Focus on Canadian English: Acoustic Realization and Listener Perception

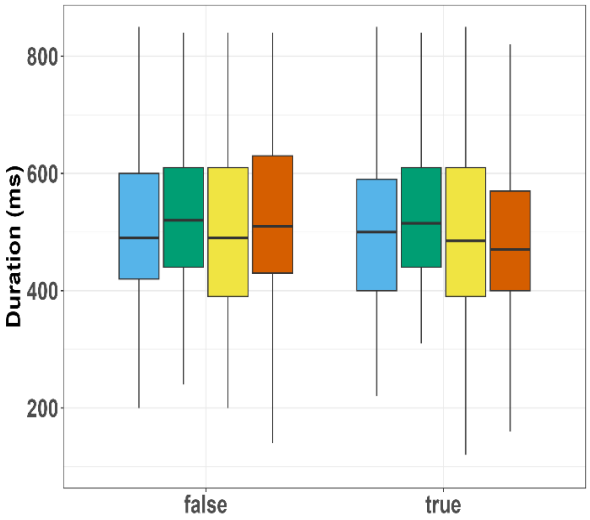
Yuting Sun, Ziyan Zou, Anja Arnhold
Department of Linguistics, University of Alberta

Introduction Focus tends to affect the acoustic realization of words. From the production perspective, on-focus is expressed with higher fundamental frequency (*f0*), longer duration and higher amplitude and post-focus is shown with compressed *f0* contours in a stress language like English [1][2][3]. From the perception view, listeners rely on those acoustic cues to identify focus, but cue weighting differs across languages (British and Irish English:[4]). For Canadian English, only few studies investigate the perception and production of focus realization [2][5], and [2] reported that speakers relied more on duration than f0 and intensity to mark focus. The present study investigates (1) to what extent focus influences the acoustic realization of Canadian English, specifically in terms of duration, intensity, and fundamental frequency (*f0*) (2) whether there was a relationship between acoustic realizations and listeners' successful identification of focus?

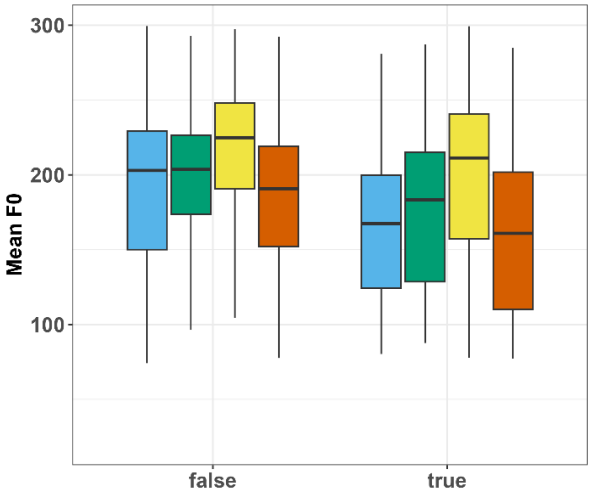
Methods Nineteen participants were recruited at the University of Alberta. None reported any hearing or visual impairments, and all were native speakers of Canadian English. This study was approved by the ethics committee (Pro00134287), and participants received course credits for their time. 160 items were created and appeared in four types of focus questions (broad, subject, object, and location), equally divided into 4 lists. This experiment was conducted interactively, involving both participants and the researcher. In each trial, participants saw a question and four words on a screen. They were required to use all four words to build a sentence to answer the given question (e.g., Object focus: What did the mayor repeat at the meeting? Answer: The mayor[*prefocal*] repeated[*prefocal*] the motto[*narrow*] at the meeting[*postfocal*]). The researcher would guess which focus condition the participant intended (resulting in Accuracy: *True* or *False*). If the researcher's guess was incorrect (*false* condition), the participant was allowed a second attempt; after two attempts, the task automatically proceeded to the next item. Separate linear mixed models were fit for each acoustic measure, using buildmer [6] to find the best fixed and random effects structure. Best models for all measures included an interaction between focus condition and accuracy as fixed effects and random intercepts for participants. To avoid additional influence of syntactic variation, this study only analyzed syntactically unmarked sentences and excluded other types of answers (e.g., it-cleft, pseudo-cleft), leaving 751 utterances, i.e. 1709 words, for analysis (only selecting subjects, object, verb and location).

Results Overall, all three cues showed similar patterns regarding the differences in focus, although no significant differences occurred for duration. For mean intensity and mean *f0*, significantly lower values appeared in postfocal condition compared to both narrow and broad focus (see Figure 1 and Table 1). Broad focus and narrow focus did not differ significantly from each other, but prefocal words showed significantly higher values than both. Interestingly, although the interaction between focus condition and accuracy significantly improved all models, no significant differences appeared between the two accuracy conditions.

Discussion and Conclusions In summary, the results show that native Canadian English speakers exhibited systematic acoustic differences across the four focus conditions. Regardless of whether listeners identified the focus correctly or incorrectly, post-focal values were consistently smaller than those in narrow focus. This supports the post-focal *f0* compression reported in [2] and extends the effect to intensity and to Canadian English. Although previous studies have suggested that duration is the crucial cue for American English listeners, the present did not find significant differences for duration, indicating a weaker role for this cue in our dataset. Future work with more participants will be needed to clarify the contribution of duration and to examine how prosodic and syntactic factors interact.



(a) Duration



(b) Mean F0

(c) Mean Intensity

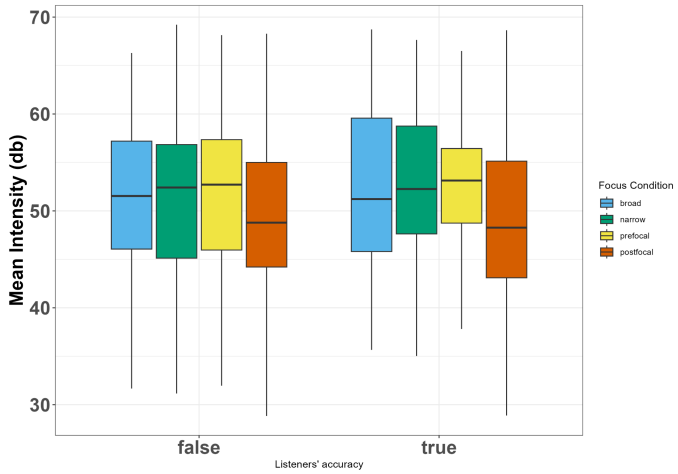


Figure 1: Acoustic measurements [duration-a, F0-b, intensity-c] by listener accuracy (true vs false) under four focus conditions (broad-blue, narrow-green, prefocal-yellow, postfocal-red).

Table 1: Results of significant pairwise comparisons among combinations of focus conditions and listener accuracy.

Table 1

Contrast	Estimate	Acoustic Cue	P Value
broad false –prefocal false	-1.37 and -13.65	mean intensity and F0	<.05
broad false –postfocal false	2.4960 and 15.18	mean intensity and F0	<.0001
narrow false –prefocal false	-2.2971 and -14.78	mean intensity and F0	<.05
narrow false –postfocal false	2.3153 and 14.05	mean intensity and F0	0.0001
prefocal false-postfocal false	4.6124	mean intensity and F0	<.0001
broad true –prefocal true	-2.4297 and -12.87	mean intensity and F0	<.05
broad true –postfocal true	3.5270 and 11.04	mean intensity and F0	<.0001
prefocal true –postfocal true	5.9567 and 23.91	mean intensity and F0	<.0001
narrow true –postfocal true	4.7661	mean intensity	<.0001
narrow true –prefocal true	-1.1906	mean F0	<.05

References

[1] Breen, M., Fedorenko, E., Wagner, M., & Gibson, E. (2010). Acoustic correlates of information structure. *Language and Cognitive Processes*, 25(7–9), 1044–1098.

[2] Kim, J., & Arnhold, A. (2024). Phonetic and phonological aspects of prosodic focus marking in Canadian English. *Laboratory Phonology*, 15(1), Article 9316.

[3] Chen, Y. (2010). Post-focus f0 compression—Now you see it, now you don't. *Journal of Phonetics*, 38(4), 517–525.

[4] 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.

[5] 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, Article 72.

[6] Voeten CC (2025).buildmer: Stepwise Elimination and Term Reordering for Mixed-Effects Regression. R package version 2.12