c) were used to create experimental items as follows. Two speakers of American English read all question-answer pairs to one another, a female speaker reading the questions and a male speaker reading the answers. These answer sentences were then edited out of their original contexts and replaced by the version of the answer produced in response to a VP-focus question. That is, the recording of the sentence “*No...because I bought a motorcycle*”, originally recorded in reply to “*Why's your wife mad...because you lost your job?*” was used as the answer sentence to the that question as well as to both “*Why's your wife mad...because the roof's leaking?*” and “*Why's your wife mad...because you bought a car?*”. 51 test stimuli were created in this way (17 SVO test sentences, each appearing in 3 question contexts). Additionally 37 filler dialogues were created that placed focus on the subject, verb, or the object of a preposition in the answer sentence, but which were otherwise similar to the experimental items.

2.1.2. Participants

Thirty native-English speakers, most of whom were UCLA undergraduate students, participated as listeners in the experiment. None had taken a course in intonational phonology or the transcription of prosody. All received compensation for their participation.

2.1.3. Procedure

Listeners participated in a ‘stress rating’ task in which they were asked to listen to each of the mini-dialogues, and to rate how “stressed” certain underlined words sounded in the male's answer sentences. These underlined words were the verbs and the nouns in each sentence. Listeners were instructed to give a rating from 1 (“not at all stressed”) to 5 (“very stressed”), and were told that “stress” referred to how strongly the speaker used

his voice to make the word stand out. Listeners were provided with a transcript of the mini-dialogues, ordered as they were presented to them, and were asked to write their ratings in above each word. (3) shows an example of what the listeners saw and how they might have responded (this hypothetical rating is shown above each underlined target word).

(3) a. Q: What did you do yesterday?

2 4

A: I bought a motorcycle.

b. Q: What did you eat at the picnic?

2 5

A: I ate a hamburger.

After completing a brief practice session and asking questions, participants listened to the 51 test and 37 filler dialogues over Sony MDR-V500 closed, dynamic headphones at a comfortable listening volume (held constant across participants) in a sound-attenuated booth. They provided prominence ratings as above for verbs and objects in each sentence (30 listeners \times 17 test sentences \times 2 words (verbs and objects) \times 3 focus conditions (sentence focus, VP focus, object focus) = 3,060 ratings). Three measures were considered in evaluating the possible effect of the experimental manipulation on participants' judgments: (a) prominence ratings for verbs, (b) prominence rating for objects, (c) the difference between the object's rating and the verb's rating for each sentence (referred to henceforth as OV_{diff}).

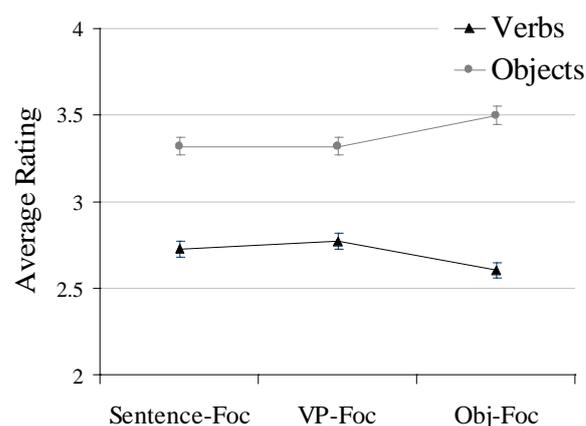
2.2. Results

Figure 1 shows mean ratings for verbs and objects in each of the three focus conditions. The results of repeated measures ANOVA indicated a significant main effect of condition on each measure, both by subjects and by items: verbs $F(2,29) = 6.9$, $p < .001$, $F(2,16) = 3.8$, $p = .02$; objects $F(2,29) = 9.9$, $p < .0001$; $F(2,16) = 4.5$, $p = .01$; OV_{diff} $F(2,29) = 17.1$, $p < .0001$; $F(2,16) = 12.5$, $p < .001$.

The results of Tukey-Kramer pairwise comparisons showed the following. Nuclear pitch accented objects were rated significantly higher in the object focus condition than in either the sentence focus ($p < .01$) or VP focus ($p < .01$) conditions. Conversely, verbs were rated lower in the object focus condition than in the sentence focus ($p < .001$) or VP focus ($p < .05$) conditions.

The relative measure, OV_{diff} , also showed a highly significant difference between object focus and the other two focus conditions: objects were rated significantly higher relative to the preceding verb in the object focus condition compared to either the sentence focus ($p < .001$) or VP focus ($p < .001$) conditions. There were no significant differences found between the sentence focus and VP focus conditions on any measure ($p > .1$ in all cases).

Figure 1: Average prominence ratings for verbs and objects in test sentences in the three focus conditions. 1 is lowest in prominence, 5 is highest. Error bars show standard error.



Although the present study was not designed to examine frequency effects, a post-hoc correlation suggests that the focus-size effect we found might have been somewhat modulated by the lexical frequency of the sentences' verb. We plotted the difference in verb ratings across focus conditions (that is, the absolute difference between a verb's prominence rating in the broad VP focus context and narrow object focus context) as a function of that verb's log frequency. There was a relatively strong negative correlation ($R^2 = .39$). This suggests that as the frequency of the verb increased, its vulnerability to the illusory effect of focus condition decreased.

3. DISCUSSION

The results of the experiment above suggest that listeners do in fact have knowledge regarding how the size of the focus constituent (at least a basic broad versus narrow contrast), relates to patterns of prosodic prominence. Importantly, their expectations are not just about the nuclear pitch accent per se, but about the relation between nuclear and prenuclear prominence: the narrower the focus on an object, the more *relatively*

prominent it should be. This is precisely the way that speakers have been reported to manipulate prominence in productions [1, 2, 5, 9, 10, 11, 13]. The fact that listeners in the above experiment were not presented with any acoustic differences, however, (i.e., the same production of an answer was presented in all three focus contexts), suggests that we were observing their expectations about how prominence can be used to express different information structures.

Other, non-information structure-related factors are known to have similar illusory effects on listeners' perception of prosodic prominence. For example, Cole and colleagues [7] attempted to model listeners' prominence judgments for words in excerpts of unscripted (American English) dialogue. One of the results from that study was that the best model of the probability that a listener would judge a word as prominent did not include only acoustic features, but also frequency information. Listeners tended to hear words as prominent simply because they were low-frequency, somewhat independent of their actual acoustic prominence. We probed for an effect of frequency and found evidence that it might be interacting with the focus effect: there was a negative correlation between the size of the focus effect for an item, and the frequency of the item's verb. A recent theory of how focus is interpreted from prosody (i.e., [6]) predicts this sort of interaction, and suggests that factors such as frequency should independently contribute to an "expected prominence" value for a word, which in turn should have consequences for the level of prominence required to mark focus. In the present experiment, it is unclear to what extent the item effect found reflects the acoustic properties of the stimuli used, and to what extent (if any) it simply reflects listeners' expectations for them. This matter is being investigated currently.

4. CONCLUSION

In this study we presented English listeners with productions of simple SVO sentences that differed only in terms of the size of their focus constituent. We found that this information structural manipulation had the effect of an auditory illusion: prenuclear prominence was heard as lower, and nuclear prominence as higher, when an object was narrowly rather than broadly focused. We interpreted this in terms of listeners' expectations for patterns found in productions. It was also

suggested that the illusory effect focus has on listeners' impressions of prominence could be used to probe for other influences.

5. REFERENCES

- [1] Baumann, S., Becker, J., Grice, M., Mücke, D. 2007. Tonal and articulatory marking of focus in German. *Proc. 16th ICPhS Saarbrücken*, 1029-1032.
- [2] Baumann, S., Grice, M., Steindamm, S. 2006. Prosodic marking of focus domains – categorical or gradient? *Proc. Speech Prosody 2006*, 301-304.
- [3] Birch, S., Clifton, C. 1995. Focus, accent and argument structure: effects on language comprehension. *Lang. & Speech* 38, 365-391.
- [4] Bishop, J. In prep. *Information Structural Expectations*.
- [5] Breen, M., Fedorenko, E., Wagner, M., Gibson, E. 2010. Acoustic correlates of information structure. *Proc. Lang. & Cog.* 25, 1044-1098.
- [6] Calhoun, S. 2006. *Information Structure and the Prosodic Structure of English: A Probabilistic Relationship*. Ph.D. dissertation, University of Edinburgh.
- [7] Cole, J., Mo, Y., Hasegawa-Johnson, M. 2010. Signal-based and expectation-based actors in the perception of prosodic prominence. *Laboratory Phonology* 1, 425-452.
- [8] Gussenhoven, C. 1983. Testing the reality of focus domains. *Lang. & Speech* 26, 61-80.
- [9] Hanssen, J., Peters, J., Gussenhoven, C. 2008. Prosodic effects of focus in Dutch declaratives. *Proc. Speech Prosody 2008*, 609-612.
- [10] Jun, S.A. 2010. Focus domains and focus types: Their realizations in English and Korean. A talk presented at the *7th Workshop on Altaic Formal Linguistics*.
- [11] Sityaev, D., House, J. 2003. Phonetic and phonological correlates of broad, narrow and contrastive focus in English. *Proc. 15th ICPhS Barcelona*, 1819-1822.
- [12] Welby, P. 2003. Effects of pitch accent position, type, and status on focus projection. *Lang. & Speech* 46, 53-81.
- [13] Xu, Y., Xu, X. 2005. Phonetic realization of focus in English declarative intonation. *J. Phon.* 33, 159-197.