Focus Placement on Adjacent Words in Yes/No Questions

Jonathan Howell

Montclair State University
howellj@montclair.edu

ABSTRACT

Semantic theory predicts four different focus configurations for a pair of adjacent words: early, late, broad and double focus. We report on a production study which confirms that speakers make a significant, though non-obligatory, distinction between the four categories. The distinction is observed in yes/no question and so cannot be associated with a specific intonational tune. The results also fail to support theories of focus projection and uniquely syntagmatic models of prominence.

Keywords: focus, prosody, prominence, double focus

1. INTRODUCTION

Given a sequence of syntactic constituents *A B*, formal semantic theories of focus (e.g. [1]) predict four possible configurations:

(1) early $[A]_F$ B   broad $[A \quad B]_F$
    late $A [B]_F$ double $[A]_F [B]_F$

Models of prosodic prominence will either have an isometric set of patterns or else a non-direct mapping in which the same prosodic pattern maps to two or more focus configurations.

Strictly syntagmatic models of prosodic prominence, such as those found in the early metrical phonology literature (cf. [2]), allow for only two patterns of prominence: weak-STRONG and STRONG-weak.

Models which adopt paradigmatic stress (e.g. primary vs. secondary stress; or different pitch accent types) can distinguish four different prosodic categories corresponding to the different focus configurations. Autosegmental-metrical phonology (cf. [3] accommodates this by positing a hierarchy of prosodic categories: prominence remains syntagmatic at each level of the hierarchy. This analysis predicts a significant difference in production for all four focus configurations.

The considerable literature on focus projection (see [4], [5] among others) conflates broad and late focus to a single weak-STRONG pattern of prominence. This analysis predicts no significant difference in production for broad and late focus.

Previous studies examining double focus have not examined adjacent foci. [6] found significant differences in duration and mean F0 across the four focus conditions in declarative sentences. [7] and [8] found that listeners tended to associate particular ToBI-annotated declarative utterances with particular focus conditions. It is unclear, however, how these results will generalize (i) to adjacent foci and (ii) to non-declarative contexts.

2. METHOD

2.1. Stimuli

Twelve verb-object pairs were selected which contained only sonorants: e.g. *email Owen*, *earn yen*, *iron wool*. A yes-no question was written for each pair with additional phonetic material inserted (e.g. *lately*, *last night, sometimes*) in order to mitigate effects of phrase-final lengthening.

For each question, four different contexts were written, corresponding to the four focus conditions: focus on the verb (early focus); focus on the object (late focus), double focus on verb and object and broad focus on the verb phrase.

(2) **Target words**: *iron wool*
    Target question: *Can you iron wool sometimes?*

(3) Focus Contexts
    Early: I know you can wash wool.
    Late: I know you can iron cotton.
    Double: I know you're not supposed to dryclean cashmere and you have to wash most sweaters by hand.
    Broad: I need your advice.

2.2. Participants

Participants were 53 students and staff from an American university in the New York metropolitan area. Participants self-reported as speakers of native American English and all were paid a nominal sum.

2.3. Recording, Annotation and Measurement

Participants were recorded reading the stimuli in a sound-attenuated room. The stimuli were presented on a computer screen using a set of MATLAB scripts. Each participant was presented with only

one condition of each item, in a Latin-square design and with a pseudo-random presentation order.

The recordings were annotated at the word and phone level using forced-alignment [9] and acoustic measures were extracted using Praat [10], including word duration, minimum F0, maximum F0 and F0 range. Data are published at [11].

2.4. Analysis

The experiment was intended to measure acoustic variation due to focus condition. In order to remove acoustic variation due to speaker and item, we performed linear residualization [12]. We computed linear regression models that predicted an acoustic measure from speaker and item. We subtracted the predicted measurement from the actual measurement for each token. The resulting residual value is expected to more closely reflect variance due to experimental condition. The following results are reported using these residual values.

We computed a linear discriminant function to predict focus condition (early, late, double or broad) from a set of residualized acoustic measures. The significance of the model was assessed by a Wilk's lambda test and by accuracy in leave-one-out classification.

3. RESULTS

We computed several discriminant function models using duration, F0 minima, F0 range and intensity maxima and combinations thereof. The Wilk's lambda was significant for all of the models.

Of these, the best-performing model used duration and F0 range ($\Lambda = .94$, $\chi^2(12) = 2.78$, p < .001). Leave-one-out classification correctly predicted 30% of the tokens. The model was able to correctly classify broad focus 21% of the time, double focus 18% of the time, object focus 32% of the time and verb focus 48% of the time.

Figure 1 displays the mean measures of F0 range, minimum F0, intensity and duration for verb and object in each of the four focus conditions. For ease of presentation, the target words *iron* and *wool* are used as labels, with *[]F* brackets indicating the intended location of focus.

Looking at means in the single focus conditions, the focused word had a greater F0 range, a lower F0 minimum, greater intensity and a longer duration than the unfocused word.

In the broad focus condition, the verb and object had comparable F0 range, while the object had a lower F0 minimum, lower intensity and longer word duration than the verb.

Figure 1: Means of F0 minima, F0 range, intensity and duration for verb and object in the four focus conditions.
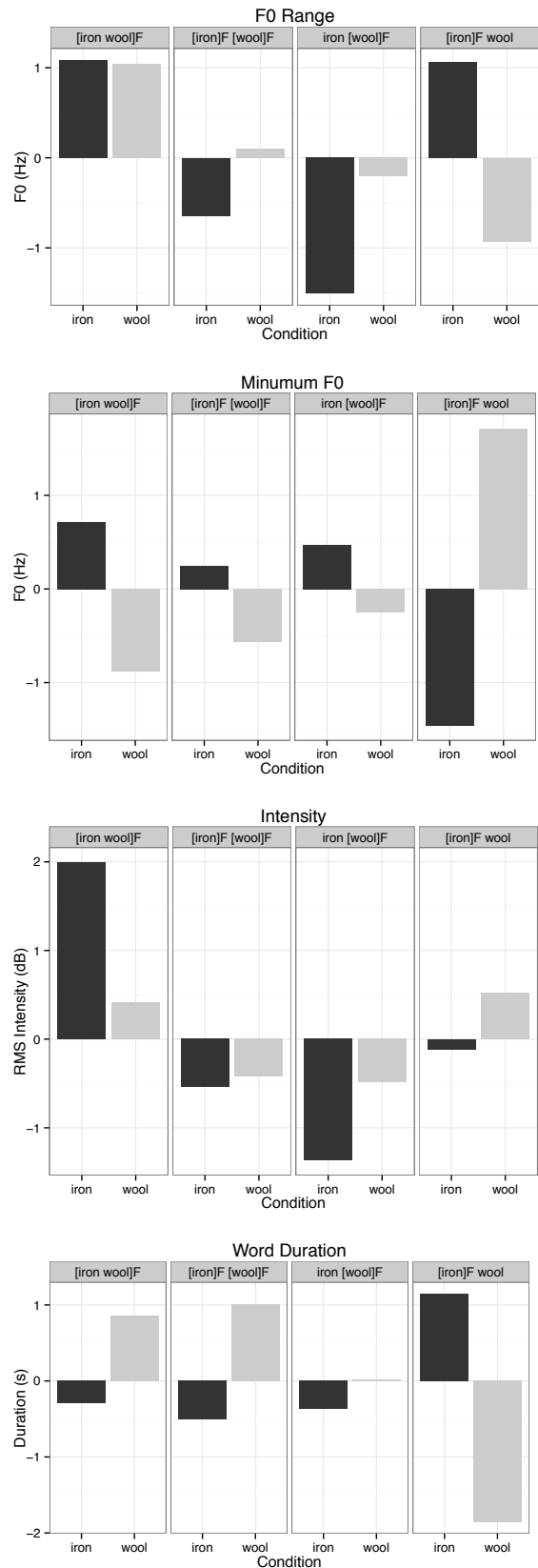
In the double focus condition, the object had a greater F0 range, a lower F0 minimum and a longer duration than the verb. The intensity of the verb and object were comparable.

4. DISCUSSION AND CONCLUSION

The results confirm a statistically significant production difference between the four focus conditions posited by semantic theory: early focus, late focus, broad focus and double focus.

The results do not support models of prosodic prominence which predict fewer than four conditions for adjacent constituents.

First, a uniquely syntagmatic model of prominence which defines prominence with respect to adjacent constituents only (strong-WEAK vs. weak-STRONG) predicts the early and late focus conditions, but fails to predict the broad and late focus conditions. Additional, paradigmatic information is required to differentiate the other two conditions.

Autosegmental-metrical theory provides for two kinds of paradigmatic information: choice of tone (i.e. pitch accent or boundary tone type) and phrasing. Since the pitch accent and boundary tones typically used in a yes-no question context (e.g. L* H- H%) differ from the pitch accent and boundary tones typically used in a declarative context (e.g. H* L- L%), it is unclear how the specific choice of tone could account for the four focus conditions attested.

Autosegmental-metrical theory also provides a distinction between pre-nuclear and nuclear pitch accents, which is a reflex of prosodic phrasing. On this approach, we expect a single phrase containing verb and object in the broad focus condition; and we expect two different phrases in the double focus condition.

Correspondingly, the verb is expected to have a pre-nuclear accent (if any) on the verb in the broad focus condition and a nuclear accent on the verb in the double focus condition. The mean values for F0 minimum and F0 range are not consistent with this contrast. The verb had a lower minimum F0 and greater F0 range in the broad rather than the double focus condition.

Phrase-final lengthening is also a reflex of prosodic phrasing: we expect the verb to exhibit phrase final lengthening in the double, but not the broad focus condition. The mean values for duration are not consistent with this contrast. The verb had greater duration in the broad rather than the double focus condition.

Second, theories of focus projection (e.g. [4],[5]) predict no production difference between broad and late focus. The results do not support the prediction, since a statistically significant difference was observed between all four conditions.

Investigations of these focus configurations by [7] and [8] found that listeners could perceive a difference between particular ToBI-annotated declarative utterances corresponding to the different focus configuration. The present study confirms that naïve speakers produce the configurations with statistically significant differences. The results also confirm that the production differences are present in yes-no question contexts.

The investigation by [6] found statistically significant differences in declarative utterances, however the authors did not attempt classification. Differences found in the present study, although highly significant, do not appear to be categorical. The measures observed are not sufficient to reliably categorize new data. Future investigations may reveal that other acoustic measures may yield better performing classifiers. However, the present data are also consistent with the presence of acoustics cues which are available, but not required, to disambiguate between focus configurations.

5. REFERENCES

[1] Rooth, M. 1992. A theory of focus interpretation. *Natural Language Semantics* 1, 75-116.

[2] Ladd, B. 1991. One word's strength is another word's weakness: Integrating syntagmatic and paradigmatic aspects of stress. *ESCOL 7*.

[3] Ladd, D.R. 2008. *Intonational Phonology*. Cambridge University Press.

[4] Selkirk, E. 1995. Sentence prosody: intonation, stress, and phrasing. In: Goldsmith, J.A. (ed), *Handbook of Phonological Theory*. London: Blackwell, 550-569.

[5] Gussenhoven, C. On the limits of focus projection in English. In: Bosch, P., van der Sandt, R. (eds), Focus. Linguistic, Cognitive, and Computational Perspectives. Cambridge, 43-55.

[6] Eady, S. J., Cooper, W.E., Klouda, G.V., Mueller, P.R., Lotts, D.W. 1986. Acoustical characteristics of sentential focus: Narrow vs. broad and single vs. dual focus environments. *Lang Speech* 29, 233-251.

[7] Jannedy, S. 2002. *Hat Patterns and Double Peaks: The Phonetics and Psycholinguistics of Broad versus Late Narrow versus Double Focus Intonations.* Doctoral dissertation: The Ohio State University.

[8] Welby, P. 2003. Effects of pitch accent position, type, and status on focus projection. *Language and Speech* 46, 53-81.

[9] Gorman, K., Howell, J., Wagner, M. 2011. Prosodylab-Aligner: A tool for forced alignment of laboratory speech. *Canadian Acoustics* 39, 192–193.

[10] Boersma, P., & Weenink, D. 2014. Praat: doing phonetics by computer [Computer program]. Version 5.4.04

[11] Howell, Jonathan, 2015, "Double Focus of Adjacent Words in Yes/No Questions", http://dx.doi.org/10.7910/DVN/29343 Harvard Dataverse Network [Distributor] V1 [Version]

[12] Breen, M., Fedorenko, E., Wagner, M., & Gibson, E. 2010. Acoustic correlates of information structure. *Language and Cognitive Processes* 25, 1044-1098.