

Phonetic Cues to Prominence in Lebanese Arabic

Dana Chahal[†]

[†] University of Balamand, Lebanon

E-mail: d.chahal@balamand.edu.lb

ABSTRACT

This paper investigates the acoustic correlates of three phonological levels of prominence in Lebanese Arabic—lexical stress, pitch accent, and nuclear accent. The paper also examines intrinsic vowel quality effects, and investigates the realizational differences between broad and narrow focus. It was found that the higher the prominence level of a particular vowel, the higher its pitch and amplitude, the longer its duration, and the more peripheral its spectral realization within the vowel space. In addition, high vowels display intrinsically higher pitch, shorter duration and lower amplitude than low vowels. It was also found that the F0 relationship between target peaks plays a significant role in signaling broad versus narrow focus. In broad focus, nuclear accented, accented, and stressed vowels may receive similar pitch values, but are significantly differentiated in terms of non-tonal cues. Under narrow focus, however, pitch and amplitude differences between the narrow focused nuclear accents and the other target syllables within the utterance are maximized, while the other non-tonal cues are not as differentiated.

1. INTRODUCTION

One of the two main functions of intonation in languages such as Arabic and English is the marking of some syllables within the phrase as more prominent than others. It has been established, in both English and various varieties of Arabic, that pitch, intensity, duration and vowel formant patterns all form acoustic correlates of prominence ([3]). Thus, prominent syllables are typically higher-pitched, longer, louder and have more distinctive vowel quality features than non-prominent syllables.

While these four phonetic features are widely accepted as cueing prominence in English and Arabic, in an accurate investigation of the phonetics of prominence in the language, it is necessary to delimit which phonological prominence level is being discussed. This is because, in both languages, F0 is a *post-lexical* tonal mark, and therefore constitutes a correlate of accented and nuclear accented syllables only. The cues for lexical stress, on the other hand, are inherently non-tonal, involving duration, intensity and vowel formant characteristics ([1]). Because accented and nuclear accented syllables are phonotactically constrained to associate with lexically stressed syllables, these post-lexical prominence levels carry both tonal and non-tonal cues.

In the model of Lebanese Arabic intonation presented in [2], Arabic distinguishes three paradigmatic phonological

prominence levels— the lexical stress, pitch accent and nuclear accent levels. These prominence levels are approximately equivalent to those proposed for English in various studies. This paper focuses on an experiment designed to investigate the F0, duration, RMS and formant frequency cues to the paradigmatic prominence levels proposed for Lebanese Arabic in [2].

Besides prominence level, the nature of the target vowel, i.e. its intrinsic vowel quality, also affects the various acoustic cues mentioned above. For example, it is found that high vowels in (American) English display shorter duration, higher F0 values, and less intensity than low vowels occurring in identical contexts [7]. Similar intrinsic vowel quality effects have been found for different varieties of Arabic ([3]). Accordingly, the second aim of the experiment is to investigate these intrinsic vowel quality effects in Lebanese Arabic.

A third aim for the paper is to investigate the surface structure realization of broad and narrow focus in Lebanese Arabic. Here, two main findings in the literature are relevant. First, various experiments show that the presence of prenuclear accents is instrumental in the perception of broad versus narrow focus in English, especially if the nuclear accent is in utterance-final position. These studies find that when an utterance having a final nuclear accented word does not contain any prenuclear accents, it is perceived as having narrow focus. However, when it does include prenuclear accents, the utterance is ambiguous between a broad or narrow focus interpretation ([5]).

Second, other perception investigations show that it is the relationship between the pitch occurring on the nuclear accent peak and that occurring on the prenuclear target words which determines a particular focus interpretation ([9]). In a two-peak utterance, listeners perceive broad focus when the two peaks display similar F0 values. Listeners perceive narrow focus on the *second* item when that item receives a larger peak value than that found on the prenuclear word. They perceive narrow focus on the *first* item when the second item is totally deaccented.

Since the presence/absence of prenuclear accents and the relationship between target peaks have been found to be instrumental in interpreting an utterance as having broad or narrow focus, the paper investigates whether such factors play a role in the production of broad versus narrow focus in Lebanese Arabic.

In addition to this qualitative investigation, the paper will examine whether broad and narrow focus affect the

examined acoustic cues quantitatively. In other words, the paper investigates whether the same prominence level will display significantly different pitch, amplitude, duration and formant frequency values depending on focus condition. Studies investigating the prosodic realization of narrow focus find that narrow focused items receive phonetically larger pitch movements than their broad focus counterparts [8]. Little work has examined differences in the non-tonal cues across different focus conditions.

2. THE EXPERIMENT: MATERIAL AND PROCEDURE

As mentioned in the previous section, the experiment presented here has three main aims. The first is to investigate whether the three prominence levels posited for Lebanese Arabic– lexical stress, accent and nuclear accent– display effects on F0, duration, RMS, and formant frequencies F1 and F2. The second is to examine interactions between intrinsic vowel quality and F0, duration, and RMS. The third aim of the experiment is to investigate qualitatively realizations of broad and narrow focus, and quantitatively possible differences in the examined acoustic cues across the two focus conditions.

The speech material for the experiment was designed as follows. The base test sentence was of the form “X Hamet Y min Z”, (X protected Y from Z), where X, Y and Z represent disyllabic CV.CV sonorant target words, with lexical stress falling on the penultimate syllable in each word. The three target words were: "lama", "muna", "lima". The stressed vowels across the three test words differed in vowel quality, illustrating the three main vowel phonemes in Arabic: /a/, /u/ and /i/. These different vowels were specifically chosen to test potential intrinsic pitch differences.

The linear order of each of the test words X, Y and Z was manipulated so that each target word appeared in the three possible phrasal positions– initial, medial and final– to allow the chosen prominence pattern to occur on all three target words. Six possible orderings were thus obtained: XYZ, XZY, YXZ, YZX, ZXY, ZYX. For each of these target utterances, four questions were designed to elicit different focus conditions. One question elicited broad focus, and three others placed narrow focus on each of the test words:

- shuu Saarel yoom
"What happened today?" (Broad focus)
- miin** Hama Y min Z
"**Who** protected Y from Z?" (Narrow focus on X)
- X Hamet **miin** min Z
"X protected **whom** from Z?" (Narrow focus on Y)
- X Hamet Y min **miin**
"X protected Y from **whom**?" (Narrow focus on Z)

(Note: the transliteration system used is based on [4]).

The material was recorded by three speakers of Lebanese Arabic (educated Tripoli dialect), one male and two females. 10 repetitions of 24 utterances (6 structural positions x 4 focus readings) for 3 speakers were analyzed, yielding a total of 720 test utterances constituting the corpus of the present experiment.

The corpus material was digitized at 22 kHz sampling frequency on a Sun workstation. Entropic ESPS Waves+ speech analysis software was used in order to generate F0, RMS and formant frequency trajectories. Standard segmentation and labeling criteria were carried out using the acoustic waveform in combination with the time-aligned displays of these three parameters.

An intonational transcription of the utterances was carried out, based on the tonal analysis posited for the language in [2], in order to label the prominence level of the target vowel. The prominence levels were hand-labeled using a separate tier in the transcriber function of ESPS Xwaves.

Figure 1 illustrates the spectrogram, syllable boundary marking, and the complete set of labels for the example utterance /muna Hamet lama men lima/ "Muna protected Lama from Lima". It also illustrates the accompanying pitch trace.

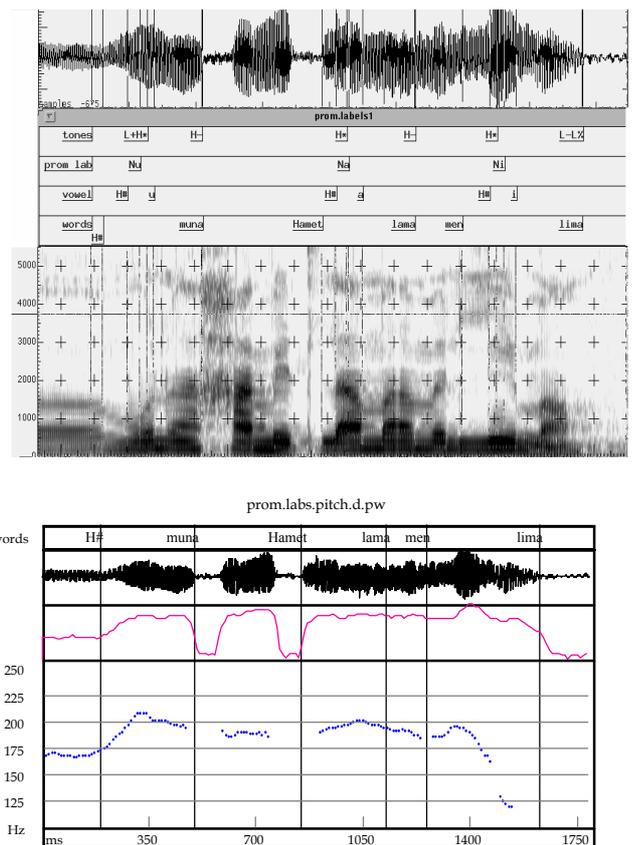


Figure 1. Illustration of the labeling system adopted in the present corpus, in line with the spectrogram, F0 and RMS displays.

The following measurements were automatically extracted using mu+ in the Splus environment, based on [6]:

- 1) The duration (in ms) of the target stressed vowel.
- 2) The F0 (in Hz), RMS (in dB), F1 (in Hz), and F2 (in Hz) extracted at the steady state of the vowel.

3. RESULTS

To test whether the prominence level of the target syllables, and the quality of their respective vowel nucleus affected any of the F0, duration, RMS, F1, and F2 measures in the corpus, a two-tailed factorial ANOVA was conducted for each speaker, with prominence status (three levels: nuclear accented, accented, unaccented), and vowel quality (three levels: /a/, /i/, /u/) as independent variables, and F0, duration, RMS, F1, F2 and duration as dependent variables. Post-hoc t-tests were also applied to determine which prominence levels, and which vowel types, contributed to the significant effects.

3.1 ACOUSTIC CUES TO PROMINENCE

The results show the three prominence levels—lexical stress, pitch accent, and nuclear accent—are generally differentiated by at least one, or a combination of F0, duration, RMS, F1 and F2. Concerning F0, duration and RMS specifically, nuclear accented vowels typically receive either higher F0, and/or longer duration, and/or higher RMS than their accented and unaccented counterparts. Accented vowels in turn receive either higher F0, and/or longer duration, and/or higher RMS than their unaccented counterparts. The strongest and most consistent differences, however, are those found between nuclear accented and unaccented vowels.

It is noted that for two out of the three speakers, F0 occupies a secondary status as an acoustic correlate differentiating the three prominence levels. This result can be explained in terms of the large number of flat hat patterns employed by these speakers. Because of flat hats, F0 differences were neutralized in these types of contours, and other non-tonal cues emerged as more primary in the judgment of the prominence of a particular token.

As for prominence effects on F1 and F2, these suggest that stressed but unaccented vowels become more peripheral in the vowel space when they are accented and/or nuclear accented. Alternatively, there is a general tendency toward centralization as the prominence level of the vowel decreases.

3.2 INTRINSIC VOWEL QUALITY EFFECTS

Despite speaker-specific differences, the vowel quality results reveal that vowel quality effects on F0, duration and RMS are most obvious at the nuclear accented level of prominence, and are mostly non-significant at the other levels of prominence, i.e. at the accented and unaccented levels. Generally, it was found that the low vowel /a/ displays lower F0, longer duration and higher RMS values than the high vowels /i/ and /u/.

3.3 REALIZATIONS OF FOCUS

This section considers the surface structure realization of broad and narrow focus in the language. It is noted that both broad and narrow focus utterances display obligatory nuclear accents, optional prenuclear pitch accents, and unaccented target syllables. As such, the presence or absence of prenuclear accents does not seem to play a crucial role in the distinction of focus condition in Lebanese Arabic. However, in the broad focus utterances, the accented, and unaccented tokens are usually realized at a high level pitch, similar to that of the nuclear accented word. In the narrow focus condition, however, the F0 peak of the non-focused target words is extremely compressed relative to the F0 peak of the narrow-focused nuclear accent. This indicates that it is the relationship among the peaks of the various target vowels of the utterance which signal a particular focus condition.

This point is further illustrated in narrow focus sentences having an utterance-final target word. It was found that these display two main patterns. In the first, all preceding material is deaccented, and the pitch range of the narrow focused target word is expanded. In the second pattern, the initial target word optionally receives an accent, and the overall realization of the contour is very similar to that of broad focus utterance. While the first set of realizations unequivocally denote a narrow focus interpretation, the second set of contours may be perceived as being ambiguous between a broad focus and a narrow focus interpretation. These results reflect those of [5] but need to be verified by future perceptual experiments.

3.4 QUANTITATIVE DIFFERENCES BETWEEN BROAD AND NARROW FOCUS

To test whether there are differences in F0, duration, RMS, F1 and F2 for target syllables having the same prominence level, but differing in focus condition, t-test comparisons were conducted for the dependent variable means, obtained in each of the broad focus and narrow focus utterances. The results of the t-tests show that for all speakers, narrow focused nuclear accented vowels display higher F0 and RMS values than their broad focused counterparts. For one speaker, narrow focused nuclear accents additionally display longer duration than their broad focus counterparts. Differences in F1 and F2 across the two focus conditions, on the other hand, emerge as largely speaker-specific.

As for the accented and unaccented prominence levels, applicable comparisons suggest that each of these two prominence levels tends to result in *lower* F0 and RMS values, but longer duration, when they occur in narrow focus versus broad focus conditions. Thus, while narrow focused nuclear accented syllables generally show higher F0 and RMS values than their broad focus counterparts, the accented and unaccented vowels within the narrow focused utterances show *lower* F0 and RMS values than their broad focus counterparts.

This inverse pattern across focus conditions suggests that in narrow focus, the difference between nuclear accented syllables, on the one hand, and accented and unaccented

syllables on the other, is attained not purely by realizing more extreme nuclear accents, but also by producing reduced contrasts among accented and unaccented syllables, in the sense of lower F0 and RMS values. In other words, in narrow focus, maximizing differences in prominence levels simultaneously increases the major acoustic cues for nuclear accented syllables, and decreases them for the other prominence levels. The decrease, however, only occurs to the extent that the accented and unaccented levels remain sufficiently distinct. In broad focus, on the other hand, nuclear accents are less distinct from their accented and unaccented counterparts, with no extreme differences occurring among the three prominence levels.

4. CONCLUSION

In the present paper, the acoustic cues to three phonological prominence levels posited for Lebanese Arabic were investigated. It was found that nuclear accented, accented and unaccented syllables are significantly distinguished in terms of a combination of F0, duration, RMS, F1 and F2 correlates. Briefly, the higher the prominence level of a particular vowel, the higher its pitch and amplitude, the longer its duration, and the more peripheral its spectral realization in the vowel space. These results generally reiterate not only the findings of various Arabic studies, but also those of English. Prominence in Arabic, like English, is thus cued by tonal as well as non-tonal cues. This confirms the typological classification of Arabic as a stress-accent language, similar to English.

The present study highlights the importance of analyzing the intonational pattern of the utterance in order to accurately investigate the phonetic cues to prominence. Analyzing the intonational pattern of utterances thus circumvents the possibility of confounding distinct prominence levels, and explains why F0 may not always feature as the most important cue to stress.

The present study also identifies certain intrinsic differences in F0, duration and RMS based on the vowel quality of the examined vowel, in particular its vowel height. All else being equal, it was generally found that the high vowels display intrinsically higher pitch, shorter duration and lower RMS amplitude than low vowels. These intrinsic differences based on vowel height resemble those found for English and Arabic. In addition, like [7], the present results indicate that these intrinsic differences are most obvious when the examined vowel is in nuclear position.

The investigation of broad versus narrow focus in the present study suggests that the relationship between target peaks within an utterance plays a role in signaling a particular focus interpretation. In broad focus, nuclear accented, accented and unaccented vowels may receive similar F0 pitch values (but are significantly differentiated in terms of non-tonal cues). Under narrow focus, pitch differences between the narrow focused nuclear accent and the other target words within the utterance are maximized.

As for the finding that narrow-focused nuclear accents receive higher F0 and RMS values than their broad focused counterparts, while narrow focused accented and unaccented tokens receive lower F0 and RMS values (but longer duration) than their broad focus counterparts, these suggest that in narrow focus, differences between the three prominence levels are achieved by maximizing the acoustic cues for nuclear accents, while maintaining minimally distinctive acoustic cues for accented and unaccented vowels.

REFERENCES

- [1] Beckman, M., & Edwards, J. (1990). Lengthenings and shortenings and the nature of prosodic constituency. In Kingston & Beckman (Eds.), *Papers in Laboratory Phonology, I: Between the Grammar and the Physics of Speech*, pp. 152-178. Cambridge: Cambridge University Press.
- [2] Chahal, D. (2001). *Modeling the Intonation of Lebanese Arabic using the Autosegmental-Metrical Framework: A Comparison with English*. Doctoral dissertation: University of Melbourne.
- [3] de Jong, K., & Zawaydeh, B. (1999). Stress, duration, and intonation in Arabic word-level prosody. *Journal of Phonetics*, 27, 3-22.
- [4] El-Imam, Y. (1990). Speech synthesis using partial syllables. *Computer Speech and Language*, 4, 203-229.
- [5] Gussenhoven, C. (1983). Testing the reality of focus domains. *Language and Speech*, 26, 61-80.
- [6] Harrington, J., Beckman, M., Fletcher, J., & Palethorpe, S. (1997). Kinematic and spectral measures of supralaryngeal correlates of the accent contrast in Australian English vowels. *Journal of the Acoustical Society of America*, 102, 5, 3205.
- [7] Ladd, D. R., & Silverman, K. (1984). Vowel intrinsic pitch in connected speech. *Phonetica*, 41, 31-40.
- [8] Norlin, K. 1989. A preliminary description of Cairo Arabic intonation of statements and questions. *Speech Transmission Quarterly Progress and Status Report*, 1, 47-49.
- [9] Rump, H., & Collier, R. (1996). Focus conditions and the prominence of pitch-accented syllables. *Language and Speech*, 39, 1, 1-17.