An analysis of prosodic prominence cues to information structure in Egyptian Arabic

Dina El Zarka[1], Anneliese Kelterer[1] , Barbara Schuppler[2]

[1]Department of Linguistics, University of Graz
[2] Signal Processing and Speech Communication Laboratory, Graz University of Technology
dina.elzarka@uni-graz.at, anneliese.kelterer@uni-graz.at, b.schuppler@tugraz.at

Abstract

This study presents the first acoustic examination of prominence relations in entire contours associated with different information structures in Egyptian Arabic. Previous work has shown that topics and foci are typically associated with different pitch events, whereas it is still a matter of debate whether and how Egyptian Arabic uses prominence relations to mark narrow focus. The analysis of data from 17 native speakers showed that narrow focus was marked by on-focus pitch expansion as well as post-focus compression. Post-focus compression was realized as a large downstep after focus, compressed pitch range, lower intensity and shorter duration. The results also showed further register lowering after a contrastive focus, but no further pitch boost of the focused word. By contrast, a contrastive topic showed higher scaling of the topic as well as an expanded pitch range of the overall contour. The findings of this study stress the significance of whole contours to convey intonational meanings, revealing gradient prominence cues to focus across the utterance, specifically post-focus register lowering to enhance the prominence of a contrastive focus.

Index Terms: information structure, narrow focus, prosodic prominence, post-focus compression, acoustic cues

1. Introduction

Prosodic prominence is one of the most important cues to information structure (IS) cross-linguistically. It has frequently been assumed that focus as the most important or newsworthy part of the information [1] [2] includes the strongest prosodic prominence in a sentence [3] [4]. Many accounts of IS distinguish *broad* focus (BF) (i.e., a larger constituent such as a VP or a sentence) from *narrow* focus on one constituent (e.g., an argument). While in broad focus all constituents are new, in narrow focus only the focus itself provides new information (IF). In addition, a narrow focus may also be *contrastive* (CF), standing in syntagmatic contrast to another focused item. In some languages, focus strength has been found to correlate with increased prominence (narrow vs. broad, contrastive vs. non-contrastive) [5] [6] [7]. In many languages, *given* information is typically non-prominent prosodically [4] [5]. Usually, one referent within the given part denotes the aboutness *topic* of the sentence, [2] [8], which may be *continuous* or *ratified* (RT), or *contrastive* (CT). Examples for these IS categories are given in Table 1.

In flexible-accent languages (e.g., English), narrow focus in an early position is marked by retracting the nuclear accent from the end of the phrase to the focus position. Non-flexible accent languages (e.g., Romance) have been considered to mark focus syntactically, but not prosodically [4] [9] [10]. Recent prosodic studies, however, have revealed prosodic strategies for focus marking also in non-flexible accent languages. First, post-focus compression may be considered the equivalent of deaccentuation [11], even in tone languages like Mandarin [12] or phrase languages like Hindi [13]. Second, peaks are aligned earlier under focus, suggesting a different accent type [14] [15] on the focused item. Such results cast doubt on the original assumption of full functional complementarity between intonational and syntactic focus marking.

Egyptian Arabic (EA) is known for its resistance to deaccentuation, also in contrastive settings [16]. Empirical results concerning focus marking in this language are varied. In an early pilot study with only one speaker, [17] found on-focus expansion of F0 plus post-focus compression. [18] [19] examined prosodic reflexes of givenness and contrast (equivalent to CF) in six speakers. She found neither F0 alignment differences nor differences in intensity or duration, but F0 expansion of contrastive items followed by F0 compression – albeit only for three of the six speakers. Based on a corpus study and a pilot experiment, [20] argued that topics and foci are preferentially associated with rising and falling contours, respectively. To test this claim, acoustic cues to the first accent of an SVO sentence under different IS conditions were investigated by [21] and [22]. The results showed that accent peaks and following valleys were indeed aligned earlier, and that peaks were scaled higher in narrow foci (IF, CF) than in topics (RT, CT) [22] (Figure 1a). Similar, though less clear results were obtained for broad (BF) versus narrow focus (IF, CF) [21]. While contrast may raise the peak of the topic considerably (CT vs. RT), it was shown to lower it in focus condition (CF vs. IF) [22]. Finally, the stressed syllable is significantly longer in narrow foci than in topics [22].

This study presents an acoustic examination of the whole utterance under the five IS conditions across a large number of speakers. Cross-linguistically, focus has been shown to exhibit greater F0 expansion, higher F0 scaling and higher durational and intensity values. Post-focus compression may involve any or all of the inverse cues. While relative prominence differences seem to involve on-focus and post-focus acoustic cues in some languages [5] [12] [14] [23] [24], other languages have been found to show only post-focus cues [13] [25]. Some studies have reported higher F0 peak values in contrastive focus [5] [6], whereas others reported lower values [22] [24]. The results of prior studies on EA were inconclusive about the relative importance of on-focus versus post-focus cues to prominence. It is also not clear whether the post-focus domain involves a narrower pitch range or rather register lowering as suggested by the contours in Figure 1b.

In this study, we test the acoustic cues of IS-related relative prominence differences within an utterance. Impressionistically, our data[1] suggests that most speakers shift the main prominence to the first word when under narrow focus and distribute prominence evenly in broad focus. Some speakers tend to reduce the prominence of topic constituents. To shed more light on these issues, we investigate the acoustic make-up of the on-focus, post-focus, on-topic and post-topic domains in terms of scaling, pitch range, duration and intensity using relative and absolute measures.

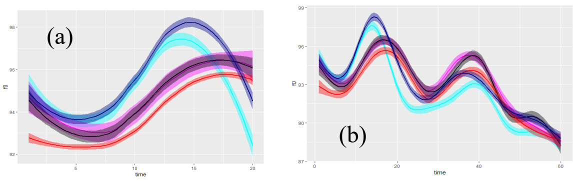

Figure 1. *F0 contours on a) the target word and b) the whole sentence (BF=black, contrastive CF=cyan, IF=dark blue) and topic (RT=red, CT=magenta) for a female speaker.*

2. Method

2.1. Speech material

We used the data of 17 native speakers of EA (11 female, 6 male) aged between 22 and 78, born and raised in Cairo or Alexandria. All but one had university-level education and knowledge of some European language. The target sentences contained three words in the default word order SVO (Table 1). They were elicited in 6 mini-dialogues in all five IS conditions with 3 or rarely 4 repetitions. Participants were presented with the target sentence on a computer screen and listened to the pre-recorded question stimuli. Sentences were recorded with a head-mounted microphone in a quiet room. Contextual information was provided to elicit more natural answers, specifically in the case of the "broad focus" question "What happened?", in order to avoid the topic construal of the subject. From a total of 1646 recorded items, 317 were excluded due to bad quality, disfluencies, a break after the target word or syntactic deviations. We only used those items for the analysis that were realized in one intonation phrase and exhibited downstep within the VP-domain (i.e. word 2 + word 3) (measured with a threshold of min. 7 Hz difference in maxF0 between the last two words, following [26]), leading to the exclusion of another 302 sentences, resulting in 1027 items from 17 speakers for the analysis.

All utterances were segmented manually into the three words: subject (W1), verb (W2) and object (W3). Additionally, the stressed syllables of each word were segmented (S1, S2, S3).

2.2. Acoustic measures and statistical method

F0 tracks were extracted automatically, manually corrected and smoothed using *mausmooth* [27] in *Praat* [28]. Then, all measures of F0 were normalized for speaker by converting Hz values to semitones with each speaker's overall mean F0 as a base for calculation (cf. [29]). Intensity values were extracted with *ProsodyPro* [30] and normalized on utterance level. We calculated three absolute F0 measures (Max_{W1}, Min_{W3}, $Range_{U}$), two relative F0 measures (Max_{W1}-Max_{S2}, Max_{S2}-Min_{W3}), three measures of mean intensity (Int_{W1}-Int_{W2}, Int_{W2}-Int_{W3}, Int_{W3}) and two durational measures (Dur_{U}, Dur_{VP}/Dur_{U}; VP = W2+W3).

To test the effect of a narrow focus on the rest of the contour, we divided the range of the utterance into two intervals covering the initial word and the VP, respectively. Interval 1 (Max_{W1}-Max_{S2}) is a measure of register lowering and interval 2 (Max_{S2}-Min_{W3}) is related to the pitch range of the VP-domain. To avoid the influence of a very late peak accent in the first word on the maxF0 of the second word, we measured the F0 difference between the maxima of W1 and S2 rather than W2.

For each acoustic measure, we built linear mixed effects regression models with the *lmer()* function of the *lme4 package* in R [31] (M_Max_{W1}, M_Min_{W3}, M_$Range_{U}$, M_Max_{W1}-Max_{S2}, M_Max_{S2}-Min_{W3}, M_Int_{W1}-Int_{W2}, M_Int_{W2}-Int_{W3}, M_Int_{W3}, M_Dur_{U}, M_Dur_{VP}/Dur_{U}). The dependent variable of each model was the respective acoustic measure. The independent variable of interest was *Condition* (CF = narrow contrastive focus; IF = narrow information focus; CT = contrastive topic; RT = ratified topic; BF = broad focus). We included the nuisance variables *Age* (group 1: 22-30 years; group 2: 31-59 years; group 3: 60-78 years), *Sex* (f, m) and *Repetition* (1-4), as well as the three nuisance variables in interaction with *Condition*. We included *Speaker* (n = 17) and *Sentence* (n = 6) as random variables. In the models of the three intensity measures, we excluded the data of four speakers due to background noise (n = 248). We fit the models by performing a stepwise backward regression with the *step()* function in R and subsequent manual reduction. After each step, we compared models in an ANOVA, comparing AIC, BIC and degrees of freedom [32]. The next section presents estimated marginal means (EMM), calculated with the *emmeans package* in R [33], and pairwise comparisons of the five conditions with Bonferroni correction averaged over the levels of the nuisance variables (if present in the final models). The data, the standard outputs of the final models and emmeans results can be found on www.dinaelzarka.com (Publications).

Table 1: *Target sentence in five IS conditions.*

hali:ma najjimit ama:ni. 'Halima put Amani to bed.'	
BF	**How come**, it's so quiet in here today?
IF	Amani's fast asleep. **Who** put Amani to bed?
CF	**Who** put Amani to bed, Halima or Nabila?
RT	**What** did Halima do?
CT	**Nabila** set the table, but **what** did **Halima** do?

3. Results

3.1. Speaker-specific strategies

Like prior studies [19][21], we found speaker-specific strategies of IS realization. Figure 2 shows different pitch strategies, exemplified by two F0 measures in three speakers. While speaker M03's topic sentences (RT and CT) have a narrow overall pitch range (height of column), those of F10 are realized in a wider pitch range. While M03 and F10 strongly compress the pitch range of the VP after focus (Max_{S2}-Min_{W3}; orange), speaker F03 shows an expanded pitch range in RT sentences and hardly any difference among the other conditions. For the first downstep interval (Max_{W1}-Max_{S2}; blue), the utterances of M03 and F10 show a considerably larger step

[1] Parts of this data set were used in [21] and [22].

down from the peak to the level of the following stressed syllable in narrow focus than in the other conditions. Generally speaking, a large downstep and a compressed pitch range after focus are more frequent as is shown by the statistical analysis of the pooled data in 3.2.

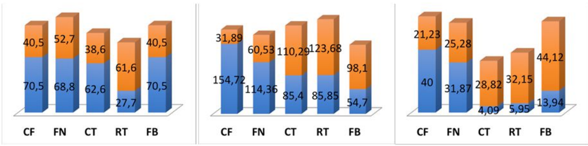

Figure 2. *Total pitch range (height of column), interval 1 (bottom, blue) and interval 2 (top, orange) for speakers F03 (left) and F10 (middle) and M03 (right). IS conditions are CF, IF, CT, RT and BF, from left to right.*

3.2. Statistical analysis across speakers

3.2.1. Whole contour

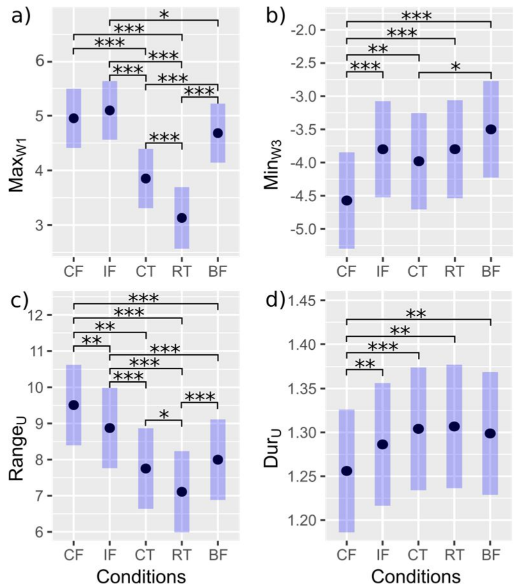

Figure 3. *EMMs (CI in blue) of (a) maxF0 in W1(M_Max_{W1}), (b) minF0 in W3 (M_Min_{W3}), (c) total pitch range (M_$Range_U$) and (d) total duration of the utterance (M_Dur_U).*

Figure 3 shows the results for the measures of the whole contour. CF, IF and BF showed the highest peaks in the first word (*Max_{W1}*) with slightly lower values for BF than IF, while values in the RT condition were lower than in all other conditions. *Min_{W3}* values were significantly lower for CF than all other conditions and significantly higher for BF than for CT. *$Range_U$* was significantly higher for CF and IF than for all other conditions and significantly lower for RT than for CT and BF. *Dur_U* was significantly lower for CF than all other conditions.

These results show that while BF patterns with the topic sentences regarding F0 range, it patterns with narrow foci regarding peak scaling. Lower peak values in CF than in IF (although not significant) replicate the results reported in [22]. The lower F0 values at phrase end (*Min_{W3}*) suggest stronger register lowering after a contrastive focus compared to all other conditions, even to IF (cf. also 3.2.3). Initial peaks of BF sentences are nearly as high as in narrow foci, but the slope of the contour is shallower and ends higher. Furthermore, there is a considerable peak lowering in RT, while RT, BF and even IF utterances end at similar pitch levels. Thus, the narrow pitch range in RT seems to be predominantly due to the lower scaling of the initial peak, while the wider pitch range of CT sentences results from both a higher peak and a lower value at the end of the utterance. At the same time, contrast raises the register of the topic word (CT) and slightly expands the pitch range of the whole utterance compared to an RT-sentence.

3.2.2. Initial downstep

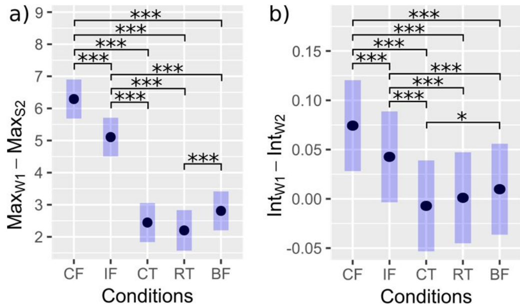

Figure 4. *EMMs (CI in blue) of (a) initial F0 downstep (M_Max_{W1}-Max_{S2}) and (b) intensity drop between W1 and W2 (M_Int_{W1}-Int_{W2}).*

Figure 4 shows the results for the first interval in terms of F0 and intensity drop. Both models clearly differentiate between two groups: the narrow focus group (CF and IF) and all other IS conditions. *Max_{W1}-Max_{S2}* was significantly higher in CF and IF than for all other conditions with the value in CF being significantly higher than in IF. The same applies to BF compared to RT. *Int_{W1}-Int_{W2}* was significantly higher for CF and IF than for all other conditions, with higher values for CF. BF had higher *IntW1-IntW2* values than CT.

These results suggest that F0 downstep and intensity drop are specifically strong after a contrastive focus. Although the F0 and intensity models show the same overall patterning, they are not fully parallel concerning the differences between CT, RT and FB. While F0 downstep is especially small after a ratified topic (RT), it is larger in the BF condition (cf. also the higher scaling of the peak in W1, 3.2.1). Some speakers realize ratified topics in a compressed pitch range and some utterances even show an upstep of the accent in *W2*. Thus, pooled across speakers, the first downstep was very small in the RT condition. The significant intensity difference between CT and FB may result from the strategy of some speakers to place the strongest accent of the phrase on the first word of the VP in the CT condition. Together with the F0 and intensity results reported for the second interval (3.2.3, Figure 5), this shows that intensity and F0 are not completely correlated, but that they may be varied independently to a certain degree.

3.2.3. Post-focus compression

Figure 5 shows the results for the second interval. *Max_{S2}-Min_{W3}* was lower for CF and IF than for all other conditions, with the lowest values in CF. Lower *Max_{S2}-Min_{W3}* values in RT compared to CT and BF were only marginally significant. CF had significantly higher *Int_{W2}-Int_{W3}* values than all other conditions, except for IF. *Int_{W2}-Int_{W3}* was also significantly higher in IF than RT and BF. CF and IF had lower *Dur_{VP}/Dur_U* values than all other conditions. CT had higher *Dur_{VP}/Dur_U* values than RT and BF, and values in RT were significantly higher than in BF. CF and IF had significantly lower I_{W3} values than all other conditions, with the lowest values in CF and a significant difference between CF and IF.

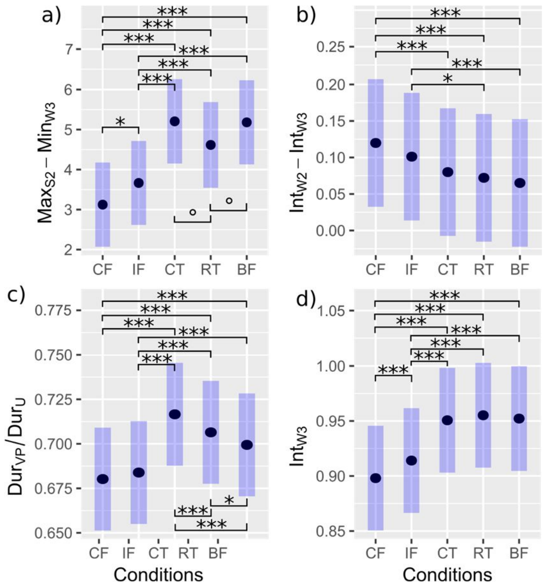

Figure 5: *EMMs (CI in blue) of (a) the second F0 interval (M_Max$_{S2}$-Min$_{W3}$), (b) mean intensity difference between W2 and W3 (M_Int$_{W2}$-Int$_{W3}$), (c) relative duration of the VP (M_Dur$_{VP}$/Dur$_{U}$) and (d) mean intensity of W3 (M_Int$_{W3}$).*

The results for the second F0 interval (Figure 5a) suggest a distinction between the narrow focus group and the other IS conditions in terms of post-focal compression, being a mirror image of the initial downstep model (Figure 4a) with two notable differences. First, the distinction between the narrow focus group and all other IS-conditions is less pronounced, and second, interval 2 in RT sentences does not show the inverse relation to interval 1. That is, while RT shows the lowest downstep from the W1 to S2, it does not show the widest pitch range of the VP. Rather, the pitch range is narrower in RT than in CT and BF sentences, even though this difference is only marginally significant. This result is in accordance with the narrower overall pitch range for RT utterances. It furthermore shows that the narrower pitch range is not only due to the narrow range of the ratified topic itself, but also of the VP-domain. The F0 model also shows slightly stronger compression after CF than after IF. Likewise, there were larger intensity differences within the VP-domain between the narrow focus sentences and most other conditions. There is, however, no significant difference between CF and IF on the one hand and IF and CT on the other hand. A clearer picture is drawn by the absolute mean intensity values in the last word (Figure 5d). In sum, the larger drop in intensity within the VP domain and specifically the lower values in W3 clearly suggest that the intensity in the final word is strongly diminished after a narrow focus, specifically after a contrastive narrow focus.

The results of the duration measure (Figure 5c) differentiate between the narrow focus group and the other IS conditions, which also exhibit significant differences between each other. The longer duration of the VP in the CT condition matches its wider pitch range (Figure 5a) and lower F0 at phrase end (Figure 3b). These results suggest that speakers expend more effort on the comment part (encoded by the VP), which is also contrastive in the context of the question used to elicit the target sentence (see Table1). When two topics are in contrast, the comments about these topics are also contrastive. Finally, the high compression values after a contrastive focus together with the very low values at phrase end suggest that the whole VP is realized in a lower pitch register in the CF condition compared to the other IS conditions, which we interpret as a stronger register lowering.

4. General discussion

This study investigated prosodic prominence cues to information structure in entire utterances. Our analysis of the pooled data showed a clear marking of narrow focus by relative prominence. This was realized by all three investigated acoustic cues, both in the focused word and in the post-focus domain as shown for many languages [11]: higher on-focus scaling and post-focus compression of pitch range as well as post-focus register lowering, lower intensity values and shorter duration of the post-focus domain and a very weak articulation at phrase end.

The CF results suggest that F0 lowering and lower intensity at the end of the phrase are important cues to relative prominence: while contrastive foci themselves were scaled even lower than non-contrastive foci, their higher relative prominence seems to be achieved by post-focus compression and additional register lowering after focus.

Another interesting result is the high scaling of the first peak in BF sentences, which groups BF with the narrow foci in this respect. This can be interpreted as a reflex of focus or novelty. In contrast, the lower scaling of ratified topics suggests that givenness induces register lowering. Both results are thus in line with the claim that focus raises the register and givenness lowers it [34]. Furthermore, a contrastive topic exhibits a higher pitch peak and an expanded register. However, contrast not only enhances the prominence of the contrastive topic itself, but also of the following comment by expanding its pitch range. All these results match the results obtained in a corpus study of spontaneous speech in EA [35].

Although the present study showed similar results for F0, intensity and duration, there were also some differences between them. For narrow focus sentences, we found specifically low intensity values at the end of utterances, while F0 downstep between W1 and W2 was a clearer cue to narrow focus than a local drop in intensity. In sum, the findings show prosodic reflexes of information structure across the entire contour as well as some independence of the individual cues.

Finally, our data also showed different speaker-specific strategies as reported by [19] and [21], which were not investigated here. In future work, we intend to examine the different speaker-specific strategies relating to whole tunes as well as their individual components to reveal the phonetic and phonological means to mark IS structure in EA. Specifically, we will relate the local acoustic cues of the first accent to the global cues of the whole contour investigated in this study.

5. Acknowledgements

The work by Anneliese Kelterer was funded by grant P-32700-N and Barbara Schuppler by grant V-638-N33, both from the Austrian Science Fund. We would like to thank Katerina Petrevska and Benjamin Daniczek for creating the segmentations and all our speakers for their contribution.

References

[1] M.A.K. Halliday, “Notes on transitivity and theme in English, Parts 2,” *Journal of Linguistics* 3(2), 199–244, 1967.

[2] K. Lambrecht, 1994. *Information structure and sentence form: Topic, focus and the mental representations of discourse referents*. Cambridge Univ. Press, 1994.

[3] R. Jackendoff, *Semantic interpretation in generative grammar.* MIT Press, 1972.

[4] D.R. Ladd, *Intonational phonology*. Cambridge Univ. Press, 2008.

[5] S. Baumann, J. Becker, M. Grice, and D. Mücke. 2007. “Tonal and articulatory marking of focus in German,” in *Proc. of the 16th International Conference of the Phonetic Sciences* 2007, 1029–1032.

[6] J. Hanssen, J. Peters, and C. Gussenhoven. “Prosodic effects of focus in Dutch declaratives,” in *Proc. of Speech Prosody 2008*, 609–612.

[7] M. Grice, S. Ritter, H. Niemann, and T. Roettger, “Integrating the discreteness and continuity of intonational categories,” *Journal of Phonetics* 64, 90–107, 2017.

[8] J. K. Gundel, *The role of topic and comment in linguistic theory*. Bloomington: Indiana Univ. Linguistics Club, 1976.

[9] D. Büring, *Intonation and Meaning*. Oxford Univ. Press, 2016.

[10] E. Vallduví, *The informational component*. New York: Garland. 1992.

[11] Y. Xu, “Post-focus compression: Cross-linguistic distribution and historical origin,” in: *Proc. of the 17th International Congress of Phonetic Sciences, 2011.*

[12] Y. Xu, “Effects of tone and focus on the formation and alignment of F0 contours,” *J. Phon.* 27(1), 55–105, 1999.

[13] U. Patil, G. Kentner, A. Gollrad, F. Kügler, C. Féry, and S. Vasishth, “Focus, word order and intonation in Hindi,” Journal of South Asian Linguistics, 1(1), 53–67, 2008.

[14] M.d.M. Vanrell, A. Stella, B. Gili-Fivela, and P. Prieto, “Prosodic manifestations of the Effort Code in Catalan, Italian and Spanish contrastive focus”, *Journal of the International Phonetic Association* 43(2), 195–220, 2013.

[15] A. Manolescu, D.J. Olson, and M. Ortega-Llebaria, “Cues to contrastive focus in Romanian,” in M. Vigáro, S. Frota, and M. J. Freitas (Eds.), *Phonetics and phonology: Interactions and interrelations* (pp. 71–90). Philadelphia: Benjamins, 2009.

[16] S. Hellmuth, “No de-accenting in (or of) phrases. Evidence from Arabic for cross-linguistic and cross-dialectal prosodic variation,” in S. Frota, M. Vigáro, and M.J. Freitas (eds), *Prosodies* (pp. 99–121). Berlin: de Gruyter, 2005.

[17] K. Norlin, “A preliminary description of Cairo Arabic intonation of statements and questions”, in *Quarterly Progress and Status Report (QPSR)* 1989, 47–49.

[18] S. Hellmuth, “The (absence of) prosodic reflexes of the given/new distinction in Egyptian Arabic”, in J. Owens, and A. Elgibali (eds), *Information Structure in Spoken Arabic* (pp.165–188). London: Routledge, 2010.

[19] S. Hellmuth, “Acoustic Cues to Focus and Givenness in Egyptian Arabic,” in Z.M. Hassan, and B. Heselwood, (eds): *Instrumental studies in Arabic phonetics,* 301–323. Amsterdam: Benjamins, 2011.

[20] D. El Zarka, *On the interaction of information structure and prosody. The case of Egyptian Arabic*. Habilitation thesis. University of Graz: Graz, 2013.

[21] F. Cangemi, D. El Zarka, S. Wehrle, S. Baumann, and M. Grice, “Speaker-specific intonational marking of narrow focus in Egyptian Arabic,” *Proc. of Speech Prosody 2016,* 335–339.

[22] D. El Zarka, B. Schuppler, and F. Cangemi, “Acoustic cues to topic and narrow focus in Egyptian Arabic,” in *INTERSPEECH 2019*, 1071–1075.

[23] M.S. Alzaidi, Y. Xu, and A. Xu, “Prosodic encoding of focus in Hijazi Arabic,” *Speech Communication* 106, 127–149, 2019.

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

[25] C. Ipek, Phonetic realization of focus with no on–focus pitch range expansion in Turkish. *Proc. of the 17th International Congress of Phonetic Sciences 2011*.

[26] F. Kügler, and C. Féry, “Postfocal downstep in German,” *Language and Speech* 60(2), 260–288, 2017.

[27] F. Cangemi, *mausmooth* [Online], 2015. Available: http://phonetik.phil-fak.uni-koeln.de/fcangemi.html.

[28] P. Boersma and D. Weenink, Praat: doing phonetics by computer (version 6.0.37) [Online], 2018. Available: http://www.praat.org/.

[29] T. Rathcke, and J. Harrington, “The phonetics and phonology of high and low tones in two falling f0-contours in Standard German”, in *INTERSPEECH 2007,* 982–985.

[30] Y. Xu, *ProsodyPro* (version 5.6.3) [Online], 2016. Available: http://www.homepages.ucl.ac.uk/~uclyyix/ProsodyPro/.

[31] D. Bates, M. Mächler, B.Bolker, and S. Walker, „Fitting Linear Mixed-Effects Models Using lme4,” *J. Statistical Software* 67(1), 1-48, 2015.

[32] R. Baayen, *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press, 2008.

[33] R. Length, *emmeans: Estimated Marginal Means, aka Least-Squares Means* (version 1.3.2) [Online], 2019. Available: https://cran.r-project.org/web/packages/emmeans/index.html.

[34] C. Féry, and S. Ishihara, “How focus and givenness shape prosody,” in M. Zimmermann, and C. Féry (eds.), *Information structure: Theoretical, typological, and experimental perspectives* (pp. 36–65). Oxford Univ. Press, 2010.

[35] D. El Zarka, B. Schuppler, “On the interplay of pragmatic and formal factors in the prosodic realization of themes in Egyptian Arabic” *Grazer Linguistische Studien (Graz Linguistic Studies)* 90(2), 33–106, 2018. (DOI:10.25364/04.45:2018.90.2)