Prosodic focus marking in Canadian English

Jiseung Kim, Anja Arnhold
University of Alberta
Jiseung1@ualberta.ca, arnhold@ualberta.ca

Abstract
The current study investigated how broad focus, narrow focus and given information are produced in Canadian English. Given previous findings that showed different varieties of English signal information structure differently, we hypothesized that the effects of focus on acoustic correlates involving duration, f0, and intensity would manifest differently in Canadian English than in Mainstream American English (MAE). Thirty-eight native speakers of Canadian English produced 24 short transitive sentences in different focus conditions: broad focus and narrow focus in different locations (Subject, Verb, Object). A total of 2,736 words were analyzed. While some acoustic correlates such as duration and maximum intensity replicated the same patterns as previous findings in MAE, mean intensity and f0 measures showed different patterns. These results suggest that speakers of Canadian English may employ a different set of acoustic correlates from speakers of MAE. The study sheds light on the role of dialect in the production of focus and givenness and expands our knowledge about the fine-grained details of the phonetic realization of prosodic focus marking in English.

Index Terms: information structure, pitch accent, Canadian English

1. Introduction
While it has been established that focus is prosodically marked in English through changes in f0, intensity, and duration ([12, 14, 16, 20, 26]), compared to our knowledge of focus realization in other languages there are relatively few studies that comprehensively examine different acoustic correlates associated with focus, using modern statistical analyses to characterize the phonetic details ([8]).

In the most comprehensive study to date, [8] investigated whether differences between focus location (S, V, O), breadth (broad and narrow), and type (contrastive and noncontrastive) are systematically distinguished, and which acoustic features are associated with these distinctions in American English. They showed that narrow focus was marked by longer duration, higher f0 mean and maximum, and higher maximum intensity compared to given information. In addition, focus breadth distinction (broad vs. narrow object focus) was marked by longer duration, higher f0 mean and maximum, and higher intensity, whereas contrastiveness was marked by longer duration, higher intensity and lower f0 mean and maximum.

There are additional inconsistent findings regarding the fine-grained details of f0 modulation in focus marking. Studies did not agree whether focused words are characterized by higher f0 than unfocused words ([3, 22]) or post-focal f0 drop ([12, 13]). Moreover, most previous studies examined a single focused word in the same location in an utterance without considering the different locations of the word in the utterance or the prosodic effects of the surrounding words ([19]). Some of the earlier studies have often built on findings from different languages, despite the fact that different languages mark focus differently ([9]). Recognizing these aspects, the current study is motivated to provide further insight into the mechanisms of prosodic focus marking in English.

In addition, it is still unclear whether and how the phonetic details of prosodic focus marking systematically differ between Canadian English and MAE. There is substantial evidence showing that prosodic focus marking is realized differently in regional dialects ([15, 21]). [10] examined four pitch accent types of Southern Standard British English (SSBE) and showed that the relation between pitch accent types and information status is not aligned with that of MAE. [1] reported two perception experiments. In Experiment 1, the eye-movements of 42 native listeners of Canadian English were tracked while listening to the auditory stimuli in SSBE. Unlike the British English listeners who associated rising pitch accents with givenness in [10], the Canadian English listeners did not make the same association. The results of Experiment 2 further showed how focus marking in Canadian English and SSBE differs: Canadian English listeners did not accept rising accents in Canadian English speech regardless of whether the target word signaled newness or givenness, but they rated rising accent on given words highly when listening to SSBE speech, suggesting that prosodic focus marking in Canadian English differs from that in SSBE.

Moreover, [18] found that duration and intensity are two primary markers of focus while f0 played a minor role in British and Irish English, contrary to the findings reported in [8]. On the other hand, [27] reported both intensity and f0 were consistently used to mark focus across varieties of South African English, while duration effects were found in some varieties. [2] (for Minnesotan and Southern Californian) and [11] (for Southern and Midland) provide cases where there are phonological or phonetic differences in focus marking between regional varieties of American English. A close examination of the phonetic realization of focus in Canadian English is necessary to shed light on the widespread assumption that Canadian English resembles American English in many aspects.

The current study investigated how various acoustic correlates are systematically modulated to distinguish constituents in broad focus, narrow focus, and given constituents in Canadian English. The main goal is to characterize the phonetic details of prosodic focus-marking and to shed light on the role of regional dialect as a potential source of phonological and/or phonetic differences between prosodic focus-marking in Canadian English and MAE.
2. Methods

2.1. Participants

Sixty-one speakers of English participated in a sentence production experiment, all undergraduate students at the University of Alberta. Out of the 61 participants, the recordings of 37 native speakers of Canadian English were used in the analysis (26 female and 11 male; median age: 19.5, age range: 17–34), excluding 18 non-native speakers, 4 native speakers who were balanced bilinguals, 1 native speaker of American English, and 1 native speaker whose audio was not recorded.

2.2. Procedures and stimuli

Participants were recorded in a sound-attenuated booth with a Countryman headset microphone (H6 Omni) connected to a Fostex field recorder (FR-2LE). They were given verbal and written instructions about the task. After silently reading a short paragraph describing a situation, they were asked to read answers to questions designed to elicit four focus conditions: broad focus (BF), subject focus (SF), verb focus (VF) and object focus (OF), as illustrated in Table 1. For example, in the SF condition, the subject constituent receives narrow focus while the verb and object are given. (Schwarzschild, 1999). All answers in the narrow focus conditions began with ‘no’, which was not included in the acoustic analysis. After the experiment, participants were asked to fill out an exit questionnaire about their language background.

Twenty-four items each occurred in the four focus conditions, balanced across four lists with a Latin Square design. All answer sentences had a simple SVO structure in which each constituent (i.e., S, V, O) consisted of three syllables, the second of which was always the lexically stressed syllable.

2.3. Acoustic and statistical analyses

Of a total of 2,664 constituents (37 participants x 24 items x 3 constituents) collected, 219 constituents were excluded from the acoustic analysis due to disfluency and 225 additional constituents were excluded from the f0 analysis due to voice quality issues. Seven acoustic correlates – word duration, mean intensity, maximum intensity, f0 range, f0 maximum, f0 minimum, f0 mean – were extracted from each constituent using Praat (7). The intensity measurements were scaled to 70dB and extracted in the nucleus of the stressed syllable. Manual inspection of the f0 values resulted in exclusion of 225 additional constituents from the f0 analysis that did not accurately represent the measured f0 values due to microprosodic f0 movements.

Linear Mixed-Effects Regression models were fit separately for these correlates as dependent variables in R ((4, 23)). Focus (BF, SF, VF, OF) and Location (S, V, O), as well as the interaction between them, were included as fixed effects and Participant and Item as random intercepts. We also tested if models were improved by including focus, location and their interaction as random slopes. All models except for f0 range, f0 maximum, f0 minimum, and f0 mean, included location as a random slope for both participant and item. Model comparisons to determine the random effect structure and inclusion of the interaction term were based on the AIC scores. Based on the resulting best models, all pairwise comparisons (via Tukey test) were computed and displayed in Figure 1 using the R package multcomp (117).

<table>
<thead>
<tr>
<th>Focus condition</th>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>BF (Broad Focus)</td>
<td>What’s going on?</td>
<td>Miranda is petting a lion.</td>
</tr>
<tr>
<td>SF (Narrow Focus on Subject)</td>
<td>Is Mark petting a lion?</td>
<td>No, Miranda is petting a lion.</td>
</tr>
<tr>
<td>VF (Narrow Focus on Verb)</td>
<td>Is Miranda distracting a lion?</td>
<td>No, Miranda is petting a lion.</td>
</tr>
<tr>
<td>OF (Narrow Focus on Object)</td>
<td>Is Miranda petting a lizard?</td>
<td>No, Miranda is petting a lion.</td>
</tr>
</tbody>
</table>

Table 1. Sample of the stimuli. The constituents in boldface (not shown for the participants) receive contrastive narrow focus, while given constituents are underlined.

3. Results

Best models of all acoustic measurements included a significant interaction between focus and location, indicating that focus conditions modulated the prosodic realization of the different constituents. Therefore, the effects of focus are discussed in terms of the results of the pairwise comparisons rather than the main effects. The results of pairwise comparisons for all measurements are shown in Figure 1 and are the basis for all p-values given below.

3.1. Word duration

Word duration significantly increased in the narrow focus condition compared to broad focus for Subjects (p<.001), Verbs (p<.001), and Objects (p<.01). The interaction between focus and location suggested that given Verbs and Objects decreased word duration compared to broad focus (p<.01 for given Verbs in the SF and OF conditions; p<.01 for the given Objects in the SF and VF conditions), whereas the word durations of Subjects with broad focus were not different from those of given Subjects (p=.66 for the VF condition; p=.79 for the OF condition).

3.2. Intensity mean & maximum

Mean intensity for Subjects and Verbs was lower in the narrow focus condition than the broad focus condition (p<.001 for Subjects; p<.05 for Verbs), Subjects and Objects with broad focus had higher mean intensity than given Subjects and Objects (p<.001 for all four conditions). Verbs in broad focus had higher mean intensity than given Verbs in the SF condition (p<.001) but not in the OF condition (p=.37).

As for maximum intensity, narrow focus Objects had higher max intensity than broad focus Objects (p<.001). Subjects (p=.95) and Verbs (p=.14) did not show broad and narrow focus distinction. The interaction between focus and location showed that broad focus constituents had higher maximum intensity than given Subject (p<.01 for the VF condition, p<.05 for the OF condition), Verb for the SF condition (p<.001), and Object (p<.001 for both SF and VF conditions).
Figure 1. Plots showing means, CI at 95% and results of pairwise comparisons of least-squares means across the constituents (S,V,O) for all acoustic measurements. The different focus conditions are denoted by color and shape of the dots. The lack of significant difference is indicated by sharing one or more of the same letter(s) (e.g., ‘a’ (BF) and ‘cde’ (SF) in word duration are significantly different, while ‘a’ and ‘ab’ are not).

3.3. f0 mean, maximum, & minimum

Broad focus Subjects had higher f0 values than narrow focus Subjects (p<.001 for mean, maximum, minimum). Verbs (p=.1 for mean and maximum and p=.64 for minimum) and Objects (p=.98 for maximum, p=.99 for minimum, p=.97 for mean) did not show a significant distinction between broad and narrow focus. Broad focus Subjects also had higher f0 maximum than given Subjects (p<.001 for both VF and OF conditions) and higher f0 mean than given Subjects (p<.01 for the VF condition, p<.05 for the OF condition), but no effect of focus was shown in f0 minimum (p=.35 for the VF condition, p=.32 for the OF condition). Broad focus Verbs had higher f0 maximum, minimum, and mean (all at p<.001) in the SF condition than in the OF condition (p=.88 for maximum, p=.99 for minimum, p=.1 for mean). Broad focus Objects had higher f0 maximum than given Objects (p<.01 for the SF condition, p<.05 for the VF condition) but not in f0 minimum (p=.94 for the SF condition, p=1 for the VF condition) or f0 mean (p=.32 for the SF condition, p=.71 for the VF condition).

3.4. f0 range

The distinction between broad and narrow focus was not significant across the constituents (S: p=.43, V: p=.50, O: p=1). Broad focus Subjects or Verbs did not differ from given Subjects (p=.27 for the VF condition, p=.87 for the OF condition) or given Verbs (p=.39 for the SF condition, p=.22 for the OF condition). Broad focus Verbs were, however, produced with larger f0 range than given Objects (p<.001 for the SF condition and p<.05 for the VF condition).

4. Discussion

The results of the acoustic analysis inform the distinction between focused and given constituents. Constituents in the broad focus condition, when compared to those in the given condition, were produced with longer duration in Verbs and Objects, higher mean and max intensity across the constituents, larger f0 range in Objects, as well as higher f0 maximum across the constituents, higher f0 minimum in Verbs, and higher f0 mean in Subjects and Verbs. These significant effects and their directions of the effects across multiple acoustic correlates are in line with the previous findings on prosodic focus marking in English ([3, 10, 12, 22]).

On the other hand, the modulation of the acoustic correlates to distinguish between narrow and broad focus was more complex. First, the results concerning word duration are in line with previous findings on the broad and narrow focus distinction reported in MAE: narrow focus was marked by longer duration across the constituents ([8, 12, 16, 20]). Mean and maximum intensity, however, provided mixed results. While maximum intensity was higher in narrow focus Objects than broad focus Objects, similarly to what have been observed in MAE ([8]), mean intensity was lower in narrow focus Subjects and Verbs than broad focus Subjects and Verbs. Moreover, broad focus Subjects showed significantly higher f0 maximum, minimum, and mean values compared to narrow focus Subjects, whereas Verbs and Objects did not distinguish broad and narrow focus based on f0 maximum, minimum, and mean.

The lack of a systematic association between narrow focus constituents and higher f0 as well as the general lack of significant differences in f0 values between broad and narrow focus in Verbs and Objects are inconsistent with previous studies that found evidence for broad and narrow focus distinction via f0 in MAE ([8, 12]) and the studies that showed higher f0 being associated with a narrow domain of focus in other Germanic languages ([24] for Dutch; [5] for German). The implication of this inconsistency is that the way Canadian English speakers modulated the acoustic correlates - especially f0 and intensity - to distinguish broad and narrow focus did not align with the way MAE speakers did, as reported in previous studies. A direct comparison between Canadian English and MAE using the same methodology may provide further evidence as to whether regional dialect is the source of variation in the phonetic modulation of broad and narrow focus in these varieties of English. Nevertheless, the results of the current study suggest one possibility in which speakers of Canadian English rely more on temporal information to distinguish broad
and narrow focus than f0 or intensity, unlike other varieties of English ([6, 8, 18, 26]).

In addition, the current study offers a case of how a more comprehensive methodology may bring about a novel insight into the inconsistent findings in the literature. For example, the current study examined a series of acoustic correlates that have been argued to be involved in prosodic focus marking, unlike some of the previous studies that focused on one or two correlates at hand. Moreover, the design of the current study also provided a more thorough look at the broad and narrow focus distinction issue by comparing the broad and narrow focus distinctions in different locations of an utterance. Given that [8] compared broad and narrow focus only in the object position of the utterances, the current study extended our understanding of how prosodic focus marking interacts with location of constituents within an utterance.

In ongoing follow-up research, we are exploring the phonological realization of focus in Canadian English, in particular the inventory of pitch accent, as well as the effects of focus on the alignment of tonal targets with segmental landmarks.

5. Acknowledgements

We thank Devon Gouzolkol and Amy Frey for help with creating the stimuli, as well as Grace Hill, Claire Casault, Clem Wisniewski, Natasha Daley, Aylish Anglin, and Zilai Wang for help with data collection and annotation. This project is supported in part by funding from the Social Sciences and Humanities Research Council of Canada.

6. References