Prosodic Effects of Discourse Salience and Association with Focus

M. Wagner[2], *M. Breen*[1], *E. Flemming*[3], *Stefanie Shattuck-Hufnagel*[3], *& E. Gibson*[3]

[1]Department of Linguistics, McGill University
[2]Department of Psychology, UMass Amherst
[3]Massachusetts Institute of Technology
chael@mcgill.ca, mbreen@psych.umass.edu

Abstract

Three factors that have been argued to influence the prosody of an utterance are (i) which constituents encode discourse-salient information; (ii) which constituents are contrastive and evoke alternatives; and (iii) which constituents interact with the meaning of focus operators such as *only* (i.e., they 'associate' with focus). One challenge for a better understanding of the prosodic effects of these factors has been the difficulty of finding a way to evaluate hypotheses quantitatively, since individual variation in productions is often large enough to wash out experimental effects. In this paper, we apply a methodology introduced in [1] which regresses out subject and item variation, uncovering otherwise hidden prosodic patterns, and show how the three factors interact in sentences containing single or multiple foci.
Index Terms: prosody, focus association, givenness, prominence, production

1. Prosodic Effects of Focus and Discourse-Salience

One of the goals of research on prosody is to determine the conditions which underlie prosodic prominence. This is particularly important is cases of association with focus [2], where the truth-conditions of a sentence change depending on the prosodic realization of material in the scope of certain focus-sensitive operators such as *only*:

(1) a. Gramma only gave *a bunny* to Maryanne.
b. Gramma only gave a bunny to *Maryanne*.

Pronunciation (1a), with prominence on *bunny*, seems compatible with a scenario in which Gramma also gave a bunny to John; (1b) does not. (1b) seems compatible with a scenario in which Gramma also gave a scarf to Maryanne; (1a) does not. In the first example *bunny* is focused and in the second *Maryanne*. The current study explores how prominence patterns reflect the intended meaning of sentences in which one or two focused elements associate with *only*. We use a normalization procedure from [1] which uncovers otherwise hidden prosodic effects.

1.1. Evoking Alternatives

One way to explain the difference between (1a) and (1b) is to analyze *only* as operating over alternatives to the the sentence in which it occurs. The alternatives for (1a) involve substitutions for *bunny*, while those for (1b) vary *Maryanne*. In the following, smallcaps denote constituents which evoke alternatives:

(2) a. 'Gramma only gave A BUNNY to Maryanne.'
Alternatives: { Gramma gave a scarf to Maryanne; Gramma gave a cake to Maryanne; ...}

b. 'Gramma only gave a bunny to MARYANNE.'
Alternatives: {Gramma gave a bunny to John; Gramma gave a bunny to Bill; ...}

Sentences involving *only* are often analyzed as presupposing the content of the sentence without only (e.g., *Gramma gave a bunny to Maryanne* in the case of (1a)), and asserting that no alternative to it is true [3]. So both (1a) and (1b) share the presupposition that Gramma gave a bunny to Maryanne, but they differ with respect to which alternative statements they exclude. The prosody of a sentence is sensitive to which alternatives are contextually relevant: Constituents which evoke alternatives are usually more prominent than those that do not.

1.2. Anaphoric Destressing

Another factor affecting prosody is that discourse-salient material is usually less prominent than discourse-new material (cf. [4] for an overview), a phenomenon often called 'anaphoric destressing.' Underlining marks discourse-salient material:

(3) a. Why did Maryanne feel special?
b. Grandma gave a bunny to Maryanne.

In a typical rendition of (3b), the noun phrase *Maryanne* is less prominent than the preceding *a bunny* because its referent is discourse-salient. However, in the following context the indirect object would typically be *more* prominent than the preceding object because it evokes alternatives:

(4) a. Why did Maryanne feel special?
b. Grandma only gave a bunny to MARYANNE.

The prosodic realization of a constituent is thus affected by two factors: (i) Which constituents evoke alternatives? (ii) Which constituents are salient in the discourse? The current study varies the information status (IS) of constituents along both dimensions. In conditions 1-5, an introductory story made all discourse referents (e.g.*John*, *Maryanne*, *bunny*, and *scarf*) salient, and the *set-up* in each condition further manipulated the IS of the target sentence. Conditions 1 and 2 vary the IS of the indirect object *Maryanne* such that it is discourse-salient in both cases, but evokes alternatives in only in condition 2:

(5) *Story*: It was Christmas, and Gramma was deciding what gifts to give to her grandchildren, John and Maryanne. She had knitted two scarves as gifts, and had also purchased a couple of stuffed bunnies. She wrapped up a scarf and a bunny for John. Then she remembered how rude Maryanne had been at Thanksgiving.

Table 1: *Information Status by Condition*

Condition	Direct Object	Indirect Object
1	GIVEN & EVOKES ALT.	given
2	GIVEN & EVOKES ALT.	GIVEN & EVOKES ALT.
3	given	GIVEN & EVOKES ALT.
4	GIVEN & EVOKES ALT.	GIVEN & EVOKES ALT.
5	GIVEN & EVOKES ALT.	GIVEN & EVOKES ALT.
6	EVOKES ALT.	EVOKES ALT.

a. Condition 1: *Set-up*: Gramma didn't give a scarf to Maryanne. *Target*: Gramma only gave a BUNNY to Maryanne.
b. Condition 2: *Set-up*: Gramma gave a scarf and a bunny to John. *Target*: Gramma only gave a BUNNY to MARYANNE.

In Condition 1, *Maryanne* is discourse-salient and does not evoke alternatives. In Condition 2, while also given, *Maryanne* evokes an alternative: *John*. In Conditions 3 and 4 the IS of *bunny* was similarly manipulated:

(6) *Story*: It was Christmas, and Gramma was deciding what gifts to give to her grandchildren, John and Maryanne. She had knitted two scarves as gifts, and had also purchased a couple of stuffed bunnies. She wrapped up a scarf and a bunny for Maryanne. Then she remembered how rude John had been at Thanksgiving.
   a. Condition 3: *Set-up*: Gramma didn't give a bunny to John. *Target*: Gramma only gave a bunny to MARYANNE.
   b. Condition 4: *Set-up*: Gramma gave a scarf to both Maryanne and John. *Target*: Gramma only gave a BUNNY to MARYANNE.

In Conditions 1-4, the discourse-salience of *bunny* and *Maryanne* is held constant while the evoking of alternatives is varied. In Conditions 5 and 6, on the other hand, the evoking of alternatives is held constant, but discourse-salience is varied. Earlier research has already shown that discourse-given material can remain accented and at the same time show quantitative signs of reduction compared to discourse-new accented material [5]. In Condition 5, both *bunny* and *Maryanne* associate with *only* and are discourse salient. In Condition 6, the story does not specifically mention Maryanne or a bunny, so while both objects evoke alternatives, they are both discourse-new:

(7) Condition 5:
Set-up: Gramma didn't give a scarf to Maryanne, and she didn't give either a bunny or a scarf to John.
Target: Gramma only gave a BUNNY to MARYANNE.

(8) Condition 6:
Set-up: Gramma picked one present and gave it to her favorite grandchild.
Target: Gramma only gave a BUNNY to MARYANNE.

Table 1 summarizes the IS of the 6 experimental conditions.

1.3. Multiple Foci

In its simplest form, the alternatives theory of association with focus [3] predicts that any alternatives evoked in the scope of a focus-sensitive operator such as *only* must be considered in the alternatives that *only* excludes—they should all 'associate' with *only*. For example, the following sentence should indicate that Gramma gave only one thing to one person:

(9) Gramma only gave a BUNNY to MARYANNE.

However, it turns out that multiple focus constructions do not necessarily work like this [6]. In condition 2 (5b), e.g., despite the fact that both direct and indirect object evoke alternatives, only the direct object associates with *only*. Similarly, in condition 4 (6b), only the indirect object associates with *only*. That there can be constituents in the scope of *only* which evoke alternatives but do not associate with it presents a problem for the theory. [7] observed a related issue. In sentences with more than one focus operator (i.e. (10)), a focus can associate either with the higher focus operator (*also*) or the lower one *only*, but the focus operators don't necessarily associate with both:[1]

(10) We only$_1$ recovered the diary entries that Marilyn$_1$ made about John.
We also$_2$ only$_1$ recovered the diary entries that Marilyn$_1$ made about Bobby$_2$.
'Also with respect to diary entries about Bobby, we only recovered the ones that Marilyn made.'

Rooth's example suggests that the evoked alternative sets may have to be indexed with respect to the focus operator that quantifies over them. Similar to the paraphase of Rooth's example in (10), our Conditions 2 and 4 can be paraphrased using a topic construction. Condition 2 could be paraphrased: 'As for Maryanne, Gramma only gave a bunny to her.'; Condition 4 could be paraphrased: 'As for bunnies, Gramma only gave one to Maryanne.' We can analyze the data by positing a contrast-operator which, like *also* in (10), associates with one focus, preventing *only* from associating with it. This is the analysis given to contrastive topics more generally in [9]: contrastive topics are focus-sensitive operators that outscope a lower focus operator.

(11) Condition 2:
Set-up: Grammar gave a scarf and a bunny to John.
Target: Gramma *contrast*$_1$ only$_2$ gave a BUNNY$_2$ to MARYANNE$_1$.

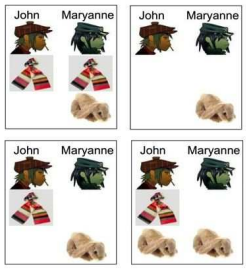
(12) Condition 4:
Set-up: Gramma gave a scarf to both Maryanne and John.
Target: Gramma *contrast*$_1$ only$_2$ gave a BUNNY$_1$ to MARYANNE$_2$.

The novel question in our experiment is whether the prosody of constituents that evoke alternatives is affected by which focus operator they actually associate with. Table 2 summarizes the focus association patterns of the three conditions in which the two objects both evoke alternatives and are discourse salient.

Table 2: *Focus Association by Condition*

Condition	Direct Object	Indirect Object
2	only	contrast
4	contrast	only
5	only	only

[1] This may be not be possible with all kinds of focus operators, however, see [8].



Gramma only gave a bunny to Maryanne.

Figure 1: *An example of the picture array that subjects chose from to indicate their interpretation of the target sentence.*

2. Method

2.1. Participants

10 pairs of English speakers from the MIT community participated in the experiment, each one received $15 for participating.

2.2. Materials

The stimuli consisted of three sections: Story, set-up, and target. The story provided a scenario for the action, and served, in all but Condition 6, to introduce two people (e.g. *Maryanne* and *John*) and two objects (e.g. *a bunny* and *a scarf*). The set-up determined the IS of the target sentence–the discourse salience of the people and objects, and which alternatives *only* associated with. The target always took the form: *Actor* only *verb*ed a *object* to *Name*. The actor, object, and name were all 2-syllables with first-syllable stress. The verb consisted of one syllable. Length and metrical stress were matched across items. The target sentence were comprised mainly of sonorants, to facilitate automatic pitch extraction. There were 10 items in total in 6 conditions, making for a total of 60 stories.

2.3. Procedure

Two participants (a *speaker* and a *listener*) sat at two computers in the same room such that neither could see the other's screen. On each trial, the speaker first read the story, set-up, and target silently. S/he then chose from an array of four pictures (Fig. 1) to assure they understood the context and gave the sentence the right meaning. The speaker then produced the set-up and target aloud for the listener, who selected the picture s/he believed the target described. Trials for which the listener chose the wrong picture were excluded, as were productions with disfluencies. Using Praat ([10]), 24 acoustic measures were automatically extracted from each of five words in the target sentence (*Gramma, only, gave bunny, Maryanne*). A participant was a speaker for a subset of half the trials, then roles were switched.

3. Results

In order to see whether the acoustic measures could discriminate the speakers' productions, we entered all 120 predictors into a series of step-wise discriminant analyses. In initial analyses, neither the full set of 6 conditions, nor pairs of conditions, were successfully discriminated. To remove variance due to speakers and items, we followed [1] and computed linear regression models in which speaker (n = 20), item (n = 20), and the interaction between them, predicted the 120 acoustic measures. From each model, we calculated the predicted

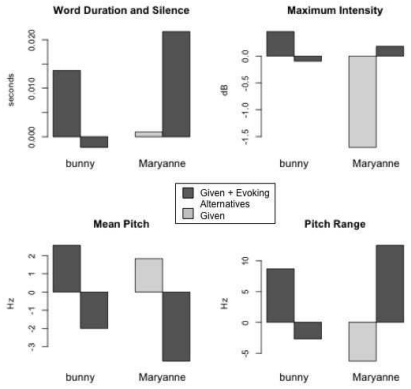


Figure 2: **Given vs. Given+Evoking Alternatives (condition 1 vs. 2):** *'Maryanne' is given in condition 1 (light grey, left-hand bar), and also in condition 2, but there it also evokes alternatives (darker grey, right-hand bar), and is longer, has greater intensity and a higher pitch range. Interestingly, the prominence of 'bunny' decreases (left vs. right bars) when that of Maryanne increases although its information status remains contant—suggesting that the relative prominence between the two arguments is what matters in coding the information status of 'Maryanne.' In condition 3 vs. 4 (no figure), similarly, 'bunny' is longer and has higher intensity when evoking alternatives (cond. 4) compared to when not (cond. 3).*

.

value of each acoustic feature per item per speaker. The difference between the predicted and actual values (i.e. the residual measure) reflects acoustic differences not due to differences between speakers or items, or their interaction. We submitted these residual measures to a stepwise discriminant analysis, to independently determine which acoustic measures speakers used to differentiate productions. Eight of the original 120 acoustic measures (duration, mean pitch, pitch range, and maximum intensity from *bunny* and *Maryanne*, respectively) resulted in better-than-chance 6-way classification of the productions by condition; moreover, many conditions were now discriminated in pair-wise comparisons.

3.1. Given vs. Given + Evokes Alternatives

Conditions 1 & 2 were well discriminated. Wilks' lambda for the comparison between 1& 2 was .81, F(8) = 5.08, $p < .001$. A leave-one-out classification successfully classified 68% of all productions; 66% of condition 1; 70% of condition 2. Conditions 3 & 4 were also well discriminated (Wilks lambda = 0.83, F(8)=4.52, $p < .001$, classification accuracy 58% condition 3, 69% condition 4). Discourse-salient constituents that evoke alternatives are produced with longer duration, higher intensity, and wider pitch range than when they do not evoke alternatives (Fig. 2). The bar plots in all Figures are averages of the normalized measures over all items—we label them with words from a particular stimulus set to make them easier to interpret.

3.2. Given + Evokes Alternatives vs. Evokes alternatives

Conditions 5 and 6 were discriminated, Wilks' lambda = .86, F(8) = 3.59, $p < .001$. Leave-one-out classification successfully classified 60% of all productions; 67% of condition 5; 52% of

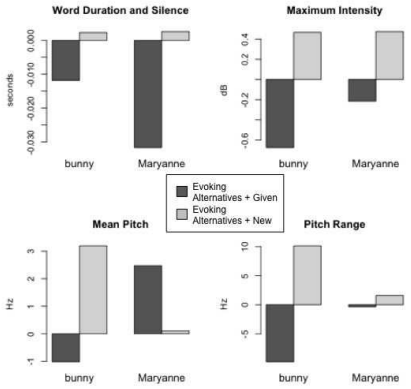


Figure 3: **Evoking Alternatives + Given vs. Evoking Alternatives and New (condition 5 vs. 6):** *Both 'bunny' and 'Maryanne' are shorter and have a lower intensity when they are discourse salient (condition 5, darker bars) compared when they evoke alternatives and are new (condition 6, lighter bars).*

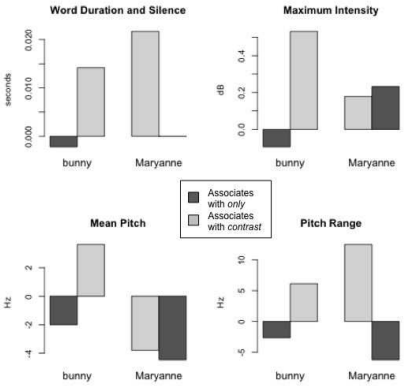


Figure 4: **Association with contrast vs. association with only (condition 2 vs. 4).** *Condition 2 are the bars on the left, condition 4 the bards on the right. When 'bunny' and 'Maryanne' associate with only (darker bars) they have a lower duration, intensity, pitch and pitch range compared to when associating with contrast (lighter bars).*

condition 6. Acoustic results show that, when the constituent evokes alternatives and is discourse-new, it is produced with a longer duration, higher intensity, and wider pitch range than when it evokes alternatives but is discourse-salient (Fig. 3).

3.3. Association with *only* vs. association with *contrast*

Conditions 2 and 4 were discriminated, Wilks' lambda = .90, $F(8) = 2.46$, $p < .05$. Leave-one-out classification successfully classified 55% of all productions; 51% of condition 2; 60% of condition 4 (Fig. 4). The acoustic results show that foci which associate with *contrast* are on average produced with longer duration, higher pitch, higher intensity, and wider pitch range than foci which associate with *only*.

4. Discussion and Conclusion

The results show that constituents which are discourse-salient but also evoke alternatives are more prominent than constituents that are discourse salient and do not (Condition 1 vs. 2, Condition 3 vs. 4). Furthermore, discourse-new constituents which evoke alternatives are more prominent than discourse-salient constituents which evoke alternatives (Conditions 5 vs 6), showing that anaphoric reduction can be observed even on accented constituents that evoke alternatives. This is similar to the finding in [5] that accented words are reduced when discourse-salient. Our study did not look at the case in which a constituent is new and does not evoke alternatives. This IS was compared to constituents that are new and *do* evoke alternatives in [11], who found the latter to be more prominent.

With respect to multiple foci, we found that foci which associate with *contrast* are more prominent than foci which associate with *only*. The results show for the first time *that* such an effect exists, but do not tease apart whether this is due to the nature of the focus operators *contrast* and *only* themselves, or whether it is a result of the scope of the focus operators (cf. [12] for a related claim on second-occurrence focus)—in our stimuli, *contrast* arguably always outscoped *only*. A further methodological result of this study is that regressing out subject and item effects following [1] can uncover effects that are otherwise washed out by variation.

5. Acknowledgements

Thanks to Nakul Vyas for theoretical discussion and development of materials, to Denise Ichinco for assistance with data analysis, to Wade Shen for automatic alignment of the productions, and to Lisa Selkirk and Jonah Katz for discussion.

6. References

[1] M. Breen, E. Fedorenko, M. Wagner, and E. Gibson, "Acoustic correlates of information structure," 2009, under Revision.

[2] R. S. Jackendoff, *Semantic Interpretation in Generative Grammar*. Cambridge, Ma.: MIT Press, 1972.

[3] M. Rooth, "A theory of focus interpretation," *Natural Language Semantics*, vol. 1, pp. 75–116, 1992.

[4] A. Cutler, D. Dahan, and W. van Donselaar, "Prosody in the comprehension of spoken language: A literature review," *Language and Speech*, vol. 40, no. 2, pp. 141–201, 1997.

[5] E. G. Bard, A. H. Anderson, C. Sotillo, M. Aylett, G. Doherty-Sneddon, and A. Newlands, "Controlling the intelligibility of referring expressions in dialogue," *Journal of Memory and Language*, vol. 42, pp. 1–22, 2000.

[6] M. Krifka, "A compositional semantics for multiple focus constructions," in *Informationsstruktur und Grammatik*, ser. Linguistische Berichte, J. Jacobs, Ed. Opladen: Westdeutscher Verlag, 1992, vol. Sonderheft 4, pp. 17–53.

[7] M. Rooth, "Focus," S. Lappin, Ed. London: Blackwell, 1996, pp. 271–297.

[8] S. Beck and S. Vasishth, "Multiple focus," *Journal of Semantics*, 2009.

[9] M. Wagner, "A compositional analysis of contrastive topics," in *Proceedings of NELS 38 2007 in Ottawa*, M. Abdurrahman, A. Schardl, and M. Walkow, Eds., vol. 2, in press, pp. 415–428.

[10] P. Boersma and D. Weenink, "PRAAT, a system for doing phonetics by computer. report 132." 1996, institute of Phonetic Sciences of the University of Amsterdam.

[11] J. Katz and E. Selkirk, "Contrastive focus vs. discourse-new: Evidence from prosodic prominence in English," 2009, Ms. MIT/UMass Amherst.

[12] M. Rooth, "Second occurrence focus and relativized stress F," 2007, Ms. Cornell University.