Open Library of Humanities

Phonetic and phonological aspects of prosodic focus marking in Canadian English

Jiseung Kim*, Department of Linguistics, University of Alberta, Edmonton, Alberta, Canada, jiseung1@ualberta.ca

Anja Arnhold, Department of Linguistics, University of Alberta, Edmonton, Alberta, Canada, arnhold@ualberta.ca

*Corresponding author.

Previous studies reported mixed findings regarding the details of the acoustic correlates of focus, as well as the relationship between information structure and pitch accent types in English. Moreover, previous studies showed that different varieties of English signal information structure differently. Therefore, we hypothesized that the way focus is expressed in terms of acoustic correlates and/or pitch accents in Canadian English would differ from 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 on different constituents (subject, verb, object). The speakers of Canadian English in the current study manifested some, but not all the acoustic effects of focus that have been reported in previous research on MAE and other varieties. Specifically, focus-induced lengthening and amplitude increase were observed but not enhancement of f0. Instead, focus-adjacent constituents were weakened in terms of duration, intensity and f0. As for the pitch accents produced by the same speakers, while the presence or absence of pitch accents was strongly associated with whether a constituent was focused or not, the different types of pitch accents did not correlate with information structure.

1. Introduction

It is well established that focus is prosodically marked in English. The components of information structure such as focus, topic, contrast, and information status (e.g., given, new, inferable) are recognized as major factors in shaping the intonation of English (e.g., Halliday, 1967; Jackendoff, 1972; Pierrehumbert, 1980; Gussenhoven, 1983) and many other languages (see overview in Kügler & Calhoun, 2020; see Krifka, 2018 on notions of information structure). Despite the substantial amount of past research, however, a closer look at the literature reveals inconsistencies and gaps in our knowledge of prosodic focus marking in English. To address some of them, the current study investigated how information structure is signalled by modulations in phonetic details and categorical variations of pitch accents in Canadian English. A variety of acoustic correlates of narrow contrastive focus and givenness was examined, and the types of pitch accents that occur in different locations of the utterance were analyzed.

In the remainder of this section, we first summarize the existing literature and highlight the need for further studies regarding three aspects of prosodic focus marking in English which our study was designed to address. Our analysis of various acoustic correlates in different focus conditions and locations puts to the test the assumption that longer duration, higher amplitude, and higher f0 are the primary acoustic consequences of focus despite the inconsistencies observed in empirical findings (cf. section 1.1). The pitch accent analysis aimed to shed light on the relation between focus and the different types of pitch accents, which appears to be less clear than previously thought (section 1.2). Moreover, by examining speakers from Western Canada, the current study tested the possibility that the phonetic and phonological details of prosodic focus marking systematically differ between Western Canadian English and Mainstream American English (section 1.3). The final two subsections discuss the current knowledge regarding prosodic focus marking in Canadian English and explain the hypotheses of the current study.

1.1. Acoustic correlates of focus

Previous studies have shown duration, f0, and intensity to be the primary acoustic correlates of focus marking in English (Weismer & Ingrisano, 1979; Folkins et al., 1976; Atkinson, 1973; Breen et al., 2010; Cooper et al., 1985; Eady et al., 1986; Eady & Cooper, 1986; Kochanski et al., 2005; O'Shaughnessy, 1979; Pell, 2001; Roettger et al., 2019). However, there is a notable variety in the reported findings, which is likely at least partially due to studies differing with regard to investigated dialects and populations, as well as methods. In particular, it is likely that the use of different subsets of acoustic correlates with varied definitions examined in these studies likely have contributed to the mixed findings, making it difficult to evaluate how focus effects systematically manifest in different acoustic correlates. Also, the effects of lexical stress and focus were not clearly delineated or controlled in many of the earlier studies, partly because lexical stress and focus placement are closely related (Turk & White, 1999). Therefore, the current study

aimed to examine the effects of focus on a variety of acoustic correlates and build on the previous findings in different varieties of English to gain a more precise understanding of how focus is produced in a single dialect, Canadian English.

Focus-induced lengthening has been consistently observed in previous studies, but there are substantial inconsistencies regarding temporal modulation of words or segments preceding or following focus. For instance, Cooper et al. (1985) and Eady & Cooper (1986) reported lengthening of the focused item without shortening of adjacent words, while Weismer & Ingrisano (1979) and Folkins et al. (1975) found post-focal shortening in addition to focus-induced lengthening on the focused item. Eady et al. (1986) replicated this post-focal shortening effect, but only in shorter sentences (5–7 syllables) and not in longer sentences (10–12 syllables) used in the experiment. Pell (2001) found the post-focal shortening effect in both short (6 syllables) and long (10 syllables) sentences. In terms of pre-focal shortening, Pell (2001) observed pre-focal shortening in long sentences, contrary to studies that did not find shortening preceding focus (Weismer & Ingrisano, 1979; Cooper et al., 1985; Eady & Cooper, 1986).

Intensity has been argued to be another primary acoustic correlate of focus, but this is partly based on earlier findings that greater intensity is associated with lexical stress in English: Fry (1955), Lieberman (1960), and Beckman (1986) observed an increase in the amplitude of the stressed syllable compared to unstressed syllables, and Turk & Sawusch (1996) showed that the contribution of intensity to predicting lexical stress in perception was significant, although its independent contribution in the absence of duration information was small. Kochanski et al. (2005) provided evidence for loudness as a better predictor for detecting prominent syllables than f0, using a corpus of spontaneous speech in seven dialects of British and Irish English. This finding is in contrast to previous studies that found f0 to be a major correlate for prominence. For instance, Xu & Xu (2005) found an expansion of f0 range to be associated with focus and a compression of f0 range with post-focal material, suggesting that f0 makes an independent contribution to the production of focus.

In addition, there is no clear agreement concerning the phonetic details of focus effects on f0. Many previous studies observed higher f0 to be associated with focus, followed by a sharp drop in f0 (Atkinson, 1973; O'Shaughnessy, 1979; Cooper et al., 1985; Eady & Cooper, 1986; Eady et al., 1986; Pell, 2001). However, how and where f0 was measured (e.g. mean or peak, syllable or whole word) and what constitutes a higher f0 also differed between studies, with some reporting a paradigmatic change, e.g., comparing the realization of a word in narrow focus to the realization of the same word in a broad focus sentence, and others a syntagmatic change, e.g. comparing a focused word to the pre- and post-focal words in the same sentence. For example, Atkinson (1973) described a lowering of f0 after a syllable in an emphasized word and discussed how the f0 fluctuation makes the syllable stand out from the rest of the utterance. He also found that this lowering is sometimes accompanied by a higher f0 (peak and mean) of

the prominent syllable compared to the same syllable in the same utterance produced without an emphasis on the word. On the other hand, O'Shaughnessy (1979) only stated that emphasis results in larger f0 obtrusions in the emphasized syllables compared to the other syllables in the utterance. Later studies such as Cooper et al. (1985), Eady et al. (1986) and Pell (2001) compared the mean f0 of the focused item with the mean f0 of the corresponding item produced in the neutral condition.

Nevertheless, a post-focal f0 decrease has been observed relatively consistently compared to an f0 increase on a focused item. In particular, Cooper et al. (1985) argued that this post-focal f0 drop, rather than the increased f0 on the focused words, is the major effect of focus in sentence-initial and sentence-medial positions (also see Sánchez-Alvarado, 2020; Arnhold, 2021), whereas Eady et al. (1986) observed a lack of post-focal f0 drop when there was an additional focused word later in the utterance (i.e., in a double-focus condition), which appears to undermine the claim. Eady et al. (1986) also reported a higher f0 peak in the sentence-initial focused word, which was not found consistently in the previous studies (Cooper et al., 1985 with f0 peak; Eady & Cooper, 1986 with f0 peak; Sánchez-Alvarado, 2020 with f0 range; Arnhold, 2021 with f0 peak). Moreover, pre-focal f0 mean lowering shown in Atkinson (1973) and Eady & Cooper (1986) was not found in Pell (2001).

More recently, Breen et al. (2010) examined a relatively large number of acoustic correlates to investigate whether and how they are employed to systematically distinguish between focus on different constituents (subject, verb, object), focus breadths (narrow object focus vs. broad focus), and types of focus (contrastive vs. non-contrastive) in American English. The results of a series of production experiments showed that the focused constituent was signalled with longer duration, higher f0 mean and maximum, and higher maximum intensity, compared to pre- and post-focal words. In addition, objects with narrow focus were marked by longer duration, higher f0 mean and maximum, and higher maximum intensity compared to objects with broad focus. Interestingly, subjects and verbs with contrastive focus were produced with lower f0 mean and maximum than those with noncontrastive narrow focus, contrary to previous studies that found higher f0 to be associated with contrastive focus (Ladd & Morton, 1997 for British English; Ito & Speer, 2008 for American English).

It is crucial to acknowledge that these previous studies examined different varieties of English using different tasks to elicit the speech data, which has likely contributed to the mixed findings on prosodic focus marking. In an attempt to address the resulting contradictions and gaps in the literature, the present study examined the prosodic marking of focus in Canadian English, taking into account a broad range of acoustic correlates and directly comparing realizations of broad focus, narrow contrastive focus and given (or background) information in the sentence-initial, -medial and -final positions (see 1.5 for the hypotheses).

1.2. Pitch accents and focus

Information structure is also phonologically signaled by pitch accents. Studies on pitch accents or focus in American English suggest that a H* (or L+H*) accent is associated with new information (or contrastive focus) and a L* or absence of pitch accent (or de-accenting) is associated with given information. For example, Pierrehumbert & Hirschberg (1990) argued that new information is rendered salient by a H* while given information does not bear a pitch accent or receives a L*. They also suggested that a L+H accent marks the salience of a scale: a L*+H accent signals a lack of speaker commitment, while a L+H* signals that the speaker asserts that the accented item should be believed over a possible alternative. In other words, under this account, a L+H* can be characterized as a contrastive accent. Similarly, Steedman (2008) characterized H* accents as signalling uncontentious rhemes (a notion that can be roughly equated with focus or new information) and L+H* as uncontentious themes (a notion corresponding to topic or givenness). Welby (2003) showed that listeners are indeed sensitive to the presence or absence of a pitch accent when perceiving the information structure of an utterance. Theoretical accounts also generally agree that the representation of information structure is closely related to the location and type of pitch accent (Büring, 2007; Selkirk, 1995; Truckenbrodt, 1995).

However, empirical studies provide evidence that the relationship between information structure and the pitch accent types is less straightforward than previously considered (Terken & Hirschberg, 1994; Ito et al., 2004; Hedberg & Sosa, 2007; Chodroff & Cole, 2019). Ito et al. (2004) found that while the L+H* pitch accent was more likely to appear in contrastive contexts than non-contrastive contexts, it did appear in non-contrastive contexts in both new and given conditions of information status. De-accenting was also not exclusively found in the given condition. Chodroff & Cole (2019) argued for a probabilistic relationship between information structure and the type of pitch accent on the target word: Given information was more likely to appear with a L* or no pitch accent, whereas new and contrastive information were more likely to appear with H* or L+H*. Their results also showed that, similarly to what Ito et al. (2004) reported, (L+)H* pitch accents were observed on target words in the given condition, and new and contrastive target words with L* or no pitch accent were also observed (also see Grice et al., 2017, for a similar conclusion on German).

The current study investigates how different pitch accents are associated with information structure in Canadian English by analyzing the presence/absence, as well as the types of pitch accents occurring in broad focus, narrow contrastive focus and on pre- and post-focal given constituents. We hypothesize that H* and L+H* are more likely to appear in the items in contrastive condition than given condition, while L* and the absence of an accent are more likely to be observed in items in given than contrastive condition. In addition, the possibility remains that the division between the types of pitch accents occurring in contrastive/given condition is not clear-cut, such that some contrastive constituents do not appear with (L+)H* or some given items appear with pitch accents that are not L*.

1.3. Dialectal variations in focus production

Dialectal variation has been argued to be one of the factors that systematically modulate acoustic correlates of focus, which can lead to phonological and/or phonetic differences. For example, differences pertaining to pitch accent inventories and tonal alignment in the varieties of Indian English (Maxwell & Payne, 2018), different pitch accent types (i.e., rising vs. falling) associated with focused elements between Indian and British English (Féry et al., 2016), and differences in the pitch accent alignment between Scottish Standard English and Southern British English (Ladd et al., 2009) have been reported. In addition, O'Reilly et al. (2010) examined f0 values and contours to investigate how broad and narrow contrastive focus are realized in Donegal varieties of Irish (Gaelic) and Irish English. The results suggest that contrastive focus is produced with a large f0 excursion and post-focal f0 lowering compared to broad focus in Irish English, as has been reported previously for other varieties of English (cf. section 1.1). However, there are some characteristics that do not align with some of the earlier results reported for different varieties of English: f0 lowering before contrastive focus was observed (contra. Xu & Xu, 2005 for American English) and broad focus was associated with a rising (L*H) pitch accent (contra Chen et al., 2007 for Southern British English).

Turning to the regional dialects of American English, Arvaniti & Garding (2007) showed evidence that phonological and phonetic differences in focus production may exist between Minnesotan English and Southern Californian English. The use of H* and L+H* in their Southern Californian speakers was modulated by the number of iterations of the target word, whereas the Minnesotan speakers used L+H* irrespective of the number of repetitions, suggesting that H* may not be used in Minnesotan English. In addition, the alignment of pitch accents investigated in this study was generally later in Southern Californian speakers than Minnesotan speakers. Arvaniti & Garding (2007) also noted other differences in phonetic details of pitch accents between their findings and some earlier studies, which the authors argued might be due to the different varieties of American English. For instance, the L tone target of the L+H* was found to be aligned within the accented syllable in their data, whereas it was observed to be aligned with the preceding syllable in earlier studies (Pierrehumbert, 1980; Beckman & Pierrehumbert, 1986). A similar cross-dialectal phonological difference in pitch accent inventories was reported in Clopper & Smiljanic (2011) as tentative findings: A relatively stronger preference for L* pitch accents was observed for Southern American English speakers than Midlands American English speakers.

Previous studies on the relationship between different pitch accents and information status also showed cross-dialectal differences in the interpretation of pitch accents. Chen et al. (2007) investigated the role of four nuclear pitch accents in the processing of information status in Southern British English, building on previous findings in German (Baumann et al., 2006) that different pitch accent types are associated with different types of information status. The authors of this eye-tracking study found that speakers of Southern British English associated fall (H*L) and

rise-fall (L*HL) pitch accents with new information, while rise (L*) pitch accents were associated with given information. This differs from how listeners of American English associate different pitch accents with information status. In an eye-tracking experiment, Watson et al. (2008) found that L+H* created a strong bias towards contrastive (and given) information, whereas H* created bias towards both contrastive and new information, suggesting that interpretations of H* and L+H* pitch accents in American English are neither mutually exclusive nor variations of a single pitch accent category.

Empirical findings thus established that different varieties of English realize focus differently. However, Canadian English has been assumed to be a subgroup of American English (Bloomfield, 1970) or intermediate between American and British English (Canepari, 2010), and little attention has been given to whether and how Canadian English may differ from the varieties of American English in terms of prosodic aspects (e.g., Wagner & McAuliffe, 2019, treat their Canadian English data as broadly representative of English). Moreover, the prosodic characteristics of Canadian English are usually not considered in the phonetic or phonological literature describing Canadian English (Chambers, 1973, 1991, 2006; Labov et al., 2005; Sundara, 2005; Boberg, 2008). A few exceptions gave minimal and/or impressionistic descriptions (Halford, 2007; Canepari, 2010), described a distinct dialect (Clarke, 2010 for Newfoundland English), or did not aim to address prosodic aspects specific to Canadian English (DePape et al., 2012). Some studies have investigated a particular phenomenon concerning phrase-level intonation called High Rising Terminal (HRT), also known as uptalk (Lacey et al., 1997; Shokeir, 2008; Sando, 2009; Ouafeu, 2009; also see e.g., Fletcher et al., 2005 on HRTs in other varieties of English). While the differences in Canadian and American English in vowel production have been relatively well studied (Joos, 1942; Clarke et al., 1995; Hagiwara, 2006), the prosodic differences between these two broadly defined varieties of English have been rarely investigated.

1.4. Focus in Canadian English

Only a few studies provide information on how focus is produced or perceived in Canadian English, although focus marking was not always the primary topic of the study. The primary goal of DePape et al. (2012) was not to investigate focus production in Canadian English, but rather to show how adults with autism spectrum disorder and varying levels of language functioning use prosodic features to mark information status. Nevertheless, the six speakers in the control group provided insight into how focus is marked in Canadian English: larger f0 falls were observed in the focused word regardless of the location of the word in an utterance (i.e., subject or object), and larger f0 rises in the focused word in sentence-final positions compared to given topics in the same position. Patience et al. (2018) examined the first pitch accents of statements and questions using a sentence repetition task. They found that H* is the dominant pitch accent type in short SVO sentences (e.g., *Bobby looked for his football*), taking up nearly 80% of all first pitch accents, while L*+H is also observed (approximately 15%).

Arnhold et al. (2020) is the only study to our knowledge that directly investigated how regional accent influences the perception of information structure in Canadian English. Experiment 1 was an eye-tracking study: Listeners of Canadian English were given auditory instructions spoken in Standard Southern British English and their eye movements were monitored while they moved a computer mouse to click and move objects presented on the monitor. The target words in the instruction sentences varied in pitch accent type (fall, rise, no accent) and information status (given, new). The results showed that Canadian listeners associated falling accents with new information, and interpreted unaccented words as given, just as British listeners responding to the same items (Chen et al., 2007). Crucially, the rising accents were not clearly associated with either given or new information. This finding, which is in contrast with the results of Chen et al. (2007) for British listeners, suggests that the way Canadian English uses pitch accents to mark information status differs from Standard Southern British English. Given the phonological differences observed between Canadian English and British English, one may assume that there are also phonetic differences in how the acoustic correlates of focus are employed in these varieties of English.

1.5. Hypotheses of the current study

We investigated the prosodic marking of information structure in Western Canadian English by analyzing a scripted speech corpus of 37 native speakers of Alberta and British Columbia English collected through a speech production experiment, which we compared to previous research on other varieties of English. Our main point of comparison was Mainstream American English (MAE), as it has been the subject of most previous research and due to the common assumption that Canadian English is vastly similar to MAE. Given the findings that suggest information structure is marked differently in different dialects of English, we hypothesized that the phonetic and phonological details of focus marking in Western Canadian English are different from those previously observed for MAE. By examining a variety of acoustic correlates to broad focus, narrow focus and given information, the current study aims to show whether and how they are modulated in a systematic way to mark information structure. By annotating and analyzing the types of pitch accents, another goal of the current study is to establish a core pitch accent inventory (limited to declarative sentences) in Canadian English and to test the relationship between information structure and the presence and types of pitch accents. Based on previous findings for other varieties of English, particularly MAE, the following hypotheses can be formulated:

H1. Acoustic marking
  a. Constituents with narrow contrastive focus have longer duration, higher intensity and higher f0 than given and broad focus constituents.

    b. Pre- and post-focal given constituents have shorter duration, lower intensity and reduced f0 compared to broad focus constituents (may only hold true post-focally).

H2. Pitch accents

    a. Constituents with narrow contrastive focus are accented significantly more often than given and broad focus constituents, and frequently carry H* or L+H* accents.

    b. Pre- and post-focal given constituents are accented significantly less often than broad focus constituents, and they carry L* accents where they are accented at all.

2. Methods

2.1. Participants

Sixty-one speakers of English participated in a sentence production experiment. They were undergraduate students recruited through the University of Alberta linguistics subject pool. Out of the 61 participants, the recordings of 37 participants who identified themselves as native speakers of Canadian English were used in the analysis, after 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. Of the 37 participants, 26 were female and 11 were male. All were from Western Canada, with 35 identifying various locations in Alberta as the place they primarily lived growing up and 2 identifying locations in British Columbia. Their median age was 19.5, ranging from 17 to 34. The study was approved by the Research Ethics Board 2 of the University of Alberta (study ID: Pro00066772).

2.2. Procedure

Participants gave written consent at the beginning of the experiment. Research assistants verbally explained the task to the participants and gave them time to read the written instructions. Participants were seated in a sound-attenuated booth with a monitor displaying the stimuli and were recorded with a Countryman headset microphone (H6 Omni) connected to a Fostex recorder (model FR-2LE) at a sampling frequency of 48,000 Hz. Participants completed two practice trials before proceeding to the main experiment. After the experiment, participants were asked to fill out an exit questionnaire containing questions about their language background and other demographic information.

2.3. Stimuli

During each trial, participants were given a short paragraph describing a situation, followed by a question and an answer about the situation. The questions were designed to elicit broad or narrow focus on different constituents (Subject, Verb, Object) in the answers. An example of a context paragraph and the Q&A pairs in different focus conditions is given in **Table 1**. The

context and question were presented aurally and in writing. Participants were asked to read the answers out loud as if they were in that situation. All answers in the narrow focus conditions began with "no", which was not included in the acoustic analysis.

Twenty-four answers each occurred in the four focus conditions – broad focus (BF), subject focus (SF), verb focus (VF), object focus (OF) – balanced across four lists with a Latin square design. Each list also contained fifteen filler trials. All three constituents are labeled as BF in the broad focus condition. However, for the three narrow focus (NF) conditions, only one constituent in each answer received narrow focus and the other two constituents without focus were always given. Therefore, Subjects received narrow focus in SF, but Subjects were given in VF and OF, where either the verb (in the VF condition) or object (in the OF condition) received narrow focus.

Word sequences containing an article or possessive pronoun in the case of subjects and objects (e.g., *the woman, her photos*) or consisting of an auxiliary and a main verb (e.g., *was moving*) were considered as single constituents in the analysis. All subject, verb and object constituents consisted of three syllables, the second of which was always the lexically stressed syllable (e.g., [mɪˈɹæn.də] *Miranda*). All materials were also chosen to consist of voiced sounds as much as possible. Read speech was used in this study in order to exclude non-prosodic strategies for marking information structure such as pronominalization and ellipsis, and to ensure that we obtained comparable utterances suitable for phonetic measurements in all focus conditions.

Context paragraph	You are in a zoo with a group of friends. Your friend Liam has wandered away from the rest of the group, so you call him to tell him about something amazing that is happening. The quality of the call isn't very good, so he asks:	
Focus	**Question**	**Answer**
Broad Focus (BF)	What's going on?	Miranda is petting a lion.
Narrow Focus on Subject (SF)	Is Mark petting a lion?	No, **Miranda** is petting a lion.
Narrow Focus on Verb (VF)	Is Miranda distracting a lion?	No, Miranda is **petting** a lion.
Narrow Focus on Object (OF)	Is Miranda petting a lizard?	No, Miranda is petting a **lion**.

Table 1: Examples of the context paragraph and the corresponding question-answer pairs. The words in boldface receive corrective narrow focus. They were not printed in boldface for the participants.

2.4. Acoustic analysis

All target utterances were manually segmented in Praat (Boersma & Weenink, 2022). Segmentation used auditory and the following visual cues, in order of importance where available: silent intervals, fricative noise, and the third and second formant. At the same time, f0 contours were inspected in pitch objects: The f0 maximum and minimum of each constituent were automatically identified and manually inspected in order to correct or exclude f0 maxima and minima that do not accurately represent the highest or lowest points in the pitch excursion due to microprosodic f0 movements. Octave jumps and other f0 measurement errors due to noise, etc. were removed from the pitch objects using the 'unvoice' button on the affected stretches. Based on these annotations, seven acoustic measurements were taken for each constituent with a Praat script (available at Arnhold, 2021): duration of the constituent (in ms), mean intensity (in dB; extracted in the middle 50% of the nucleus of the three syllables and scaled to 70dB), f0 range (f0 maximum – f0 minimum), f0 maximum, f0 mean, and f0 minimum of the constituent (all f0 measurements in semitones; st henceforth).

Of a total of 888 utterances (37 participants x 24 items), 69 were sorted out due to disfluency. The remaining 819 utterances contained 2,457 constituents (three each), of which 12 constituents were further removed due to internal pauses. Thus, 2,445 constituents underwent acoustic analysis for the duration and intensity measures. A total of 225 constituents were further excluded from the f0 analyses after manual inspection because of creaky voice and other issues making reliable f0 measurements impossible.

Linear Mixed-Effects Regression models were fit separately to the six measured acoustic correlates – duration, relative mean intensity (see section 3.1.2 for details of calculation), f0 range, f0 maximum, f0 mean, and f0 minimum – using the lme4 package (Bates et al., 2015) in R (R core team, 2021). The base model included the focus condition of the utterance (Focus: BF, SF, VF, OF), the constituent on which the measurement was obtained (Constituent: S, V, O), and an interaction between them as fixed effects. Participant and Item were included as random intercepts. We compared the fitness of the base model against more complex models that included Focus, Constituent, or their interaction as random slopes. Model comparisons to determine the random effects structure, as well as whether the interaction term should be retained in the fixed effects, were done using the ANOVA function to determine if there was a significant difference in fit. A more complex model was only selected when it provided a significantly better fit based on the AIC scores (Matuschek et al., 2017). All final models except those for f0 range, f0 maximum, f0 mean, and f0 minimum included Constituent as random slope for both Participant and Item. Based on the resulting models, pairwise comparisons were computed via Tukey test using the emmeans package (Lenth, 2018). The figures in the results section showing the estimates and the results of the pairwise comparisons in the form of compact letter display (Piepho, 2004) were produced using the multcomp package (Hothorn et al., 2008; Bretz et al., 2010). All differences reported below were significant and associated with a p-value smaller than 0.05.

2.5. Pitch accent analysis

A group of native speakers of Canadian English who did not participate in the experiment annotated the evaluated 819 utterances using the Tone and Break Indices (ToBI) transcription system originally developed for MAE (Beckman & Ayers, 1997). The MAE ToBI system was used to identify pitch accents of Canadian English, given the widespread assumption that Canadian English is similar to MAE in terms of speech prosody, which the current study aims to put to the test. The annotators were seven undergraduate students who received basic training for the ToBI transcription using existing training materials (Veilleux et al., 2006) and instruction from the second author. They annotated a pitch accent for each constituent unless clearly absent, in addition to the boundary tones at the end of each utterance. For categorizing different pitch accents, annotators were instructed to pay attention to both the shape and meaning of the pitch accents, e.g., whether they perceived a constituent as prominent or as implying a different pragmatic meaning from the example recordings of different pitch accents provided by Veilleux et al. (2006). For an illustration of average f0 contours for the most frequent accent types and a discussion of the correspondence between contours and classification, see section 3.2.

The inter-annotator reliability was assessed by calculating Cohen's Kappa value (Cohen, 1960) in which 0 indicates chance agreement (i.e., annotators selected the same categorical value by chance) and a positive value suggests that the agreement was not due to chance. A trained phonetician (first author) transcribed approximately 45% of the data (370 utterances) using the same MAE ToBI annotation conventions, blind to the original annotation. The inter-annotator reliability between the original annotation and the subset annotation amounted to the kappa value of 0.24, which is considered fair (Cohen, 1960). The rate of agreement, calculated as a percentage of matched annotation out of all 370 utterances, is 68%. When the inter-annotator reliability was calculated between the presence or absence of pitch accent, which formed the basis of the first GLMM that examined the likelihood of pitch accent annotation (accent vs. no accent), the kappa value increased to 0.49 or moderate agreement, and the rate of agreement increased to 79%. The kappa value for the comparison of H* vs. non-H* after excluding the 343 'no accent' constituents from either of the annotators was 0.43, which is considered fair according to Cohen (1960). In this comparison, the rate of agreement was 84%.

To examine the relationship between the information structural condition of a constituent and the pitch accent annotation in terms of likelihood of annotation for any pitch accent or specific pitch accent types, a series of binomial Generalized Linear Mixed-Effects Models (GLMM) was fit using the lme4 package (Bates et al., 2015) in R. The dependent variable for the first GLMM was the presence of pitch accent annotated for each constituent, which is coded as 1 if annotated and 0 if no pitch accent was annotated. The second GLMM, motivated by the frequent occurrence of H* accents across the focus conditions (see section 3.2), was fit for accented constituents only (i.e., excluding unaccented constituents) to examine the probability of the

pitch accent annotated for a constituent being a H* variation (i.e., inclusive of the downstepped H* and the H* with a delayed peak) versus any other pitch accent. The dependent variable was coded as 1 if the pitch accent was a H* variation and 0 for any other pitch accent. The base model for both GLMMs included Focus (recoded as BF, NF, GV)[1] and Constituent (S, V, O) as fixed effects. The first GLMM failed to converge when the interaction between focus and constituent was added, and therefore it was not included in the final model. The second GLMM included the interaction term. Participant and Item were included as random intercepts for both models. Any more complex random structure resulted in convergence failure and therefore was not used in the final models. The model selection process was the same as the linear mixed-effects models used in the acoustic analysis.

3. Results

3.1. Results of the acoustic analysis

All models showed a significant interaction between Focus and Constituent (see the fixed effects summary tables in section 1 of the supplementary materials). Therefore, in this section, we focus on the results of the pairwise comparisons between individual factor level combinations, which are based on the Linear Mixed Effects Models for each acoustic correlate. Results of these comparisons are presented with compact letter display. The output of the pairwise comparisons are given in section 2 of the supplementary materials.

3.1.1. Duration

Figure 1 summarizes the results of the pairwise comparisons for constituent duration (also see Tables 1 and 8 in the Supplementary Materials), with the compact letters shown above each plotted median value denoting whether the differences between the conditions were significant. That is, if two conditions share one or more letter(s), they are not significantly different, whereas two conditions that do not share any of the letter(s) are significantly different. For example, the duration of subjects in the SF condition is marked with 'cd', whereas the duration of subjects in the BF condition is marked with 'a'. The fact that they do not share a letter indicates that they differ significantly.

For additional clarity, the focus conditions without statistically significant difference are grouped together in grey boxes within each of the constituent panels, while the conditions that are significantly different from each other are be shown in separate boxes.

---

[1] Based on the pitch accent annotation data it was expected that, in the binomial GLMM, the constituents with narrow focus would be more likely to be annotated for a pitch accent than the given constituents which preceded or followed another constituent that received narrow focus. In order to clearly examine the probability of pitch accent annotation depending on the focus condition of a constituent, the focus conditions were recoded as BF (constituents with broad focus), NF (constituents with narrow focus, e.g., subjects in SF), and GV (given constituents, e.g., subjects in VF and OF).

For example, subjects (left panel of **Figure 1**) were significantly longer in the subject focus condition than the other conditions, and there was no significant difference among the other conditions for this constituent. Note that the grey boxes only highlight a lack of significant differences within constituents, whereas differences between constituents are only indicted by the letters. For example, in broad focus, subjects (labeled 'a') differed significantly from both verbs ('de') and objects ('eg'), while verbs and objects did not differ significantly from each other, as indicated by the fact that both of their labels contain the letter 'e'.

As **Figure 1** shows, subjects in the subject focus condition (SF, label 'cd'), which are narrow-focused, were significantly longer than broad-focused subjects (BF, label 'a'). The difference between the estimated marginal means (EMMs) of the conditions is 57.62 ms. The same pattern of narrow-focused constituents being significantly longer than broad-focused constituents holds for verbs ('fgh' for VF; 'de' for BF with an EMM difference of 50.38 ms) and objects ('h' for OF; 'eg' for BF with an EMM difference of 21.81 ms).

Constituents in narrow focus were also significantly longer than given constituents. Subjects in the VF and OF conditions (given while verbs and objects received NF, respectively) are marked with 'ab', which does not share a letter with the letter display ('cd') for subjects in the SF condition, where subjects received NF. The EMM differences between SF, where subjects were in narrow focus, and the other two conditions were 44.73 ms for the VF condition and 45.63 ms for OF. This pattern is again repeated with verbs ('bc' vs. 'fgh' with EMM differences of 73.89 ms comparing VF to SF and 74.10 ms comparing VF to OF) and objects ('cdf' vs. 'h' with EMM differences of 54.30 ms for OF vs. SF and 64.65 ms for OF vs. VF).
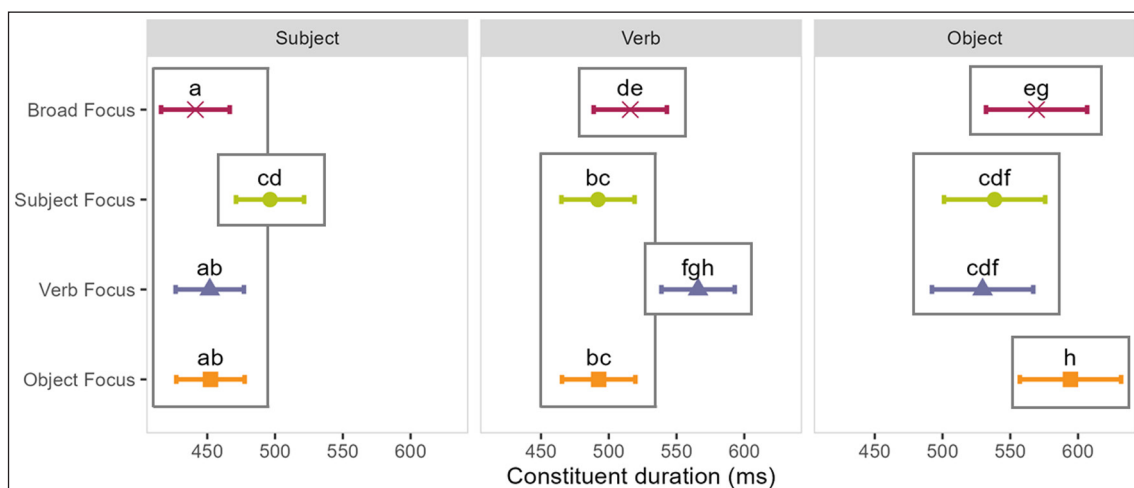

Figure 1: Durations of subjects, verbs, and objects in the four focus conditions. The letters and boxes reflect the results of the pairwise comparisons.

Turning to comparing the constituents with BF with the given constituents, where another constituent was narrow-focused in the utterance, subjects with BF did not significantly differ from given subjects ('a' vs. 'ab'), whereas verbs and objects with BF were significantly longer than given verbs and objects ('de' vs. 'bc' for verbs with EMM differences of 23.54 ms for BF vs. SF and 23.72 ms for BF vs. OF; 'eg' vs. 'cdf' for objects with EMM differences of 32.49 ms for BF vs. SF and 42.84 ms for BF vs. VF).

Lastly, comparing differences between the constituents, an effect of phrase-final lengthening was observed in a progressive manner: The phrase-final constituents, objects, were significantly longer than verbs in most conditions, and verbs in turn were significantly longer than subjects in all conditions.

3.1.2. Intensity

The intensity measure was calculated for each constituent as the difference between the mean intensity (dB) taken at the middle 50% portion of the vowel of the stressed syllable, which was always the second syllable of the constituent, and the average of the mean intensity (dB) values taken at the middle 50% portion of the vowels of the unstressed syllables preceding and following the stressed syllable. The larger the relative intensity, the greater the amplitude of the stressed syllable compared to the surrounding unstressed syllables within the constituent. **Figure 2** shows that the mean estimates were always greater than 0 in all conditions, meaning that the stressed syllable was louder than the unstressed syllables in a constituent regardless of the constituent or focus conditions.
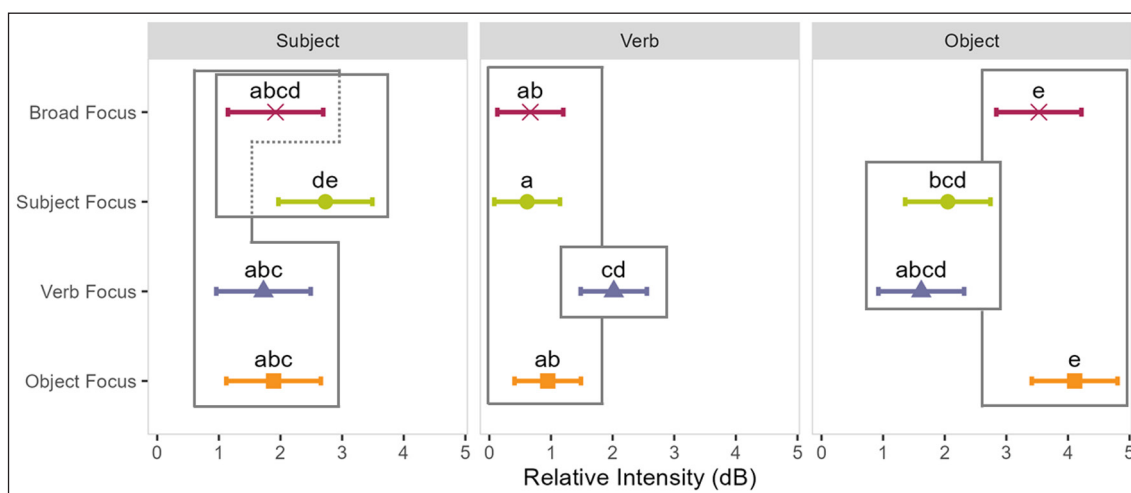

Figure 2: Intensity of subjects, verbs, and objects in the four focus conditions, calculated as the difference between the mean intensity of the stressed (i.e., second) vowel of the constituent and the average of the mean intensity of the preceding (i.e., first) and following (i.e., third) unstressed vowels of the same constituent. The letters and boxes reflect the results of the pairwise comparisons.

As shown in **Figure 2** (also see Tables 2 and 9 in the Supplementary Materials), the relative intensity of the narrow-focused constituents was significantly higher than for the given constituents, regardless of the constituent. That is, the degree to which the stressed syllable was louder than the surrounding unstressed syllables was greater when the constituent was narrow-focused than when it was given. The comparison between NF and BF, however, was less straightforward: The relative intensity values of narrow-focused verbs were significantly higher than in broad focus, but subjects did not differ significantly between BF and SF and neither did objects show a significant difference between BF and OF. This indicated that, in terms of relative intensity, the object constituent was prominent in the broad focus condition, which was in line with previous literature on focus projection (e.g., Gussenhoven, 1983; Féry & Samek-Ludovici, 2006; Wagner & McAuliffe, 2019). In addition, the object constituent was the only one showing significantly higher relative intensity in broad focus than when it was given (i.e., in SF and VF), while given subjects and verbs did not differ significantly from broad focus.

A syntagmatic comparison between the constituents in different focus conditions also showed that in BF, objects had higher relative intensity than both subjects and verbs, which did not differ significantly from each other. Further, the narrow-focused subjects and objects showed significantly higher relative intensity values than the two following (for subjects) or preceding (for objects) given verbs in SF and OF, respectively. In contrast, the relative intensity of the narrow-focused verbs was not significantly higher than the preceding subjects or the following objects in VF. In addition, subjects also had lower relative intensity than objects in OF, while in SF, the relative intensity of the verbs immediately following the narrow-focused subjects was significantly lower than the relative intensity of the following objects.

An evaluation of absolute mean intensity in the middle 50% portion of the vowel of the stressed syllable of each constituent (in dB, with all sounds scaled to 70dB before measurements) showed slightly different effects of the focus manipulation: The difference between narrow focus and given realizations of the same constituent observed here was not consistent, with only objects showing significant differences between narrow focus and both given conditions. Instead, analysis of this absolute measure indicated post-focal drops in intensity, with verbs having lower values in SF than in all other conditions and objects having lower values in SF and VF than in BF and OF. Finally, syntagmatic comparisons with the absolute measure also displayed lowering of intensity over the course of the utterance, with significant differences between the constituents appearing in all focus conditions (see Tables 3 and 10, and Figure 1 in the Supplementary Materials for details).

3.1.3. F0 range

As shown in **Figure 3** (also see Tables 4 and 11 in the Supplementary Materials), the f0 range did not significantly differ between narrow and broad focus for any of the three constituents. For example, the f0 range values of the subjects were not significantly different between the BF and

the SF conditions, and the same null effect was found for the f0 range values of the verbs and the objects. However, there were significant differences between the f0 range of the constituents in the narrow-focus condition and the given condition, as illustrated in **Figure 3**. For subjects, the f0 range in SF was significantly larger than the f0 range of the given subjects when either the verbs were focused or the objects were focused. Similarly, verbs with narrow focus had significantly larger f0 range values compared to verbs in the SF or the OF condition. Objects with narrow focus also showed significantly larger f0 range values compared to objects in the SF or the VF condition. Comparing broad focus and given realizations only showed significant differences for objects, which had smaller ranges in SF and VF than in BF.

Lastly, f0 range did not show a general decrease over the course of an utterance and few significant differences appeared between the constituents within a given focus condition.
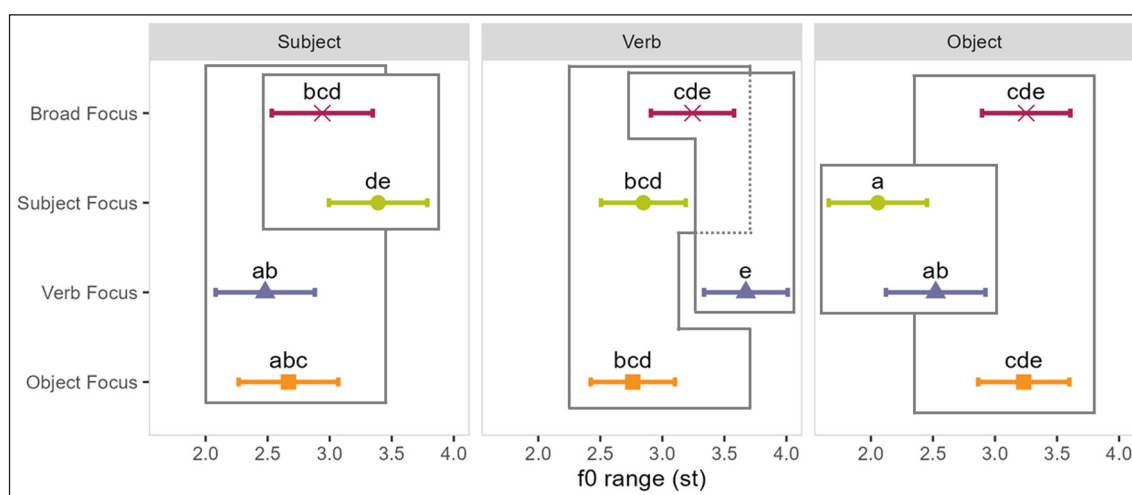

Figure 3: f0 range (in st) of subjects, verbs and objects in the four focus conditions. The letters and boxes reflect the results of the pairwise comparisons.

3.1.4. F0 maximum

No significant difference was found between the f0 maxima of verbs and objects in the BF and narrow focus conditions (VF and OF, respectively), as illustrated in **Figure 4** (also see Tables 5 and 12 in the Supplementary Materials). The f0 maximum of subjects in the BF condition was significantly higher than that in the SF condition, contrary to expectations. However, as expected, given subjects had lower maxima than subjects in BF. A similar pattern was found in the f0 maxima of given objects, which were significantly lower than those of objects in the BF and OF conditions. The f0 maximum of verbs in the SF condition was significantly lower than that of verbs in the BF and the VF conditions, which did not differ from that of verbs in the OF condition. That is, the presence of NF did not result in an f0 increase, but an f0 decrease was observed following a constituent with NF.

Finally, f0 maxima displayed a general downtrend over the course of an utterance. This resulted in significant differences between all constituents within a focus condition in all cases except for subjects and verbs in VF, i.e., a drop was only observed post-focally.
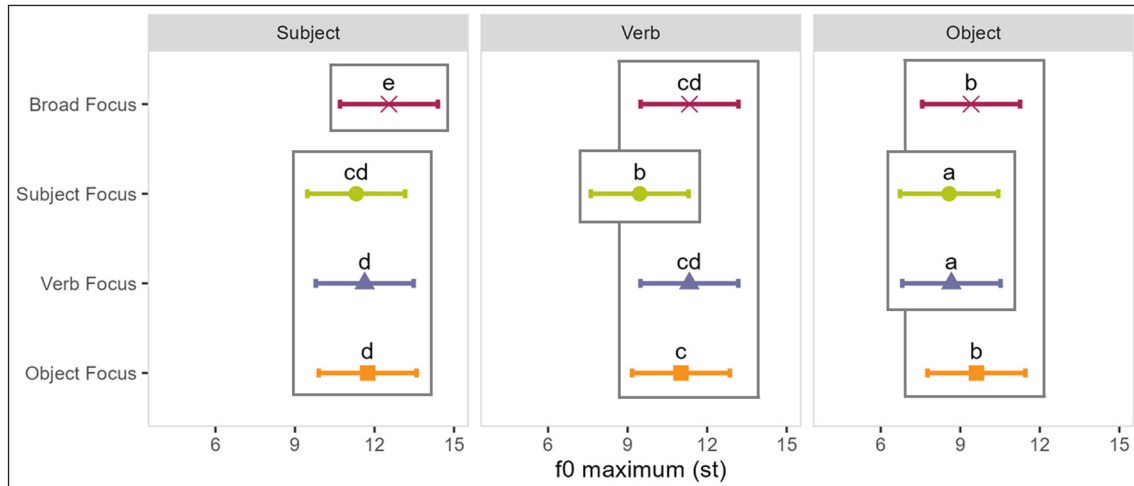

Figure 4: f0 maximum (in st) of subjects, verbs and objects in the four focus conditions. The letters and boxes reflect the results of the pairwise comparisons.

3.1.5. F0 mean

The f0 mean of subjects in the SF condition was significantly lower than in the BF condition, as shown in **Figure 5** (also see Tables 6 and 13 in the Supplementary Materials). The f0 mean of subjects in the SF condition was also significantly lower than in VF and OF. However, the f0 mean values of verbs and objects in narrow focus were significantly higher than those in the SF condition, while not significantly differing from the f0 mean in the other given condition (OF and VF, respectively). This suggests that a narrow-focused constituent was followed by a sharp drop in f0, particularly for subjects, while the constituent with the narrow focus did not show an increase of f0 itself. This generalization is further supported by directly comparing constituents within each focus condition: The f0 mean showed a decrease from subjects to verbs, and then from verbs to objects, regardless of focus. One exception was the f0 mean values of verbs and objects, which did not differ when subjects were narrow-focused. It is possible that the f0 level of the verbs following the narrow-focused subjects were already low enough to be close to the floor of the speaker's range, keeping the following element—objects—from manifesting a further decrease in f0.

3.1.6. F0 minimum

As shown in **Figure 6** (also see Tables 7 and 14 in the Supplementary Materials), the f0 minimum of subjects in the SF condition was significantly lower than in the BF and given conditions. For
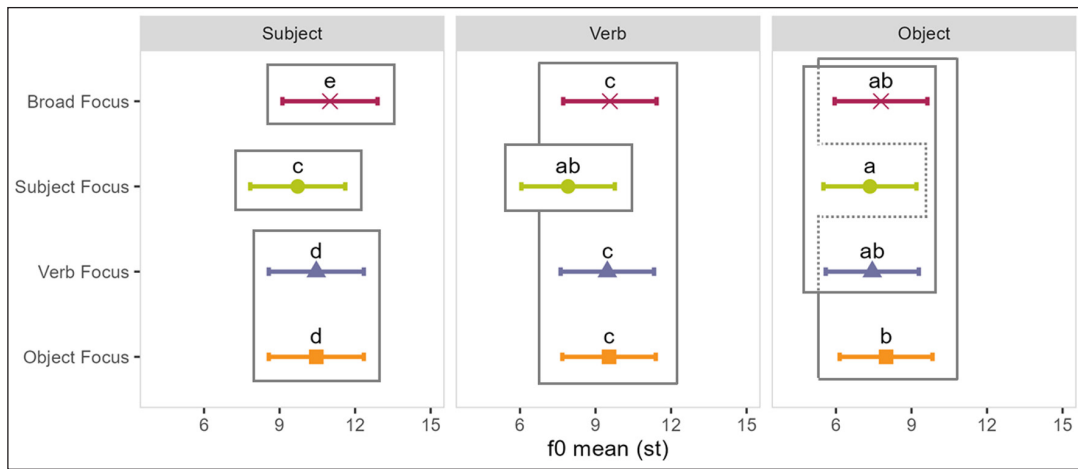
Figure 5: f0 mean (in st) of subjects, verbs and objects in the four focus conditions. The letters and boxes reflect results of pairwise comparisons.
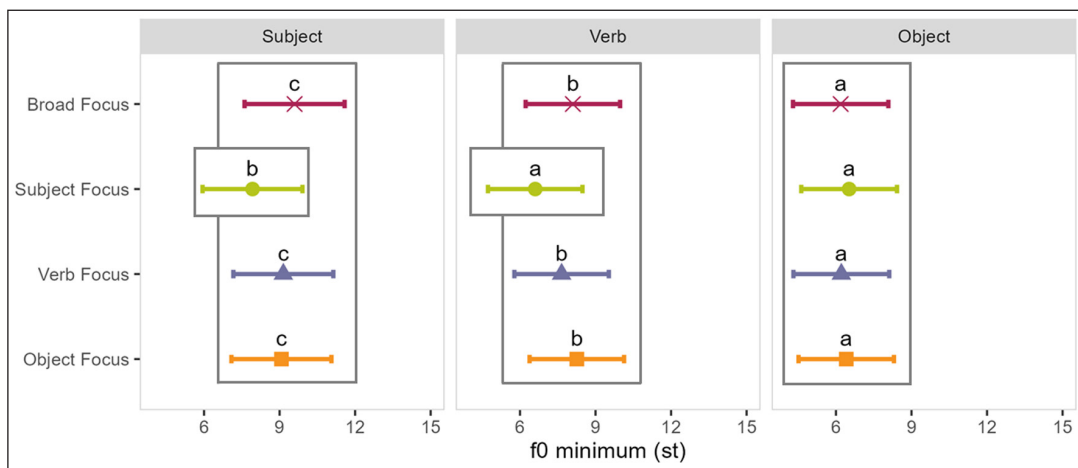

Figure 6: f0 minimum (in st) of subjects, verbs and objects in the four focus conditions. The letters and boxes reflect the results of the pairwise comparisons.

verbs, the f0 minimum in the VF condition was significantly higher than in the SF condition, although not significantly different from the BF and OF conditions. No significant difference was found for objects between the OF, BF and the given conditions, suggestive of a floor effect. Pairwise comparisons between the constituents within the same focus condition suggested that the f0 minimum of objects was significantly lower than that of verbs, which in turn was significantly lower than that of subjects, similar to the pattern observed in f0 means. The only difference was found, again, in verbs and objects in the SF condition, which did not differ significantly from each other. A post-focal f0 drop was observed in verbs after subject focus, but no focus-induced f0 increase was found on the focused constituent itself.

3.2. Interim summary of the results of the acoustic analysis

The acoustic correlates examined in the current study were systematically modulated by the focus conditions. The presence of narrow focus resulted in increased magnitudes of various correlates, most notably constituent duration, intensity and f0 range of the focused constituent relative to their realization in given conditions. The other f0 measures did not show enhancement effects due to narrow focus across the board: narrow-focused subjects were actually lower in f0 mean and f0 minimum than given subjects and comparable to given in f0 maximum. The f0 maximum, mean, and minimum of verbs were higher in the narrow focus condition than those in the post-focal given condition (i.e., subject focus) but not different from those in the pre-focal given condition (i.e., object focus). As for objects, the presence of narrow focus manifested as enhancement in some contexts (f0 mean compared to post-focal condition and f0 maximum) but not in the others (f0 mean compared to pre-focal condition and f0 minimum).

The comparison between narrow and broad focus revealed a complex pattern: while the durations of all constituents were longer under narrow focus than broad focus, only verbs showed enhancement of relative intensity under narrow focus compared to broad focus, and none of the constituents showed significant differences in f0 range between the narrow and broad focus conditions. With regards to the other f0 measures, objects did not show significantly differences between broad focus and narrow (object) focus, while narrow-focused subjects and verbs were even lower in f0 maximum, mean and minimum than subjects and verbs in broad focus, respectively.

Finally, comparing broad focus and given conditions showed that subjects showed lower f0 mean and f0 maxima in given conditions than in broad focus, while given objects had shorter durations, lower relative intensity, f0 ranges and f0 maxima than in broad focus. For verbs, durations were shorter in both given conditions (SF and OF) than in broad focus, while f0 maximum, f0 mean and f0 minimum were lower in SF compared to broad focus, suggesting clearer marking post-focally than pre-focally.

Overall, the hypotheses in H1 (see 1.5) were partially supported: narrow focus did not induce enhancement of all acoustic correlates in an across-the-board manner, although enhancement of duration, intensity, and f0 range in the narrow-focused constituents compared to given constituents were consistently observed. Given constituents differed not only from narrow focus, but also from broad focus realizations, though less extensively. As hypothesized, several measures showed consistent marking of givenness only post-focally.

3.3. Results of the pitch accent analysis

Figure 7 displays all constituents submitted to the pitch accent analysis, divided by the presence or absence as well as the types of pitch accents. The most frequent label was "no accent", i.e., unaccented constituents (35% of the data; 860 constituents). Among the accented constituents, the dominant pitch accent type was H* – a total of 625 constituents were annotated with a H*.

The number of the H* variations (i.e., including the constituents in which a delayed peak and/ or downstepping was additionally identified) added up to 1,160 constituents, which took up approximately 73% of all accented constituents and 47% of all constituents overall. Variations of the L+H* pitch accent accounted for approximately 14% (228 constituents) and variations of the L* accent accounted for 12% (183 constituents) out of all accented constituents. The least common pitch accents were L*+H (16 constituents) and H+!H* (1 constituent).

Figure 8 illustrates time-normalized average f0 contours for the three most common main accent types, i.e., variations due to a delayed peak and downstepping are combined into a single category. These were obtained from f0 measurements at 10 equidistant points on each constituent. Averages of these measurements, as well as error bars representing standard errors, appear in **Figure 8** for each constituent and focus condition by main accent type. Average contours are clearly distinct for the three accent types H* (and its variations), L* (and its variations) and L+H* (and its variations), as well as consistent across constituents, with H*-accents showing a small peak followed by a slight fall, L* accents displaying a dip in f0 and L+H* accents showing a low turning point (L+) followed by a rise to a peak (H*) followed by a small fall. The difference between the H* contours and the "no-accent" contours was less than prominent. Here, differences in perceived prominence leading to the respective classification (cf. section 2.5) may not be reflected in average f0 contours. Finally, **Figure 8** illustrates that average contours for all constituents and accent types were consistent across focus conditions, further suggesting that annotations reflect prosodic realizations rather than experimental conditions as such. Instead, some accents were never observed in certain experimental conditions – for example, no L+H* accent was observed for objects in the given conditions – i.e., SF and VF.
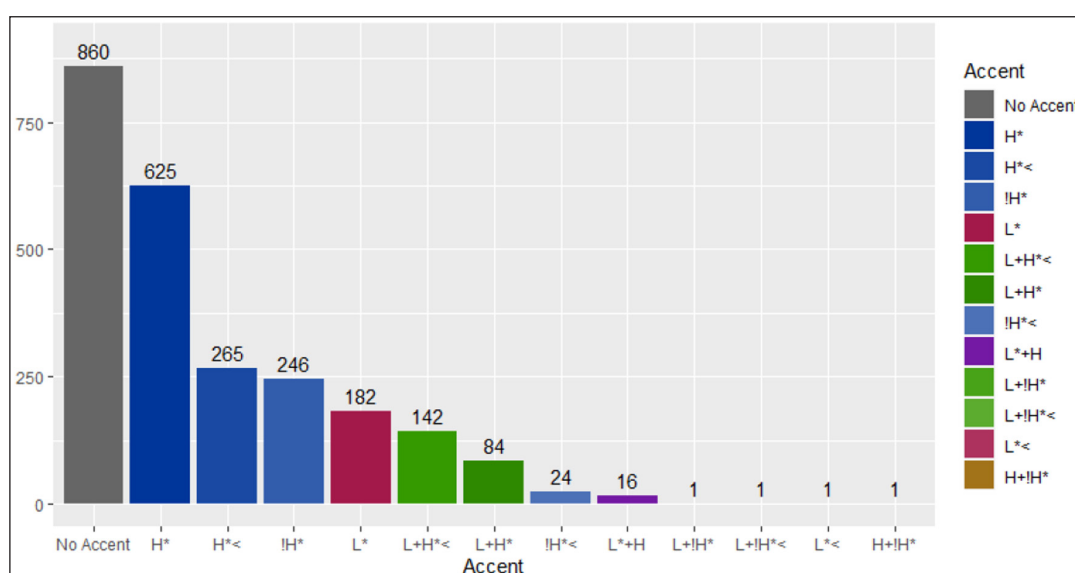

Figure 7: Number of constituents without pitch accent annotation and with different accent types.

To directly illustrate the correlation between the frequency of accent types and experimental conditions, **Figure 9** shows the occurrences of each main pitch accent type divided by focus condition and constituent. The most noticeable pattern was the high number of unaccented constituents (labeled "No Accent") in the given conditions across the constituents. It is also noteworthy that, while the H* pitch accent is the dominant type across the conditions, verbs with BF and VF are annotated with L* more frequently than subjects and objects in any condition. In addition, the L+H* accent was rare on verbs and objects, particularly in given conditions, but appeared relatively frequently on subject constituents in all conditions.
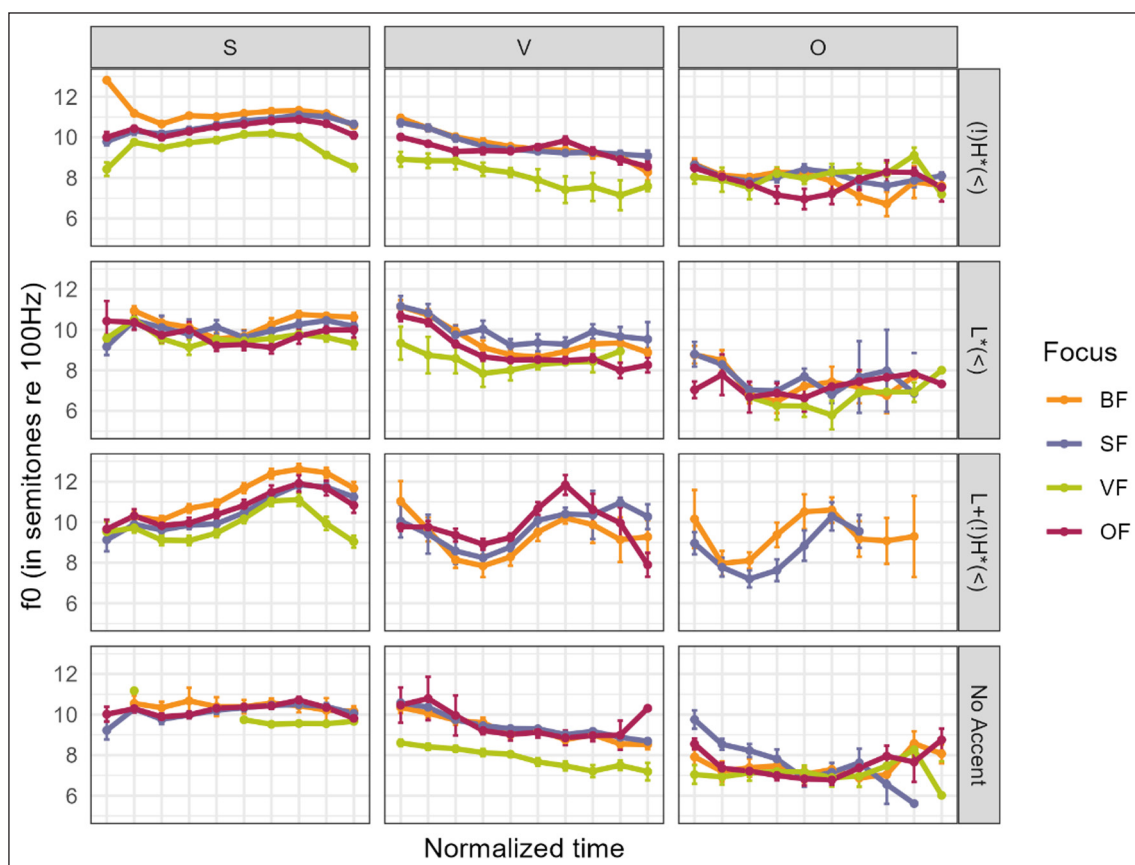

Figure 8: Averaged f0 contours of subjects, verbs and objects annotated with three most common pitch accents (the top three rows of panels) and averaged f0 contours of constituents with no accent (bottom row of panels) in different focus conditions.

The first GLMM examined the likelihood of pitch accent annotation (vs. no accent) depending on location and focus condition (see Tables 15 and 16 in the Supplementary Materials). Recall that focus was re-coded for the analyses of pitch accents to reflect the information status of the individual constituent as being in broad focus (BF), narrow focus (NF) or given (GV) rather than

Figure 9: Number of constituents without annotated pitch accents and with different accent types on subjects, verbs and objects by focus conditions.

reflecting the information structure of the whole utterance (BF, SF, VF, OF). The results showed significant main effects of focus and constituents: Constituents with NF were significantly more likely to be annotated as accented than constituents with BF, which in turn were more likely to be accented than given constituents. The low probability of pitch accent annotation for the given constituents corresponds to the high frequency of unaccented given constituents shown in **Figure 9**. Comparing the different constituents, subjects were significantly more likely to be accented than verbs, which were more likely to be accented than objects.

Focusing on the accented constituents in **Figure 9**, it appears that the majority of them was annotated with a H* variation, while far fewer constituents were annotated with a different type of pitch accent, regardless of the focus conditions and for all three constituents. The second GLMM therefore examined the probability of H* annotation compared to all other pitch accent types, excluding the unaccented constituents (see Tables 17 and 18 in the Supplementary Materials). **Figure 10** shows the results of the inverse calculation of the estimates of the fixed effects in the model (specified in the logit scale), with higher values indicating more frequent/probable H* accents than lower values. The interaction between Focus and Constituent was significant, and the results of the multiple comparisons using the Tukey test are indicated by the compact letter display. The results of the second GLMM showed that, overall, accented constituents showed a probability of H* annotation higher than 50% across the focus conditions and locations. The results of the multiple comparisons between the levels of Focus and Constituent showed that narrow-focused verbs were significantly less likely to be annotated with a H* ('a') than given
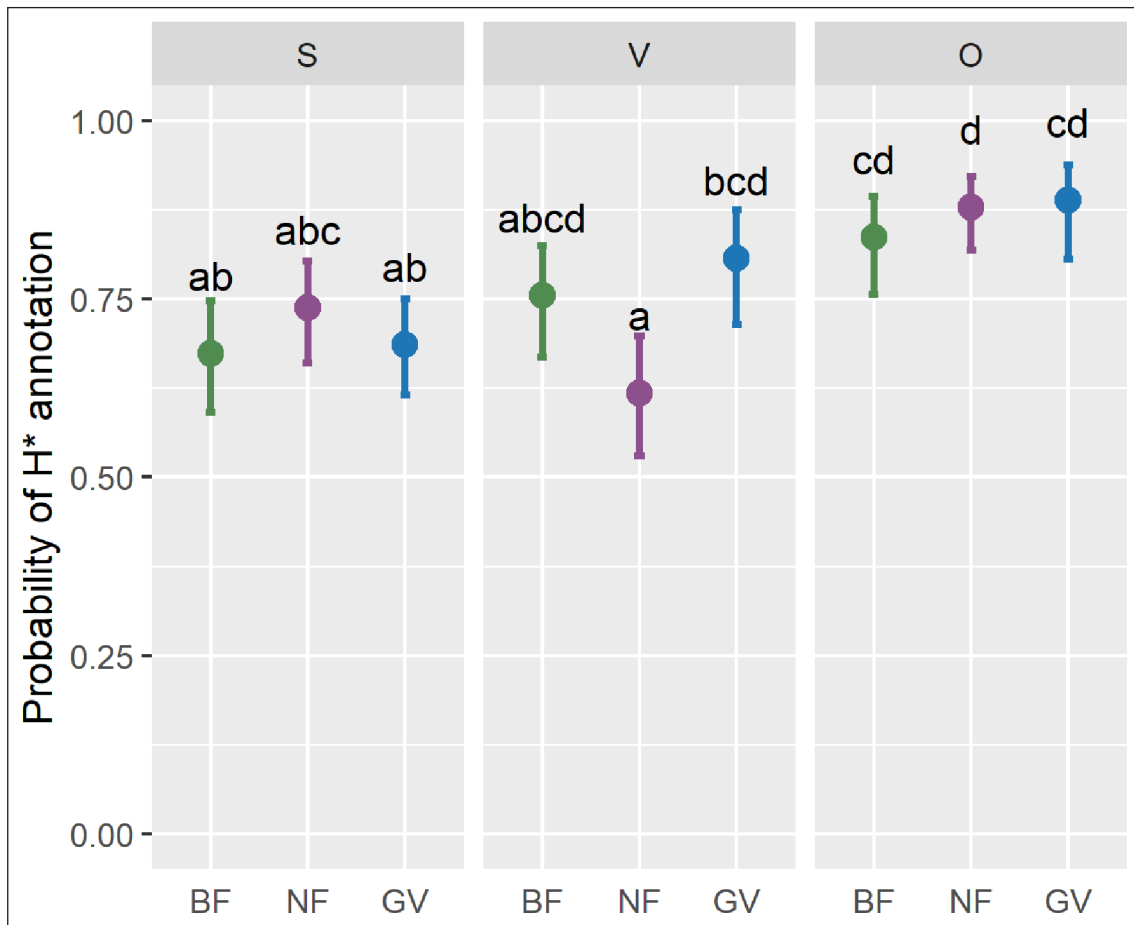
Figure 10: Probability of H* pitch accent annotation in different focus and location conditions.

verbs ("bcd") or objects in any focus condition ('d' or 'cd'). Moreover, objects in any focus condition were more likely to be annotated with H* accents than broad-focused subjects ('ab') or given subjects ('ab'). All other combinations of focus and location did not differ significantly from each other, suggesting an overall high probability of H* accents.

3.4. Interim summary of the results of the pitch accent analysis

The pitch accent analysis showed that the presence of narrow focus led to higher probability of the focused constituents being annotated as accented, compared to broad focus and given constituents. It also showed the prevalence of H* accent under the presence of broad or narrow focus in the variety of Canadian English examined in the current study. The average f0 contours in different focus conditions and constituents suggest that the f0 shapes differ between the broad pitch accent categories such as H* and L+H*, although further research is needed to evaluate how the different types of focus map onto the different f0 shapes. Overall, the hypotheses in H2

(see 1.5) were partially supported: narrow-focused constituents were most frequently annotated as H*, which was far more frequent than L + H* across all conditions. In addition, while given constituents were less likely to be annotated as accented than broad-focused constituents, L* accents were not consistently associated with the given, accented constituents.

4. Discussion

One of the main goals of this study was to examine the fine-grained phonetic modulations as a result of prosodic focus marking in Canadian English and to show how much they align with the previous findings in MAE in order to further delineate the effects of dialects in prosodic focus marking in English. The results of the acoustic analysis showed significant effects of focus on all constituents: All acoustic correlates except for f0 range were substantially modulated when a narrow contrastive focus was present. Importantly, the directions of the effect were often inconsistent with the current knowledge about prosodic focus marking in MAE. A summary of the results in **Table 2** indicates whether the results support or contradict the hypothesis regarding the acoustic marking of focus and givenness (H1) that (a) constituents with narrow contrastive focus have longer duration, higher intensity and higher f0 than (given and)[2] broad focus constituents and (b) pre- and post-focal given constituents have shorter duration, lower intensity and reduced f0 (may only hold true post-focally) compared to broad focus constituents (cf. section 1.5).

The duration row of the middle three columns in **Table 2** indicates that all three constituents were significantly lengthened under the effect of focus, replicating a well-established phenomenon in American English with speakers of Western Canadian English, focus-induced lengthening (Folkins et al., 1975; Weismer & Ingrisano, 1979; Cooper et al., 1985; Eady & Cooper, 1986; Eady et al., 1986; Pell, 2001; Breen et al., 2010; Chodroff & Cole, 2019). This result is also consistent with DePape et al.'s (2012) data that showed Canadian speakers using lengthening to mark focus (also see Arnhold, 2021). Another finding of the current study that is consistent with previous research is a post-focal shortening effect (Cooper et al., 1985; Eady et al., 1986; Arnhold, 2021), supporting H1b given in section 1.5. The durations of given verbs and objects, when following a focused constituent, were significantly shorter than broad focus verbs and objects. Pre-focal shortening was also observed in the given verbs that were significantly shorter than broad focus verbs when preceding focused objects (in line with Pell, 2001). There was no significant difference between the durations of given verbs undergoing focus-adjacent shortening, indicating that the degrees of shortening did not systematically vary depending on

---

[2] While (H1a) predicts consistent differences between narrow focus and both broad focus and given, **Table 2** only shows comparisons between narrow focus and the broad focus baseline for conceptual and expositional clarity. For example, there are no checkmarks for focus constituent enhancement in the rows for f0 range, which was higher in narrow focus than in given conditions, but not higher than in broad focus, for all constituents.

	Pre-focal weakening			Focused constituent enhancement			Post-focal weakening		
	S	**V**	**O**	**S**	**V**	**O**	**S**	**V**	**O**
Duration		✓Pre-OF		✓	✓	✓		✓Post-SF	✓Post-SF & Post-VF
Relative intensity					✓				✓Post-SF & Post-VF
F0 range									✓Post-SF & Post-VF
F0 maximum	✓Pre-VF & Pre-OF			!	!			✓Post-SF	✓Post-SF & Post-VF
F0 mean	✓Pre-VF & Pre-OF			!	!			✓Post-SF	
F0 minimum				!	!			✓Post-SF	

Table 2: Summary of results concerning the modulation of the acoustic correlates to focus and givenness. A checkmark (✓) denotes that the results supported H1a with respect to comparisons between narrow and broad focus, whereas an exclamation mark (!) denotes the results contradicted H1a (i.e., significant difference in the opposite direction). Significant weakening adjacent to constituents with narrow focus supporting H1b is indicated with a checkmark in corresponding cells, specifying the condition(s) where it occurred. An empty cell indicates neither enhancement nor weakening was found.

the pre-focal or post-focal contexts. Given the short lengths of the utterances used in the current study, a more fine-grained durational analysis may shed more light on the details of focus-adjacent shortening effects.

The results of the current study also supported previous studies in regard to post-focal f0 lowering. Post-focal verbs showed significantly lower f0 maximum, mean and minimum than broad focus verbs, and post-focal objects had significantly lower f0 range and f0 maximum than broad focus objects, regardless of whether the preceding narrow focus was on subjects or verbs. For American English, Couper-Kuhlen (1984) argued that contrastive focus had a sharp f0 fall, confirming a previous observation in O'Shaughnesy (1979). Cooper et al. (1985) also noted that the effect of focus on sentence-initial and medial words manifested in post-focal f0 lowering rather than increased f0 on the focused word itself (also see Sánchez-Alvarado, 2020). Eady & Cooper (1986) compared the f0 contours of questions and statements with sentence-initial focus and observed a low f0 following the focused word in statements but not in questions. For

Canadian English, DePape et al. (2012) showed larger f0 decrease for focus than given words in both the sentence-initial and final positions, and Arnhold (2021) showed that verbs and objects following subjects with narrow noncontrastive focus had significantly lower f0 maxima than those following subjects in broad focus. Our results replicate the findings of these earlier studies on North American dialects of English showing reduction of given information, in contrast to some other varieties of English that may not use the same strategy (Lim, 2004 for Singapore English; Gut, 2005 for Nigerian English; Ouafeu, 2007 for Cameroon English).

The relative intensity measure showed enhancement of stressed syllables in narrow focus compared to given realizations for all constituents, though the comparison between narrow focus and broad focus was only significant for verbs. This finding is consistent with previous production studies on English that showed increased intensity as an effect of focus (O'Shaughnessy, 1979 and Breen et al., 2010 for American English; Ouafeu, 2007 for Cameroon English). It also corroborated perceptual evidence for intensity as a robust cue to focus shown for MAE (Beckman, 1986; Turk & Sawusch, 1996; Kochanski et al., 2005). The current study further replicated the pre-focal weakening in terms of lower intensity in subjects preceding a narrow-focused objects shown in Breen et al. (2010) if the comparison was between subjects and objects in object focus sentences (syntagmatic comparison, as used by Breen at al.). However, note that the pre-focal subjects did not show significantly lower intensity compared to subjects in broad focus (paradigmatic comparison). An alternative evaluation of an absolute intensity measure (reported in detail in the Supplementary Materials) revealed post-focal intensity lowering consistent with what Arnhold (2021) found in the speech of a similar group of Western Canadian speakers: Verbs and objects following subjects with narrow focus had significantly lower intensity than those following broad focus subjects – and, in the current study, objects following narrow focus verbs additionally had lower intensity than broad focus objects (verb focus was not investigated in Arnhold, 2021).

On the other hand, some of the findings in the current study were not entirely consistent with previous findings on the effects of focus in MAE. Specifically, the subjects and verbs in narrow focus were marked with lowered f0 maximum, mean and minimum compared to broad focus, contradicting previous research suggesting that increased f0 marks focus. In addition, f0 maximum, mean and minimum of focused objects did not differ significantly from those in broad focus. This is inconsistent with the results of studies that found an association between increased f0 and focus in American English (Atkinson, 1973; O'Shaughnessy, 1979; Cooper et al., 1985; Eady & Cooper, 1986; Eady et al., 1986; Bartels & Kingston, 1994; Pell, 2001). The f0 range results are also not entirely consistent with previous findings. The current study did not find significant expansion of f0 range in focused constituents, unlike what Xu & Xu (2005) found for American English and Ladd & Morton (1997) found for British English. However, post-focal objects had significantly smaller f0 range than broad focus objects (with a lack of such effect in pre-focal constituents), which is in line with the suppressed f0 range of post-focal syllables (along

with a lack of pre-focal f0 range suppression) observed in Xu & Xu (2005). Again, our results are most in line with those reported for subject focus in Western Canadian English by Arnhold (2021), who observed larger f0 ranges due to lowered f0 minima, but no significant increase of f0 maxima in focused subjects compared to broad focus, while post-focal verbs and objects showed lowered maxima and reduced f0 ranges.

Overall, the effects of focus manifested via lengthening of the focused constituents, enhancing of f0 range and relative intensity focused compared to given items, and weakening of focus-adjacent constituents in terms of duration, absolute intensity and f0. Importantly, the focused constituents were reliably distinguished from the broad focus constituents via duration and, for verbs, intensity enhancement, whereas f0 parameters were not enhanced relative to broad focus to mark narrow focus. Rather, the focused constituents were even weakened in some contexts. In contrast, focus-adjacent reduction was reliably observed across the acoustic measurements (duration, absolute intensity, f0) and constituent locations (S, V, O). The results of the current study thus offer fresh insights on how acoustic correlates are modulated in relation to prosodic focus marking by raising two possibilities: (1) Presence of focus may be more reliably marked by weakening of focus-adjacent materials rather than enhancement or strengthening of focused materials in English, or (2) focus is marked by temporal (and intensity) enhancement of the focused materials and post-focal weakening in Canadian English, unlike MAE, in which enhancement of focused materials is observed in an across-the-board fashion. Supporting the second possibility, our results generally replicate the results of Arnhold (2021) for subject focus in Western Canadian English and extend them to verb and object focus. Importantly, while a lack of raised f0 peaks has been observed in some previous studies on MAE (Cooper et al., 1985; Eady and Cooper, 1986; Sánchez-Alvarado, 2020), it had always been restricted to sentence-initial subjects. In our study, the lack of on-focus enhancement was more pervasive both regarding the positions/constituents on which it occurred and regarding the measures for which it was observed. However, follow-up research directly comparing the two regional varieties with identical experimental materials would be needed to determine which of these interpretations is accurate.

The results of the pitch accent analysis of our data showed that whether or not a pitch accent was identified and annotated was strongly associated with the presence of narrow contrastive focus and broad focus. Narrow-focused constituents, regardless of focus location in the utterance, were almost always annotated for a pitch accent, showing annotation probability higher than 90%. The results of the first GLMM showed a three-way distinction between the three focus conditions: Narrow focus constituents were more likely to be annotated for a pitch accent than broad focus constituents, which in turn were more likely to be annotated than given constituents. This is, of course, connected to the findings from the analyses of f0 measures, which showed significant post-focal weakening in terms of f0 maximum, minimum and mean, as well as pre-focal weakening in terms of f0 maximum and mean. This observation regarding phonetic focus

marking fits with the observed phonological focus marking in terms of de-accentuation of given constituents. Interestingly, the acoustic analyses showed no enhancement of constituents in narrow focus compared to broad focus, suggesting that the fact that narrow focus constituents were annotated as accented significantly more often than broad focus ones is at least partly due to changes in duration (and relative intensity), which did show significant on-focus enhancement (cf. **Table 2** above).

The second GLMM tested the probability of H* annotation compared to any other pitch accent. It showed that H* annotation was less likely in narrow focus verbs than given verbs or objects in any focus condition. This may be interpreted as indicating that contrastiveness was more frequently marked with a pitch accent other than H* on verbs compared to objects. However, given the high probability of H* annotation in narrow focus verbs (approximately 62%), further research is needed to test whether marking contrastiveness via pitch accent manifests differently on verbs and objects.

The high probability of accent annotation for narrow focus constituents and the high probability of H* annotation support H2a (section 1.5). However, contrary to the hypothesis, L + H* was rarely found in accented constituents (see **Figure 7**), although L + H* has been argued to be associated with narrow contrastive focus in MAE (Pierrehumbert & Hirschberg, 1990). The low probability and number of L + H* in constituents with narrow contrastive focus in the current study suggest that Canadian English may differ from MAE in terms of pitch accent inventory. It is also possible that H* and L + H* are not phonologically distinct categories in (Western) Canadian English, similar to what has been argued for other varieties of English (Ladd & Schepman, 2003; Jepson et al., 2021 for British English; Arvaniti & Garding, 2007 for Minnesotan English; Arvaniti et al., 2022 for some of their British English listeners).

In addition, despite the predominance of the H* pitch accents across different focus conditions, other types of pitch accents were also found among different focus conditions, including given constituents. The relation between pitch accent type and information structure was not clear-cut, and rather, it was the location of the constituents that seemed to be more closely related to pitch accent types. That is, while given constituents were significantly less often accented than broad focus constituents according to the results of the first GLMM, the L* pitch accent does not seem to occur more frequently for given constituents than broad or narrow focus constituents, contrary to the hypothesis H2b in Section 1.5. **Figure 9** shows that L* (blue bars) occurred more frequently in verbs with broad/narrow focus than given verbs, whereas L* occurrence did not differ across the focus conditions in subjects and objects. Similarly, L + H* accents (red bars) appeared mostly on subjects independent of focus condition. These observations are somewhat in line with previous studies in American English that suggested a lack of one-to-one mapping between information status and types of pitch accents (Ito et al., 2004; Cole et al., 2019; Chodroff & Cole, 2019), but unlike those studies, our data do not suggest a probabilistic relationship between pitch accent

type and information structure. This lack of distinction between different pitch accent types, at least as a means of encoding focus/givenness, fits the results of Arnhold et al. (2020) in which Canadian listeners associated givenness with a lack of pitch accent but not with the presence of a rising accent, a salient cue to givenness for British listeners. While the presence vs. absence of pitch accent seems strongly associated with focus and givenness in Canadian English, no clear distinction between different pitch accent types has emerged in this respect so far.

It is widely agreed that there is no universal strategy regarding prosodic focus marking across dialects of English and across languages (Cruttenden, 2006; Ladd, 2008). For example, post-focal compression (i.e., reduction of post-focal material in terms of duration, f0, and/or intensity) has been observed in many different languages including Dutch, German, French, Arabic, but not in other languages such as Cantonese and Taiwanese (Xu, 2011). Relatedly, deaccenting of given constituents is cross-linguistically common, but not universal (Cruttenden, 2006). In perception, Dutch listeners associate given information with deaccented materials like listeners of MAE, Italian listeners only do so in a limited context (Swerts et al., 2002). Like MAE, focus is prosodically marked by nuclear pitch accents in German, but givenness is marked in a more gradient manner in German (Röhr & Baumann, 2010).

Differences have also been found among varieties of English. For instance, speakers of Singapore English did not distinguish words with contrastive focus and words with broad focus in terms of f0 maximum and maximum intensity (Lim, 2004). Examining four different varieties of South African English, Zerbian (2013) found that duration was an acoustic correlate of prosodic focus marking in General South African English and the prestige form (acrolect) of Black South African English, but not in the non-prestige form (mesolect) of Black South African English and the younger generation (postacrolect) Black South African English. In contrast, f0 peak on a focused word was observed in acrolect and mesolect Black South African English but not in General South African English and postacrolect Black South African English. The current study adds to the body of literature on cross-dialectal and cross-linguistic variations in how focus is prosodically marked, by putting forth the possibility that speakers of Western Canadian English may differ from speakers of MAE in employing acoustic correlates of focus and pitch accents. Our results follow the footsteps of previous studies that showed how prosodic focus marking in a variety of English does not resemble MAE or Standard British English.

5. Conclusion

The current study showed that prosodic marking of narrow contrastive focus in Western Canadian English differs from the current understanding of focus marking in MAE in phonological as well as phonetic aspects. Two tentative conclusions can be drawn: First, speakers of Western Canadian English seem to primarily rely on enhancing the focused item in terms of duration to mark focus, unlike MAE speakers who have been argued to modulate f0 in addition to duration and intensity

on the focused material to signal focus. In addition, Western Canadian English speakers appear to consistently mark givenness via weakening or reduction of the acoustic correlates adjacent to the focused item. Second, speakers of Western Canadian English showed clear association between presence of pitch accent and narrow contrastive focus, suggesting a close relation between information structure and presence/absence of pitch accent. Moreover, the pitch accent inventory of Canadian English may consist of a different set of pitch accents than that of MAE or the contrast between H* and L+H* is absent in Canadian English. As our participants all hailed from Western Canada, further study on whether the regional variety of Canadian English leads to systematic variations in focus production in terms of acoustic correlates or pitch accent inventories and alignment is necessary before reaching a clearer conclusion about the production of focus in Canadian English as a whole.

Additional file

The additional file for this article can be found as follows:

- **Supplementary Materials.** The output of the statistical models. DOI: https://doi.org/10.169 95/labphon.9316.s1

Acknowledgements

Funding information

Competing interests

The authors have no competing interests to declare.

References

Arnhold, A. (2021). Prosodic focus marking in clefts and syntactically unmarked equivalents: Prosody–syntax trade-off or additive effects? *The Journal of the Acoustical Society of America, 149*(3), 1390–1399. DOI: https://doi.org/10.1121/10.0003594

Arnhold, A., Porretta, V., Chen, A., Verstegen, S. A. J. M., Mok, I., & Järvikivi, J. (2020). (Mis) understanding your native language: Regional accent impedes processing of information status. *Psychonomic Bulletin & Review, 27*(4), 801–808. DOI: https://doi.org/10.3758/s13423-020-01731-w

Arvaniti, A., & Garding, G. (2007). Dialectal variation in the rising accents of American English. In J. Cole & J. H. Hualde (Eds.), *Papers in Laboratory Phonology 9: Change in Phonology* (pp. 547–576). Mouton de Gruyter.

Arvaniti, A., Gryllia, S., Zhang, C., & Marcoux, K. (2022). Disentangling emphasis from pragmatic contrastivity in the English H* ~ L+H* contrast. *Proc. Speech Prosody 2022*, 837–841. DOI: https://doi.org/10.21437/SpeechProsody.2022-170

Atkinson, J. E. (1973). *Aspects of intonation in speech: Implications from an experimental study of fundamental frequency*. [Doctoral Dissertation, University of Connecticut]. University of Connecticut.

Bartels, C., & Kingston, J. (1994). Salient pitch cues in the perception of contrastive focus. *The Journal of the Acoustical Society of America, 95*(5), 2973–2973. DOI: https://doi.org/10.1121/1.408967

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. DOI: https://doi.org/10.18637/jss.v067.i01

Baumann, S., Grice, M., & Steindamm, S. (2006). Prosodic marking of focus domains—categorical or gradient? *Proceedings of Speech Prosody 2006*, 301–304. DOI: https://doi.org/10.21437/SpeechProsody.2006-73

Beckman, M. E. (1986). *Stress and non-stress accent*. De Gruyter. DOI: https://doi.org/10.1515/9783110874020

Beckman, M. E., & Ayers, G. M. (1997). Guidelines for ToBI labelling. The OSU Research Foundation. https://www.ling.ohio-state.edu/research/phonetics/E_ToBI/

Beckman, M. E., & Pierrehumbert, J. (1986). Intonational structure in Japanese and English. *Phonology*, *3*, 255–309. DOI: https://doi.org/10.1017/S095267570000066X

Bloomfield, M. W. (2013). *Essays and explorations: Studies in ideas, language, and literature*. Harvard University Press. DOI: https://doi.org/10.4159/harvard.9780674733046

Boberg, C. (2008). Regional Phonetic Differentiation in Standard Canadian English. *Journal of English Linguistics*, *36*(2), 129–154. DOI: https://doi.org/10.1177/0075424208316648

Boersma, P., & Weenink, D. (2022). Praat: Doing phonetics by computer (6.2.14). http://www.praat.org/

Breen, M., Fedorenko, E., Wagner, M., & Gibson, E. (2010). Acoustic correlates of information structure. *Language and Cognitive Processes*, *25*(7–9), 1044–1098. DOI: https://doi.org/10.1080/01690965.2010.504378

Bretz, F., Hothorn, T., & Westfall, P. (2010). *Multiple comparisons using R*. Chapman and Hall/CRC. DOI: https://doi.org/10.1201/9781420010909

Büring, D. (2007). *Semantics, intonation, and information structure*. Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780199247455.013.0015

Canepari. (2010). *The pronunciation of English around the world: Geo-social applications of the natural phonetics & tonetics method*. LINCOM Europa.

Chambers, J. K. (1973). Canadian raising. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique*, *18*(2), 113–135. DOI: https://doi.org/10.1017/S0008413100007350

Chambers, J. K. (1991). Canada. In J. Cheshire (Ed.), *English around the world: Sociolinguistic perspectives* (1st ed., pp. 87–107). Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511611889.007

Chambers, J. K. (2006). Canadian raising retrospect and prospect. *The Canadian Journal of Linguistics/La Revue Canadienne de Linguistique*, *51*(2), 105–118. DOI: https://doi.org/10.1353/cjl.2008.0009

Chen, A., den Os, E., & de Ruiter, J. P. (2007). Pitch accent type matters for online processing of information status: Evidence from natural and synthetic speech. *The Linguistic Review*, *24*(2–3). DOI: https://doi.org/10.1515/TLR.2007.012

Chodroff, E. R., & Cole, J. (2019). The phonological and phonetic encoding of information status in American English nuclear accents. In *Proceedings of the 19th International Congress of Phonetic Sciences*. Article 187. https://assta.org/proceedings/ICPhS2019Microsite/

Clarke, S. (2010). *Newfoundland and Labrador English*. Edinburgh University Press. DOI: https://doi.org/10.1515/9780748631414

Clarke, S., Elms, F., & Youssef, A. (1995). The third dialect of English: Some Canadian evidence. *Language Variation and Change, 7*(2), 209–228. DOI: https://doi.org/10.1017/S0954394500000995

Clopper, C. G., & Smiljanic, R. (2011). Effects of gender and regional dialect on prosodic patterns in American English. *Journal of Phonetics, 39*(2), 237–245. DOI: https://doi.org/10.1016/j.wocn.2011.02.006

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37–46. DOI: https://doi.org/10.1177/001316446002000104

Cole, J., Hualde, J. I., Smith, C. L., Eager, C., Mahrt, T., & Napoleão de Souza, R. (2019). Sound, structure and meaning: The bases of prominence ratings in English, French and Spanish. *Journal of Phonetics, 75*, 113–147. DOI: https://doi.org/10.1016/j.wocn.2019.05.002

Cooper, W. E., Eady, S. J., & Mueller, P. R. (1985). Acoustical aspects of contrastive stress in question–answer contexts. *The Journal of the Acoustical Society of America, 77*(6), 2142–2156. DOI: https://doi.org/10.1121/1.392372

Couper-Kuhlen, E. (1984). A new look at contrastive intonation. In R. J. Watts & U. Weidmann (Eds.), *Modes of Interpretation: Essays presented to Ernst Leisi on the occasion of his 65th birthday* (pp. 137–158). Gunter Narr.

Cruttenden, A. (2006). The de-accenting of given information: A cognitive universal? In G. Bernini & M. L. Schwartz (Eds.), *Eurotyp, 8, Pragmatic Organization of Discourse in the Languages of Europe*. De Gruyter. DOI: https://doi.org/10.1515/9783110892222.311

DePape, A.-M. R., Chen, A., Hall, G. B. C., & Trainor, L. J. (2012). Use of prosody and information structure in high functioning adults with autism in relation to language ability. *Frontiers in Psychology, 3*. DOI: https://doi.org/10.3389/fpsyg.2012.00072

Eady, S. J., & Cooper, W. E. (1986). Speech intonation and focus location in matched statements and questions. *The Journal of the Acoustical Society of America, 80*(2), 402–415. DOI: https://doi.org/10.1121/1.394091

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. *Language and Speech, 29*(3), 233–251. DOI: https://doi.org/10.1177/002383098602900304

Féry, C., Pandey, P., & Kentner, G. (2016). The prosody of Focus and Givenness in Hindi and Indian English. *Studies in Language, 40*(2), 302–339. DOI: https://doi.org/10.1075/sl.40.2.02fer

Fletcher, J., Grabe, E., & Warren, P. (2005). Intonational variation in four dialects of English: The high rising tune. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 390–409). Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199249633.003.0014

Folkins, J. W., Miller, C. J., & Minifie, F. D. (1976). Rhythm and syllable timing in phrase level stress patterning. *Journal of Speech, Language, and Hearing Research, 18*, 739–753. DOI: https://doi.org/10.1044/jshr.1804.739

Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America, 27*, 765–768. DOI: https://doi.org/10.1121/1.1908022

Grice, M., Ritter, S., Niemann, H., & Roettger, T. B. (2017). Integrating the discreteness and continuity of intonational categories. *Journal of Phonetics*, *64*, 90–107. DOI: https://doi.org/10.1016/j.wocn.2017.03.003

Gussenhoven, C. (1983). Focus, mode and the nucleus. *Journal of Linguistics*, *19*(2), 377–417. DOI: https://doi.org/10.1017/S0022226700007799

Gut, U. (2005). Nigerian English Prosody. *English world-wide: A journal of varieties of English*, *26*(2), 153–177. DOI: https://doi.org/10.1075/eww.26.2.03gut

Hagiwara, R. E. (2006). Vowel production in Winnipeg. *The Canadian Journal of Linguistics/La Revue Canadienne de Linguistique*, *51*(2), 127–141. DOI: https://doi.org/10.1353/cjl.2008.0022

Halford, B. K. (2007). Adolescent intonation in Canada: Talk units in in-group conversations. *Anglia – Zeitschrift Für Englische Philologie*, *125*(1). DOI: https://doi.org/10.1515/ANGL.2007.4

Halliday, M. A. K. (1967). *Intonation and grammar in British English*. De Gruyter Mouton. DOI: https://doi.org/10.2307/3723540

Hedberg, N., & Sosa, J. M. (2007). The prosody of topic and focus in spontaneous English dialogue. In C. Lee, M. Gordon, & D. Büring (Eds.), *Topic and Focus* (Vol. 82, pp. 101–120). Springer Netherlands. DOI: https://doi.org/10.1007/978-1-4020-4796-1_6

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, *50*(3), 346–363. DOI: https://doi.org/10.1002/bimj.200810425

Ito, K., & Speer, S. R. (2008). Anticipatory effects of intonation: Eye movements during instructed visual search. *Journal of Memory and Language*, *58*(2), 541–573. DOI: https://doi.org/10.1016/j.jml.2007.06.013

Ito, K., Speer, S. R., & Beckman, M. E. (2004). Informational status and pitch accent distribution in spontaneous dialogues in English. *Proceedings of Speech Prosody 2004*, 4. DOI: https://doi.org/10.21437/SpeechProsody.2004-65

Jackendoff, R. (1972). *Semantics in generative grammar*. MIT Press.

Jepson, K., Zhang, C., Lohfink, G., Marcoux, K., & Arvaniti, A. (2021). H* and L+H* in English and Greek. *Proceedings of Phonetics and Phonology in Europe*.

Joos, M. (1942). A phonological dilemma in Canadian English. *Language*, *18*(2), 141. DOI: https://doi.org/10.2307/408979

Kochanski, G., Grabe, E., Coleman, J., & Rosner, B. (2005). Loudness predicts prominence: Fundamental frequency lends little. *The Journal of the Acoustical Society of America*, *118*(2), 1038–1054. DOI: https://doi.org/10.1121/1.1923349

Krifka. (2018). Basic notions of information structure. Acta Linguistica Hungarica. *Acta Linguistica Hungarica*, *55*, 243–276. DOI: https://doi.org/10.1556/ALing.55.2008.3-4.2

Kügler, F., & Calhoun, S. (2020). Prosodic encoding of information structure: A typological perspective. In C. Gussenhoven & A. Chen (Eds.), *The Oxford handbook of language prosody* (pp. 454–467). Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780198832232.013.30

Labov, W., Ash, S., & Boberg, C. (2005). *The atlas of North American English: Phonetics, phonology and sound change*. Walter de Gruyter. DOI: https://doi.org/10.1515/9783110206838

Lacey, C., Rampersaud, S., & Tennant, J. (1997). Observations sur les finales à montée élevée dans les phrases déclaratives en anglais canadien. In H. Gezundhajt & P. Martin (Eds.), *Promenades phonétiques* (pp. 131–143). Éditions Mélodie.

Ladd, D. R. (2008). *Intonational phonology* (2nd ed.). Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511808814

Ladd, D. R., & Morton, R. (1997). The perception of intonational emphasis: Continuous or categorical? *Journal of Phonetics, 25*(3), 313–342. DOI: https://doi.org/10.1006/jpho.1997.0046

Ladd, D. R., & Schepman, A. (2003). "Sagging transitions" between high pitch accents in English: Experimental evidence. *Journal of Phonetics, 31*(1), 81–112. DOI: https://doi.org/10.1016/S0095-4470(02)00073-6

Ladd, D. R., Schepman, A., White, L., Quarmby, L. M., & Stackhouse, R. (2009). Structural and dialectal effects on pitch peak alignment in two varieties of British English. *Journal of Phonetics, 37*(2), 145–161. DOI: https://doi.org/10.1016/j.wocn.2008.11.001

Lenth, R. (2018). Emmeans: Estimated marginal means, a.k.a. least-squares means. https://CRAN.R-project.org/package = emmeans

Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *The Journal of the Acoustical Society of America, 32*(4), 451–454. DOI: https://doi.org/10.1121/1.1908095

Lim, L. (2004). Everything you wanted to know about how stressed Singaporean Englishes are. *Program for Southeast Asian Studies* (pp. 429–444).

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language, 94*, 305–315. DOI: https://doi.org/10.1016/j.jml.2017.01.001

Maxwell, O., & Payne, E. (2018). Pitch accent types and tonal alignment of the accentual rise in Indian English(es). *Speech Prosody 2018*, 942–946. DOI: https://doi.org/10.21437/SpeechProsody.2018-190

O'Reilly, M., Dorn, A., & Chasaide, A. N. (2010). Focus in Donegal Irish (Gaelic) and Donegal English bilinguals. *Speech Prosody*, 4. DOI: https://doi.org/10.21437/SpeechProsody.2010-83

O'Shaughnessy, D. (1979). Linguistic features in fundamental frequency patterns. *Journal of Phonetics, 7*(2), 119–145. DOI: https://doi.org/10.1016/S0095-4470(19)31045-9

Ouafeu, Y. T. S. (2007). Intonational marking of new and given information in Cameroon English. English World-Wide. *A Journal of Varieties of English, 28*(2), 187–199. DOI: https://doi.org/10.1075/eww.28.2.05oua

Patience, M., Marasco, O., Colanton, L., Klassen, G., Radu, M., & Tararova, O. (2018). Initial pitch cues in English sentence types. *Speech Prosody 2018*, 463–467. DOI: https://doi.org/10.21437/SpeechProsody.2018-94

Pell, M. D. (2001). Influence of emotion and focus location on prosody in matched statements and questions. *The Journal of the Acoustical Society of America, 109*(4), 1668–1680. DOI: https://doi.org/10.1121/1.1352088

Piepho, H. (2004). An algorithm for a letter-based representation of all-pairwise comparisons. *Journal of Computational and Graphical Statistics*, *13*(2), 456–466. DOI: https://doi.org/10.1198/1061860043515

Pierrehumbert, J. (1980). The phonology and phonetics of English intonation [PhD thesis]. MIT.

R Core Team. (2021). R: A language and environment for statistical computing (4.1.2). R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Roettger, T. B., Mahrt, T., & Cole, J. (2019). Mapping prosody onto meaning – the case of information structure in American English. *Language, Cognition and Neuroscience*, *34*(7), 841–860. DOI: https://doi.org/10.1080/23273798.2019.1587482

Sánchez-Alvarado, C. (2020). Syntactic and prosodic marking of subject focus in American English and Peninsular Spanish. In A. Morales-Front, M. J. Ferreira, R. P. Leow & C. Sanz (Eds.), *Issues in Hispanic and Lusophone Linguistics* (Vol. 26, pp. 184–203). John Benjamins Publishing Company. DOI: https://doi.org/10.1075/ihll.26.09san

Sando, Y. T. (2009). Upspeak across Canadian English accents: Acoustic and sociophonetic evidence. *Proceedings of the 2009 Annual Conference of the Canadian Linguistic Association* (pp. 1–12).

Selkirk, E. (1995). Sentence prosody: intonation, stress, and phrasing. In *The handbook of phonological theory* (Vol. 1, pp. 550–569). Blackwell. DOI: https://doi.org/10.1111/b.9780631201267.1996.00018.x

Shokeir, V. (n.d.). Evidence for the stable use of uptalk in South Ontario English. *University of Pennsylvania Working Papers in Linguistics*, *14*(2), Article 4.

Steedman, M., & Lee, C. (2008). Information-structural semantics for english intonation. In M. Gordon & Büring (Eds.), *Topic and focus: Cross-linguistic perspectives on meaning and intonation* (pp. 245–264). Springer. DOI: https://doi.org/10.1007/978-1-4020-4796-1_13

Sundara, M. (2005). Acoustic-phonetics of coronal stops: A cross-language study of Canadian English and Canadian French. *The Journal of the Acoustical Society of America*, *118*(2), 1026–1037. DOI: https://doi.org/10.1121/1.1953270

Swerts, M., Krahmer, E., & Avesani, C. (2002). Prosodic marking of information status in Dutch and Italian: A comparative analysis. *Journal of Phonetics*, *30*(4), 629–654. DOI: https://doi.org/10.1006/jpho.2002.0178

Terken, J., & Hirschberg, J. (1994). Deaccentuation of words representing 'given' information: Effects of persistence of grammatical function and surface position. *Language and Speech*, *37*(2), 125–145. DOI: https://doi.org/10.1177/002383099403700202

Truckenbrodt, H. (1995). *Phonological phrases—Their relation to syntax, focus, and prominance.* [Doctoral Dissertation, Massachusetts Institute of Technology].

Turk, A. E., & Sawusch, J. R. (1996). The processing of duration and intensity cues to prominence. *The Journal of the Acoustical Society of America*, *99*(6), 3782–3790. DOI: https://doi.org/10.1121/1.414995

Turk, A. E., & White, L. (1999). Structural influences on accentual lengthening in English. *Journal of Phonetics*, *27*(2), 171–206. DOI: https://doi.org/10.1006/jpho.1999.0093

Veilleux, N., Shattuck-Hufnagel, S., & Brugos, A. (2006). Transcribing prosodic structure of spoken utterances with ToBI. *MIT OpenCourseWare*. https://ocw.mit.edu

Wagner, M., & McAuliffe, M. (2019). The effect of focus prominence on phrasing. *Journal of Phonetics*, *77*, 100930. DOI: https://doi.org/10.1016/j.wocn.2019.100930

Watson, D. G., Tanenhaus, M. K., & Gunlogson, C. A. (2008). Interpreting pitch accents in online comprehension: H* vs. L+H*. *Cognitive Science*, *32*(7), 1232–1244. DOI: https://doi.org/10.1080/03640210802138755

Weismer, G., & Ingrisano, D. (1979). Phrase-level timing patterns in English: Effects of emphatic stress location and speaking rate. *Journal of Speech, Language, and Hearing Research*, *22*(3), 516–533. DOI: https://doi.org/10.1044/jshr.2203.516

Welby, P. (2003). Effects of pitch accent position, type, and status on focus projection. *Language and Speech*, *46*(1), 53–81. DOI: https://doi.org/10.1177/00238309030460010401

Xu, Y. (2011). Post-focus compression: Cross-linguistic distribution and historic origin. In W. S. Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII)* (pp. 152–155).

Xu, Y., & Xu, C. X. (2005). Phonetic realization of focus in English declarative intonation. *Journal of Phonetics*, *33*(2), 159–197. DOI: https://doi.org/10.1016/j.wocn.2004.11.001

Zerbian, S. (2013). Prosodic marking of narrow focus across varieties of South African English. English World-Wide. *A Journal of Varieties of English*, *34*(1), 26–47. DOI: https://doi.org/10.1075/eww.34.1.02zer