Acoustic cues to focus and givenness in Egyptian Arabic

Abstract

Hellmuth (2009) showed that contrastively focussed items are realised in an expanded pitch range in Egyptian Arabic (EA), whereas items following such a focus are realised in a compressed pitch range (cf. Norlin 1989 for EA; and Chahal 2001 for LA). No equivalent variation in f0 excursion was found depending on whether the item was new to the discourse or given (repeated from earlier in the discourse). The present study presents analysis of F0 excursion, duration, overall intensity and spectral tilt, in a directly parallel dataset, in order to test whether any of these correlates are employed in the expression of givenness in EA; focus is found to be marked by f0 excursion only, but no prosodic correlates of givenness are observed.

1 Introduction

This paper explores the prosodic expression of focus and givenness in Egyptian Arabic (EA), defined here as the colloquial variety of Arabic spoken in Cairo and by educated speakers throughout Egypt.

Krifka (2006) suggests that there are just three basic information structure (IS) categories: focus, givenness and topic. It is focus and

givenness that are relevant here, and for the purposes of the present paper we adopt Krifka's definitions as in (1) and (2) below (paraphrased):

(1) Focus indicates the presence of alternatives that are relevant for the interpretation of linguistic expressions.

(2) Givenness indicates whether the denotation of an expression is present in the common ground or not, and/or the degree to which it is present in the immediate common ground.

In practical terms then, here focus denotes a choice between alternatives, similar to the commonly used notion of contrastive focus, (cf. Rooth 1996), whereas givenness denotes availability, or prior mention, in the discourse (Halliday 1967;cf. Schwarzschild 1999)[1].

IS categories such as focus and givenness can be marked in different ways in different languages, with use of syntactic as well as phonological or phonetic cues. A common way in which languages have been observed to express focus and givenness prosodically involves the presence vs. absence of intonational pitch accents, known as deaccentuation (Cruttenden 2001). It has been observed in EA however that deaccentuation generally fails to

[1] Kiss (1998) distinguishes i) identification focus, of which contrastive focus is a subtype, and which equates to the definition of focus used here; and ii) information focus, which equates to Halliday's (1967) notion of new (vs. given) information. Krifka (2006) argues for a simple two-way distinction between focus and givenness, which is adopted here.

occur in contexts which would condition deaccentuation in other languages such as English (Hellmuth 2005). The canonical context for de-accenting in English and other Germanic languages is when a given item follows a focused item (Selkirk 2000;Cruttenden 2006), so that the second occurrence of the lexical item 'car' in the following sequence would be realised without an accent: *"Was that a blue car? No, it was a RED car."* (capital letters denote focus). In a parallel sequence in EA, the word 'car' would not be de-accented, and the same lack of de-accenting is found in parallel contexts in a number of Romance languages (Swerts et al. 2002;Cruttenden 2006;Swerts 2007;Ladd 2008). Nonetheless, there are reports in the literature of use of gradient variation in f0 excursion as a marker of focus in Arabic (EA: Norlin 1989; Lebanese Arabic: Chahal 2001), with the result that even if both words in a phrase like 'red car' bear a pitch accent, variation in the excursion size of the accent f0 (on both words) results in greater prominence on the focused word, and reduced prominence on the following word.

The intonational patterns of spoken colloquial EA have been investigated instrumentally (Abdalla 1960;Norlin 1989), and impressionistically, within the British School of intonational description, (Soraya 1966;Mitchell 1993), and in a number of Autosegmental-Metrical (AM) analyses which combine instrumental and impressionistic analysis (Rifaat 1991;Rastegar-El Zarka 1997;Hellmuth 2006b). A shared finding of all of these descriptions is that in EA a pitch movement (a 'pitch accent') is observed on almost every content word, rather then being somewhat more

sparsely distributed across the utterance as in, say, English. This generalisation has been shown to hold across a range of speaking styles (Hellmuth 2006b) and has been formalised in terms of association of a pitch accent with every Prosodic Word (Hellmuth 2007). Crucially, for our present purposes, the generalisation holds even when a word is given, and would be de-accented in a language like English, as in the example below from Soraya (1966: 177), reproduced in (3) below with his notation; in (3b) the lexical item [itta=faːʔul] is repeated from the preceding question but nonetheless bears a tone (a pitch accent), albeit a low level one:

(3) a. b-it-ħibb itta.faːʔul
you-like optimism
'Do you like to be optimistic?'

b. ˋʔaah ˋʕaːjiz itta=faːʔul
yes I-want optimism
'Yes I want to be optimistic'

- high level tone
ˋ falling tone
= low level tone
. tail pitch

As part of a wider study of the properties of intonational pitch accents in EA, Hellmuth (2009) investigated both the presence/absence of pitch accents and their f0 excursion properties in words appearing under varying information structure conditions. An experimental paradigm was used in which a pair of lexically distinct SVO sentences were placed in frame paragraphs designed to manipulate the focus status of the subject of the sentence (focussed '+F' vs. non-focussed '-F') and the givenness status of the object of the sentence ('+f' vs. '-f'), in a 2 x 2 comparison permitting a fourway comparison of values of f0 excursion. An earlier study, Norlin

(1989), which measured f0 excursion in the SVO productions of a single speaker, reported that a post-focal object was not fully deaccented but showed reduced f0 excursion ('f0 compression')[2]. Unfortunately Norlin did not report the focus context used to elicit the SVO sentences, so it is not possible to know whether the effect he observed results from the post-focal position of the object or its own givenness status (or both). The 2 x 2 design used in Hellmuth (2009) was designed to overcome this problem and disambiguate between potential prosodic marking of focus and givenness, and data was collected with six speakers (3 male, 3 female).

Fig. 1 below, reproduced from Hellmuth (2009:177), shows mean values of f0 excursion in the subject (*xxn*) by focus condition and by speaker in the SVO dataset. The patterns of f0 excursion produced by female speakers are as expected, with greater mean f0 excursion in focussed (+F) contexts as compared to non-focussed (-F) contexts. The male speakers exhibit considerably more variation. Among female speakers statisitical analysis revealed that mean values of *xxn* varied significantly between [+F+f]~[-F+f] and [+F-f]~[-F-f] only, directly reflecting focus status.

As for f0 excursion in objects (*yxn*), this could have varied either according to whether the object follows a focus or not, or according to the givenness status of the object itself, or both. Fig. 2, reproduced from Hellmuth (2009: 178) below, shows mean values of f0 excursion in the object (*yxn*). These data showed more homogeneity across all speakers, and

[2] Soraya (1966:177) also reports that low level tones are found in post-focal positions.

the trend observed was found to be in the direction expected if f0 excursion marks the post-focus status of the object (±F), rather than its own givenness status (±f).

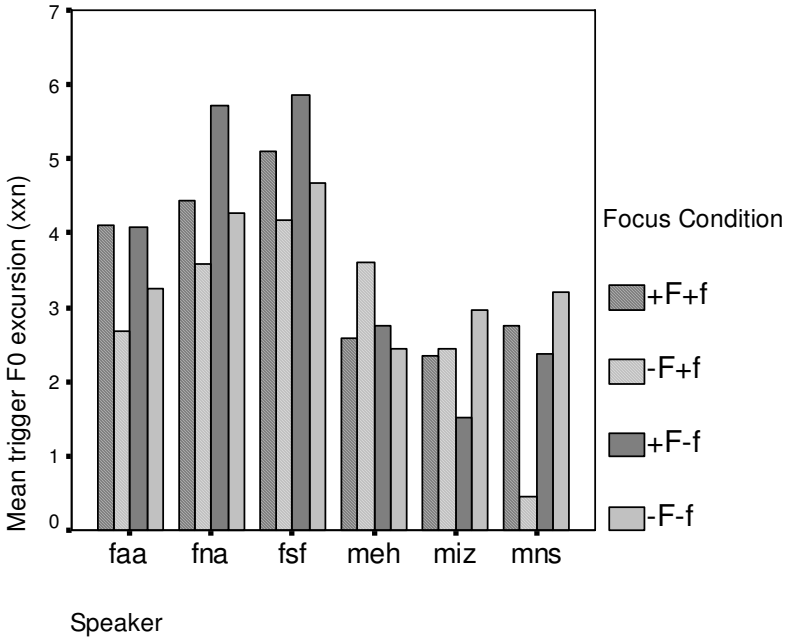

Fig. 1 Mean f0 excursion in the subject (*xxn*) (Hellmuth 2009: 177).

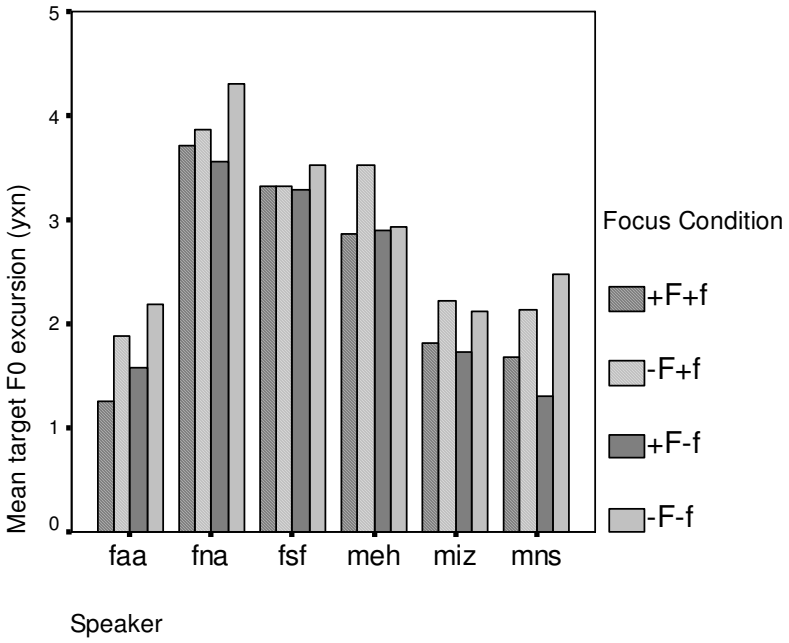

Fig. 2 Mean f0 excursion in the object (*yxn*) (Hellmuth 2009: 178).

Mean f0 excursion in the object was generally smaller when it followed a focussed subject (in +F conditions), than when it followed a non-focussed subject (-F conditions), indicating f0 compression of post-focal items, and these differences were statistically significant. Crucially however, when grouped by the givenness status of the object itself (+f vs. -f) the values of *yxn* showed no significant differences. The Hellmuth (2009) study thus showed that in EA variation in f0 excursion varies according to focus status, but does not vary according to givenness status.

The quantitative investigation in the Hellmuth (2009) study was limited to f0 excursion and duration (which did not vary with either focus or

givenness)[3]. The purpose of the present study is to test whether a key potential alternative prosodic correlate - intensity - reflects givenness status in EA. We therefore here investigate both overall intensity and selective intensity (spectral tilt), alongside f0 excursion and duration, in a dataset directly parallel to the one analysed in Hellmuth (2009), recorded with the same speakers and at the same recording session[4].

What findings are expected? It is possible that we will find that intensity does vary with givennes status, even though f0 excursion does not, and this would match findings for EA's Afroasiatic neighbour Beja, which is reported to use intensity as a marker of focus (Vanhove 2005). Alternatively, it is possible that there are simply no prosodic correlates of givenness realised on overt lexical items in EA.

The acoustic correlates of prominence investigated in the present paper are thus f0 excursion, duration of the accented syllable (or rhyme), overall intensity and spectral tilt. These acoustic cues are known to be used in a number of languages to mark accentual prominence at the word-level (van Heuven & Sluijter 1996), and may be considered in two subsets: melodic cues (f0 excursion) vs. dynamic cues (duration and intensity). Languages are known to vary in whether only melodic cues or both melodic and dynamic cues are used to mark prominence (Beckman 1986;Ladd 2008).

[3] Hellmuth (2006a) also showed that, in the same dataset, there were no effects of either focus or givenness status on the peak alignment of pitch accents realised on either subject or object.

[4] A further potential prosodic correlate would be vowel quality, which we leave for future research.

Relatively little work has been done on the correlates of prominence in spoken varieties of Arabic (see Chahal & Hellmuth to appear for an overview). Although the consensus thus far seems to be that both melodic and dynamic cues are used to mark prominence in Arabic, some authors such as Mitchell (1993) suggest that there may be cross-dialectal variation. DeJong & Zawaydeh (1999) investigated duration and f0 (inter alia) as potential correlates of word-level prominence in Ammani Arabic. They found that both duration and f0 were greater in stressed syllables than in unstressed syllables, but could not determine whether increased f0 was a word-level or phrase-level cue because the tested 'stressed' syllables happened also to bear intonational pitch accents. They chose to analyse f0 as a cue to phrase-level prominence on the assumption that pitch accents on stressed syllables are optional in that dialect (ibid. p20)[5]. As discussed above however, pitch accents are not optional on stressed syllables in EA (Hellmuth 2007), thus opening up the possibility that the range of cues used at differing prominence levels may differ in EA from other varieties of Arabic. Since post-focal given items in EA are both accented and stressed, the present study may shed some light on whether dynamic cues (as well as melodic) are used to mark prominence at different levels in EA.

In the remainder of this paper we set out the methodology of the experimental study (section 2) and the results of the quantitative investigations (section 3). The findings are discussed in section 4.

[5] Similar results were found by Chahal (2001) for Lebanese Arabic (LA), see section 4.

2 *Methodology*

Two lexically distinct ditransitive sentences, as shown in 4) below, were each placed in one of four context paragraphs designed to manipulate the information structure of the sentence.

(4) Target sentences used in the focus experiment.

			direct object	*indirect object*	
A	ˈhuwwa he *'He*	bijˈdarris teaches *teaches*	ˈʕarabi Arabic *Arabic*	li-l-ʔaˈgaːnib to-the-foreigners *to foreigners*	bi-l-lajl in-the-evening *in the evenings'*
B	ˈhijja she *'She*	biˈtaʕmal makes *makes*	zaˈlabja cake *cake*	li-wiˈlaːd-i for-children-my *for my children*	kull joːm every day *every day'*

The *focus* status of the direct object and the *givenness* status of the indirect object were systematically varied, resulting in the four possible focus/givenness combinations between the direct and indirect objects shown in (5) below. A translation of the eight context paragraphs used to elicit the four focus conditions for each target sentence is provided in (6).

(5) Four *focus* x *givenness* conditions in which sentences were elicted.

		direct object	*indirect object*	
1	[+F+f]	contrastive focus	new	'he teaches Arabic $_{[+F]}$ to foreigners$_{[+f]}$'
2	[+F-f]	contrastive focus	given	'he teaches Arabic $_{[+F]}$ to foreigners $_{[-f]}$'
3	[-F+f]	no contrast, new	new	'he teaches Arabic $_{[-F]}$ to foreigners $_{[+f]}$'
4	[-F-f]	no contrast, new	given	'he teaches Arabic $_{[-F]}$ to foreigners $_{[-f]}$'

(6) Context paragraphs used to elicit focus/givenness conditions.

A1	My friend said they heard my father's brother teaches English in the evenings but I told them no, *he teaches Arabic$_{[+F]}$ to foreigners$_{[+f]}$ in the evenings*. My mother's brother teaches English.
A2	My uncle is very busy with teaching. *He teaches Arabic$_{[-F]}$ to foreigners$_{[+f]}$ in the evenings* and he also has a job at the Ministry of Education.
A3	My friend said they heard my father's brother teaches English to foreigners in the evenings but I told them no, *he teaches Arabic$_{[+F]}$ to foreigners$_{[-f]}$ in the evenings*. My mother's brother teaches English.
A4	My uncle teaches foreigners. *He teaches Arabic$_{[-F]}$ to foreigners$_{[-f]}$ in the evenings* and he also has a job at the American Embassy.
B1	My neighbour said she heard our maid makes balaħ-iš-šaam every day but I told her no, *she makes zalabya$_{[+F]}$ for my kids$_{[+f]}$ every day*. I make the balaħ-iš-šaam.
B2	Our maid loves making sweet things. *She makes zalabya$_{[-F]}$ for my kids$_{[+f]}$ everyday* and she makes the best kunaafa in Cairo.
B3	My neighbour said she heard our maid makes balaH-iš-šaam for my kids every day but I told her no, *she makes zalabya$_{[+F]}$ for my kids$_{[-f]}$ every day*. I make the balaħ-iš-šaam.
B4	Our maid loves making sweet things for my kids. *She makes zalabya$_{[-F]}$ for my kids$_{[+f]}$ every day* and she often brings them cakes she has made at home.

Differences in givenness (of the indirect object) were generated based on whether or not the word is repeated from earlier in the paragraph (that is, whether or not the word is 'textually given' (Halliday 1967)), and thus unambiguously available or unavailable in the common ground. The direct object was designed to be 'new' [+f] in all contexts, and its focus status was manipulated by means of a direct contrast with a previously mentioned alternative ('English'~'Arabic', 'balaħ-iš-šaam'~'zalabja', which are two different kinds of cake).

A final adverbial phrase, [bi-l-lajl] ‘in the evenings’ or [kull joːm] ‘every day’, was included at the end of the target sentences in order to elicit a pre-nuclear rather than nuclear pitch accent on the indirect object, since pitch excursion in phrase-final pitch accents is known to be vulnerable to the effect known as final lowering, resulting in highly compressed pitch range, due to proximity to the phrase boundary (Chahal & Hellmuth to appear). The context paragraph was designed so that the final adverbial phrase would always be given and not inadvertently elicit focus on the adverbial by generating alternatives (such as teaching in the evenings vs. (inferred) teaching during the day).

The full dataset comprises 8 paragraphs (2 lexical sets x 4 focus contexts) which were interspersed among a somewhat larger number of paragraphs related to a separate study, then pseudo-randomised and divided into sets such that no set contained two paragraphs from any one lexical set. Each of the 8 sets of paragraphs was read 3 times by 6 speakers of EA, yielding 18 tokens x 8 targets (N=144) for analysis.

The paragraphs were presented in written form, in printed Arabic script. Presentation of a text in written Arabic carries the risk of inadvertently eliciting a more formal register of Arabic than the desired colloquial EA. This was mitigated by inclusion of specifically colloquial lexical items in both target and filler paragraphs (such as the various cake names in lexical set B), and by use of a modified, colloquial, spelling

system similar to that used in informal written texts such as personal letters, following Siemund et al. (2002).

Speech recordings were made in southern Cairo, with speakers recruited at a private English school among students in classes at pre-intermediate level or lower in English. As a result none had advanced proficiency in any other language besides Arabic. Six speakers participated, three female and three male. All were mother tongue speakers of EA, born and raised in Cairo, aged between 21-34 years, and none had any auditory or speech production difficulties. Recordings were made on the school premises and the speakers were paid a small fee for their participation.

It would have been ideal to have subjects read each set of paragraphs on different occasions, to exclude the possibility of interference between different contexts, however, this was not possible in the recording time available. To reduce potential interference, after the third repetition of each set the speaker performed a different style of task, which they found diverting and in most cases thought was the real purpose of the recording session. This rotation of tasks served to break up the pace and pattern of the recording session, and facilitate interpretation of each paragraph from its own internal structure, rather than in comparison with paragraphs in other sets.

The fact that recordings were made in a single session - albeit broken up among different tasks - means that it is possible to extract reliable intensity measurements from the data. Recordings were made using a head-

mounted AKG C-420 cardioid condenser microphone (frequency range 20-20,000 Hz; SNR 64dB). Participants were asked to avoid repositioning the microphone, and the gain control was left unchanged through the session for each individual speaker. The data was saved directly to digital format using ProTools 6.0 on MBox at 44100Hz 16bit, then re-sampled at 22050Hz 16bit for analysis.

The 144 recordings were hand-labelled by the author, with reference to F0 and spectrogram using Praat 5.1 (Boersma & Weenink 2007), to identify: i) the start and end point of the stressed vowel of each target word (direct object + indirect object); ii) the position of the low (L) and high (H) turning points of the rising pitch movement associated with the stressed vowel of each target word; and iii) the intensity peak within each target stressed vowel (identified in an Intensity object generated with minimum pitch 100Hz and a time step of 0.8/100Hz, with mean pressure subtracted). A sample labelling scheme is provided in Fig. 3 below.

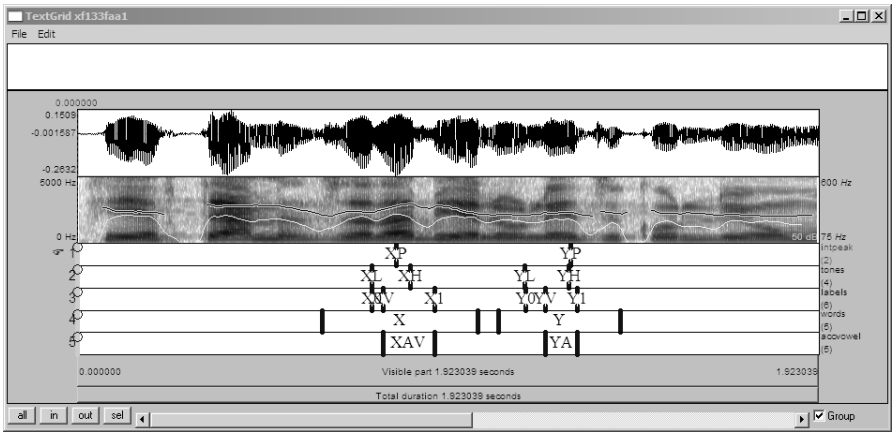

Fig. 3 Sample labelling scheme (in token 133faa1)

The following dependent variables were extracted from the data: i) f0 excursion in the direct object (*xxn*) and indirect object (*yxn*); ii) stressed syllable duration in the direct object (*xdur*) and indirect object (*ydur*); iii) overall intensity at various points in the direct object (*xmp*, *xip*, *xav*) and indirect object (*ymp*, *yip*, *yav*); and iv) two measures of spectral tilt, at different points in the indirect object only (*h1h2mp*, *h1h2ip*, *h1a3mp*, *h1a3ip*).

In order to calculate f0 excursion in the direct object (*xxn*) and indirect object (*yxn*), the f0 at each of the L and H points in each word was extracted in semitones, and f0 excursion within each word calculated (Hf0 - Lf0). Semitones were used to permit comparison among speakers of different genders (Nolan 2003)[6]. F0 excursion was the primary dependent variable of interest in Hellmuth (2009), and we expect to reproduce the findings of that study here, in this previously uninvestigated but directly parallel dataset. No significant differences in stressed syllable duration across focus contexts were found in the earlier study, but duration is included here for completeness.

The primary dependent variable of interest in the present study is intensity, and both overall intensity (across all frequencies) and a measure of spectral tilt (at different frequencies) are investigated. Overall intensity (in dB) was extracted at two relevant points in each target word: '*mp*', which is

[6] In figures, pitch traces are presented scaled logarithmically in Hertz for ease of interpretation.

the midpoint of the stressed vowel, and '*ip*', which is the intensity peak within the stressed vowel (following van Heuven & Sluijter 1996). Average intensity over the whole stressed vowel, '*av*' was also calculated.

Two measures of spectral tilt were extracted at the midpoint (*mp*) and intensity peak (*ip*) of the stressed vowel of the indirect object in each token: H1-H2, which is the amplitude at the first harmonic minus the amplitude at the second harmonic (Sluijter et al. 1995), and H1-A3, which is the amplitude at the first harmonic minus the amplitude of the third formant (Campbell & Beckman 1997). The spectral tilt measures were extracted using a Praat script developed and made available by Bert Remijsen[7]. A multivariate Analysis of Variance (MANOVA) was carried out with all ten measures as dependent variables, and with Focus Set and Gender as fixed factors using SPSS 16.0. To compensate for the risk of Type I error, due to the fact that measures of each of the various dependent variables were taken from data from the same set of participants, a more restrictive significance level ($\alpha = 0.005$) was applied following a Bonferroni adjustment (α divided by the number of tests: $0.05/10 = 0.005$) (Rietveld & van Hout 2005:177)[8]. The full MANOVA table is provided in the Appendix and relevant details discussed in section 3 below.

[7] http://www.ling.ed.ac.uk/~bert/praatscripts.html (accessed 22nd April 2009).
[8] We are grateful to an anonymous reviewer for suggesting this approach.

3 Results

In this section we present the results of each of the quantitative analyses in turn, for f0 excursion, duration, overall intensity and spectral tilt. For f0 excursion we also present sample f0 pitch traces, as an example of the use of f0 excursion in the expression of focus (but not givenness).

3.1 F0 excursion

Averaged across all speakers, F0 excursion is relatively larger in focussed direct objects than in non-focussed ones, as illustrated in Fig. 4 (though the trend is reversed among male speakers in the [+F+f]~[-F+f] pair).

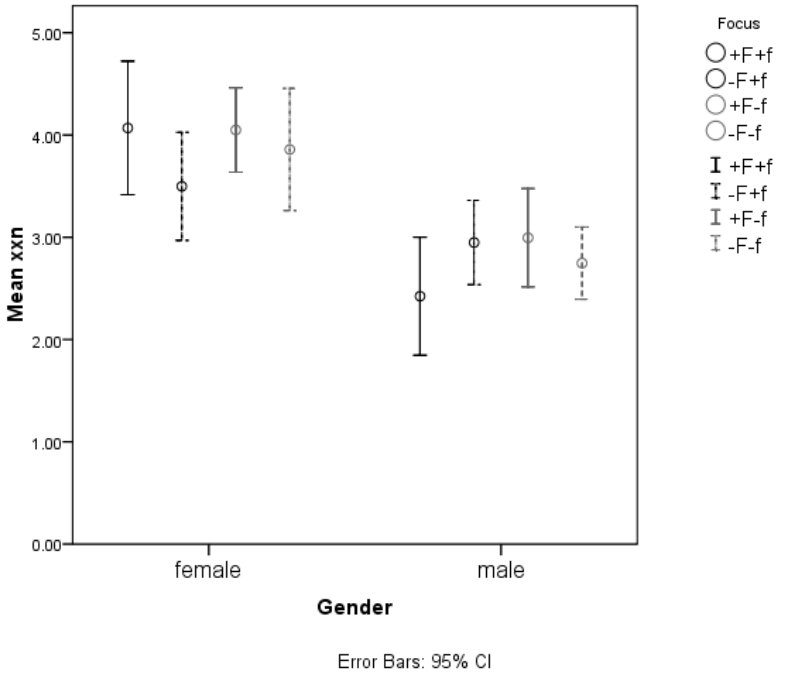

Fig. 4 Means and 95% confidence intervals for values of f0 excursion in direct objects (*xxn*), by focus and by gender in semitones.

If we break this down by individual speaker as in Fig. 5 below, we see the same pattern of speaker variation as was observed in the Hellmuth (2009) dataset. For f0 excursion in the ± focussed direct object (*xxn*) we expect to see increased excursion in focussed (+F) cases, compared to non-focussed (-F) cases: in the present dataset the expected pattern is only produced by two of the female speakers (*faa*, *fna*); the other speakers show either the reverse pattern (*fsf*, *miz*) or no particular pattern at all (*meh*, *mns*). This matches the findings in Hellmuth (2009) in which there was also considerably more variation in the productions of the male speakers than the female speakers.

The MANOVA analysis (see Appendix) shows a main effect of Gender (as is to be expected: ($F=40.474$, $df=1$, $p<0.001$)) but no main effect of Focus Set ($F=0.634$, $df=3$, $p=0.594$) and no interaction ($F=1.709$, $df=3$, $p=0.168$). The expected pattern of ±F effects on *xxn* is thus observed descriptively for two of the three female speakers but does not result in a statistically significant effect of focus overall. This contrasts slightly with the more homogenous findings in the parallel SVO dataset investigated in Hellmuth (2009), in which a main effect of Focus was found among all female speakers, but the trend in the present dataset is in the same direction.

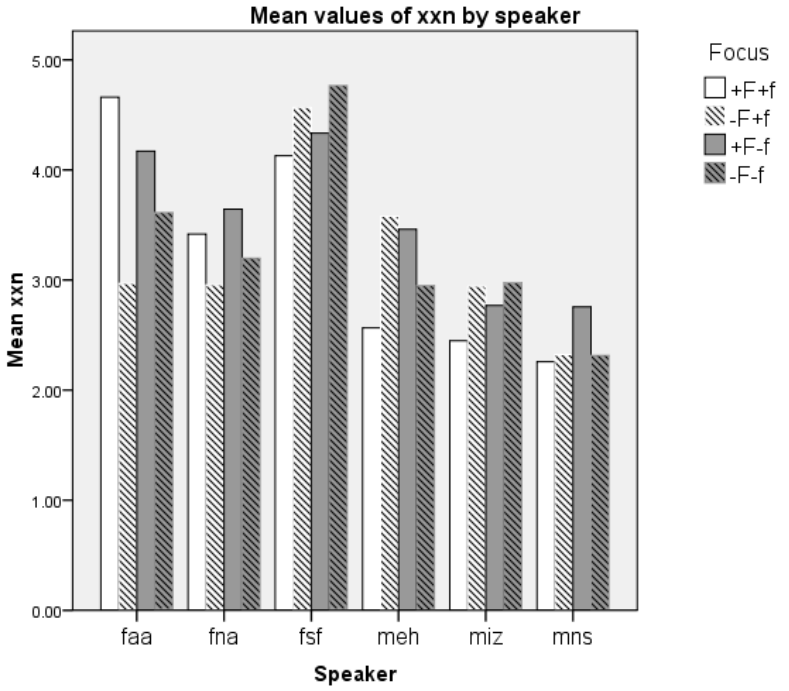

Fig. 5 Mean values of f0 excursion in the direct object (*xxn*) by speaker.

Of more interest here is potential variation in f0 excursion in the indirect object (*yxn*), since this is where independent effects of focus and givenness can be observed (if any). In the original SVO study, post-focal compression was observed in post-focal contexts, but there was no effect of givenness status on f0 excursion. If we look at the *yxn* results for each individual speaker, as shown in Fig. 6 below, we can see that three of the speakers followed this pattern (*faa*, *meh*, *miz*), whereas a mixed pattern was observed in two speakers (*fna* and *fsf*) and one speaker showed the reverse pattern (*mns*). Crucially however, no speaker shows a pattern in which it is givenness status (±f) rather than post-focal status (±F) which determines levels of f0 excursion in the indirect object[9], and none of the differences in values of *yxn* in different focus conditions are significant. The MANOVA analysis for *yxn* shows a main effect of Gender (as is to be expected: $F=39.725$, $df=1$, $p<0.001$) but no main effect of Focus Set ($F=0.332$, $df=3$, $p=0.802$) and no interaction ($F=1.528$, $df=3$, $p=0.210$).

[9] The nearest to this is speaker *miz* who shows greater f0 excursion in given conditions (-f) than in new conditions (+f).

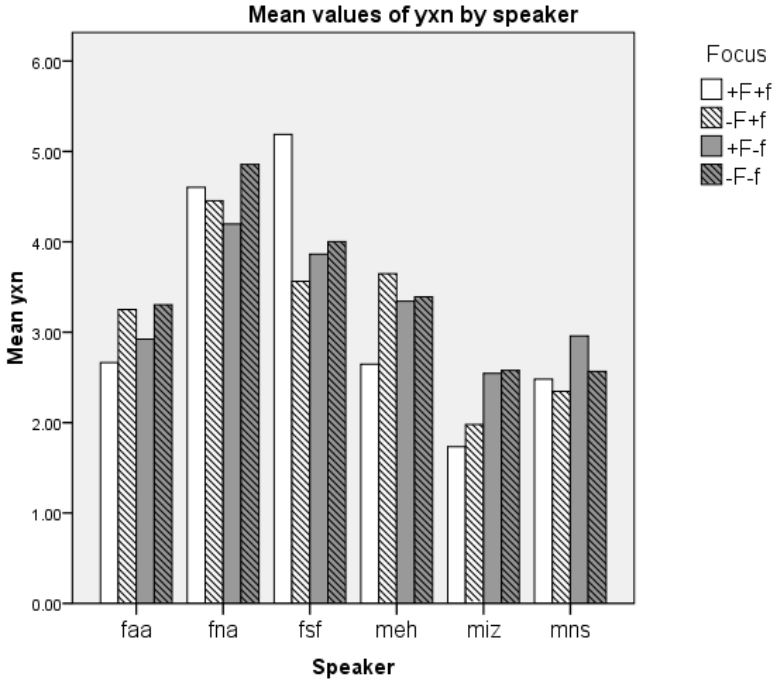

Fig. 6 Mean values of f0 excursion in the indirect object (yxn) by speaker.

In sum then, the general pattern of f0 excursion variation observed in the SVO study is reproduced here, albeit to a lesser degree: f0 excursion is increased on words bearing focus (observed in two speakers of the six here), and compressed on words following a focus (observed in three of the six speakers). However, the givenness status of the post-focal word is not reflected in f0 excursion (observed in five of the six speakers). We can see this pattern in the qualitative pitch contours produced by the speaker (*faa*) who most closely followed the overall trend, illustrated in Figures 7-10 below (also showing overall intensity, see section 3.3 below). Note

expanded f0 excursion in the direct object [za'labja] in Figs. 7 and 9 as compared to Figs. 8 and 10.

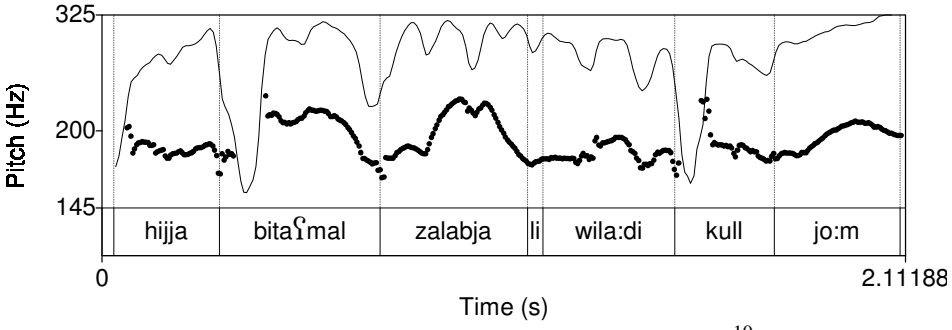

Fig. 7 Pitch trace and intensity curve in token 131faa1[10] [+F+f]

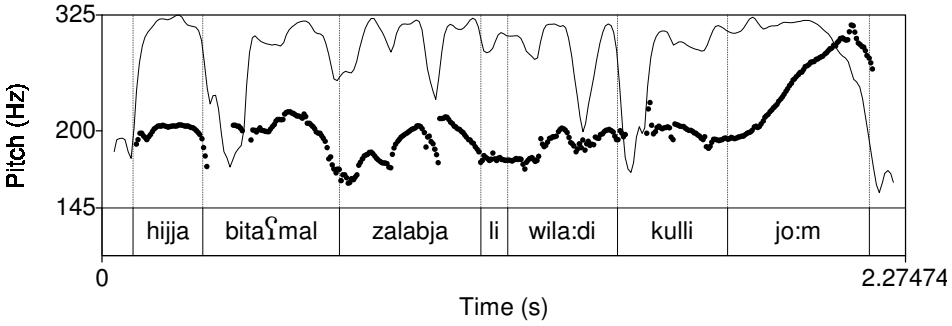

Fig. 8 Pitch trace and intensity curve in token 132faa1 [-F+f]

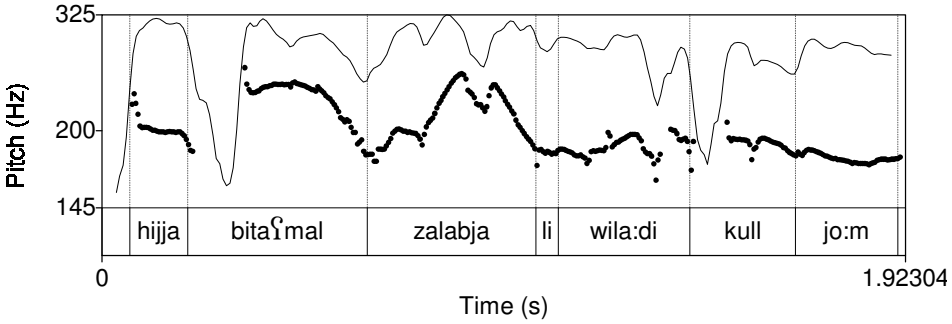

Fig. 9 Pitch trace and intensity curve in token 133faa1 [+F-f]

[10] In Figures 7-10 the pitch trace is shown in bold and the intensity curve with a thin line.

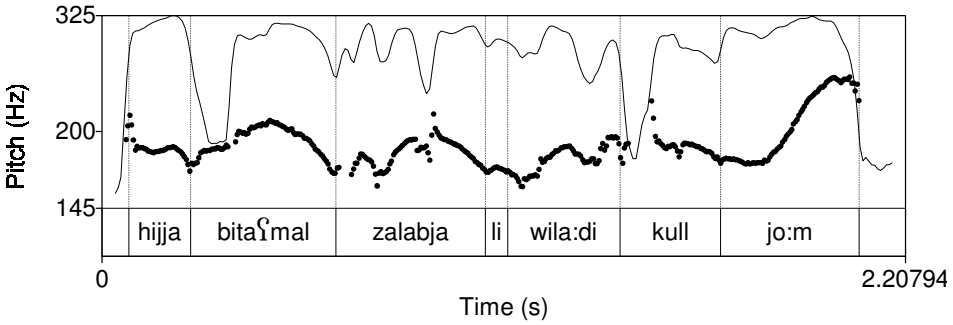

Fig. 10 Pitch trace and intensity curve in token 134faa1 [-F-f]

3.2 Duration

The original Hellmuth (2009) study found no differences in the duration of the accented syllable of the target words (subject/object) across focus conditions. The same holds in the present dataset also: there is no variation in the duration of the accented syllable of the direct object (*xdur*) or indirect object (*ydur*) across focus conditions. The MANOVA result shows no main effect of Gender or Focus Set, and no interaction (see Appendix), for the duration of the direct object *xdur*, nor for the duration of the indirect object.

Any variation in the duration of the stressed syllable of the target words in the present dataset (direct object and indirect object) is thus attributable to individual speaker variation (e.g. differences in speech rate) or the differing syllabic structure properties of the target words themselves. In sum, neither focus nor givenness influence stressed syllable duration.

3.3 Overall intensity

In this section we report the result of overall intensity measurements taken at various points in the stressed vowel of the indirect object ([ʔaˈgaːnib] / [wiˈlaːdi]) only. Intensity was not measured in the direct object ([ˈʕa.ra.bi] / [za.ˈlab.ja]) because the vowel quality in the stressed syllable is different in the two target words.

Overall intensity is reported here in dB, and was measured in Praat in an Intensity object created for each sound file, generated with minimum pitch 100Hz and a time step of 0.8/100Hz, with mean pressure subtracted. Three measuring points were used (van Heuven & Sluijter 1996): *ymp*, the midpoint of the accented vowel of the indirect object; *yip*, the intensity peak within the accented vowel of the indirect object (identified during labelling, as described in section 2), and average intensity over the whole accented vowel of the indirect object (*yav*).

The MANOVA analysis (see Appendix) shows no main effect of Gender or Focus Set and no interaction for any of the three overall intensity variables *ymp*, *yip* and *yav*, indicating that intensity values in the direct object do not vary significantly according to focus condition.

A more revealing way to analyse the intensity of the indirect object might be to address its intensity relative to that of the preceding direct object, in order to take into account the relative intensity of each whole utterance. We calculated the intensity differential (*idav*) between the average intensity in the accented vowel of the direct object (*xav*) and the average intensity in

the accented vowel of the direct object (*yav*). Unlike f0 excursion we do not necessarily expect an effect of post-focal compression (yielding a positive differential in every case, so that the direct object would be realised at greater intensity than the indirect object), but we might expect the differential to vary according to focus condition. We do indeed find variation in positive vs. negative differentials among speakers, as illustrated in Fig. 11 below, with five of the six speakers, consistently showing a positive differential, and one a negative differential (speaker meh). The differentials are all small however (none greater than 3dB) and although the MANOVA analysis for *idav* shows a main effect of Gender ($F=32.513$; $df=1$; $p<0.001$), which we attribute to the variant differential of speaker meh, there is no main effect of Focus Set ($F=0.198$; $df=3$; $p=0.898$) and no interaction ($F=0.504$; $df=3$; $p=0.680$).

In sum then, there are no significant differences due to focus condition observed in values of overall intensity in the dataset.

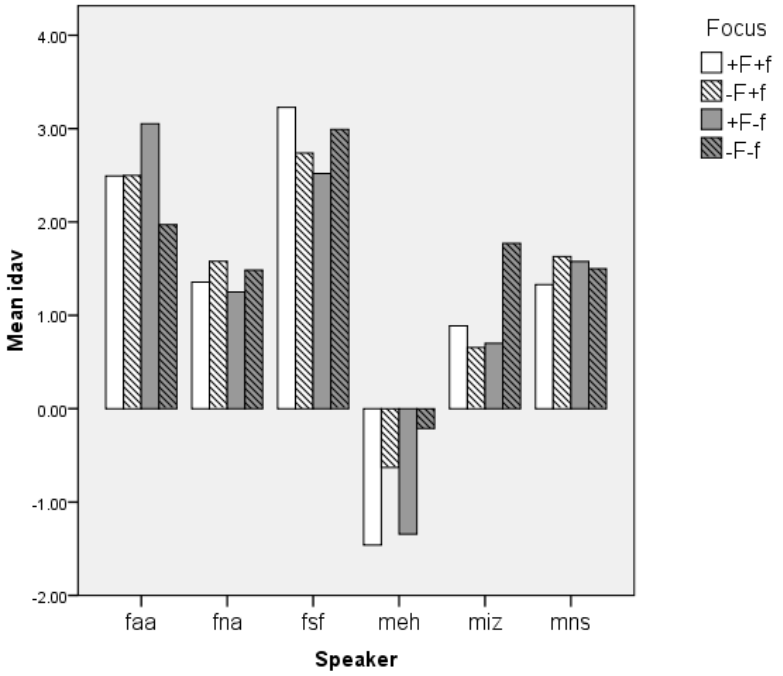

Fig. 11 Mean values of the differential in dB between average intensity in the accented vowel of the direct object and of the indirect object (*idav*).

3.4 Spectral tilt

In this section we report the result of spectral tilt measurements taken at two points in the accented vowel of the indirect object ([ʔaˈgaːnib]/[wiˈlaːdi]).

Initially it was planned to use two different measures of spectral tilt: H1-H2, which is the amplitude at the first harmonic minus the amplitude at the second harmonic (Sluijter & van Heuven 1995), and H1-A3, which is the amplitude at the first harmonic minus the amplitude of the third formant (Campbell & Beckman 1997). The results for H1-H2 were very variable

however, showing as many cases with negative slope as with positive slope, and this variation was not systematic (by speaker or by focus condition) which we interpret as indicating free variation in voice quality both across and within speakers. We report only the H1-A3 results therefore, as these are undisturbed by voice quality fluctuations.

The results show that spectral tilt values do not vary significantly according to focus condition, as illustrated in Fig. 12 and 13 below.

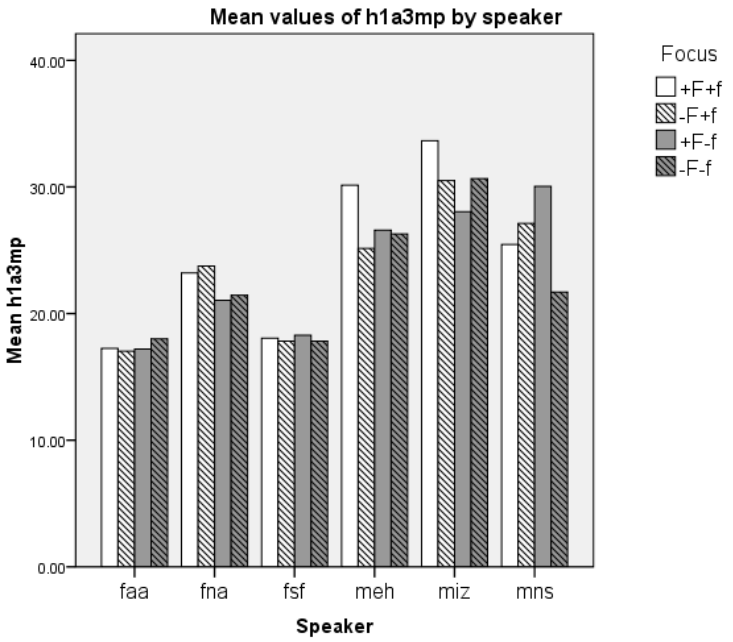

Fig. 12 Mean values of H1-A3 at midpoint of indirect object accented vowel.

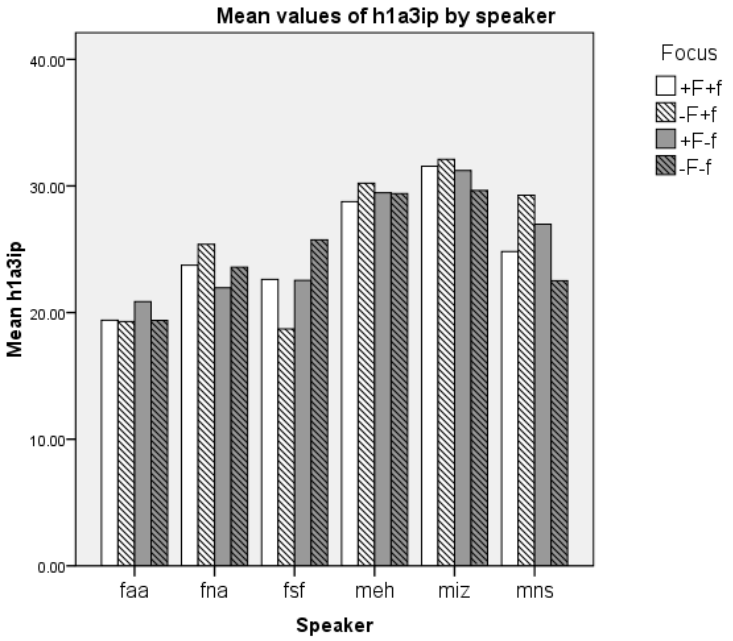

Fig. 13 Mean values of H1-A3 at intensity peak of indirect object accented vowel.

Any small differences observed are not statistically significant: the MANOVA analysis shows a main effect of Gender, as is to be expected, in both measuring points (*h1a3mp*: F=126.297; df=1; p<0.001; *h1a3ip*: F=68.767; df=1; p< 0.001), but there is no main effect of Focus Set and no interaction, at either measuring point (see Appendix).

In summary then, there appears to be no variation in selective intensity (spectral tilt) according to either focus or givenness.

3.5 Summary of results

The VOO dataset analysed here shows a similar basic pattern to the original SVO dataset (reported in Hellmuth 2006a, 2009), in that although there is considerable speaker variation in behaviour, there is a trend towards greater f0 excursion on focussed items, and compressed f0 excursion after such an item. This trend is not matched in changes in duration however (also similar to the original SVO results).

The main purpose of the present study was to explore whether there is any variation in intensity as a marker of givenness, whether in terms of overall intensity or selective intensity (spectral tilt). No variation in either intensity measure was observed neither according to post-focal status nor the givenness status of the indirect object itself. The key finding of the study is therefore that intensity does not appear to vary with the givenness status of an object.

4 Discussion

In the quantitative study described above, we found that the pattern of variation in f0 excursion observed in the Hellmuth (2009) study in SVO sentences was reproduced here in VOO sentences, although the effect was somewhat smaller so that the differences observed did not reach significance. Crucially, we again found that variation in f0 excursion in the indirect object reflects its position (following or not following a focus) rather than its own

status (given vs. new). There was no accompanying effect on accented syllable duration, which also matches the findings of Hellmuth (2009).

The aim of the present paper was to rule out the possibility that some other acoustic correlate besides f0 excursion or duration reflects the givenness status of the indirect object itself in these data, and the most likely potential correlate for investigation was intensity (cf. van Heuven & Sluijter 1996). The results show however that intensity patterns (whether overall intensity at all frequencies, or selective intensity, that is, spectral tilt) prove to be similar to accented syllable duration patterns: there is no variation in intensity either, according to either post-focal position or givenness status.

How are we to interpret the lack of prosodic marking of givenness in EA? One possible explanation of the apparent lack of prosodic reflexes of givenness status could come from appeal to some form of complementarity in the division of labour between phonology and syntax in the expression of information structure categories. It has been noted that the primary marker of givenness in EA is in fact to delete the item altogether, which is possible because of the availability of subject 'pro-drop' and object cliticisation in the language (Jelinek 2002). It is perhaps therefore unsurprising to find no prosodic marker of givenness on arguments in the language (see discussion in Hellmuth 2010). A problem with this approach however is that it assumes a functional complementarity between phonology and syntax in the expression of IS in EA, which does not hold in other IS contexts such as prosodic vs. syntactic marking of thetic statements (Hellmuth 2010).

Instead, we explore here an interpretation of the lack of marking of givenness in EA within a prosody-syntax relationship mediated through metrical structure (Ladd 2008); in this approach IS distinctions such as ±focus or ±givenness are realised in terms of distinctions of (phonological) metrical prominence. The phonetic realisation of an IS category will in turn depend on how each level of metrical prominence is realised in a particular language.

The results of the experimental study presented here found no prosodic marking of the given vs. new distinction, but did match the findings of Hellmuth (2009) in that there was prosodic marking of focussed items in EA, though by manipulation of f0 excursion only with no accompanying dynamic cues such as duration or intensity.

This contrasts with the findings of Chahal (2001) for Lebanese Arabic, who found significant differences in mean values of f0, duration, intensity and F1/F2, distinguishing three levels of metrical prominence[11]:

(7) $\underset{①}{\begin{bmatrix}\text{+nuclear accent}\\ \text{+accent}\\ \text{+stress}\end{bmatrix}}$ vs. $\underset{②}{\begin{bmatrix}\text{-nuclear accent}\\ \text{+accent}\\ \text{+stress}\end{bmatrix}}$ vs. $\underset{③}{\begin{bmatrix}\text{-nuclear accent}\\ \text{-accent}\\ \text{+stress}\end{bmatrix}}$

As already noted in section 1, EA, unlike LA, is a variety in which every content word routinely bears a pitch accent, and in which post-focal items

[11] The correlates of the [+stress] vs. [-stress] distinction in LA are not known; for EA they have been shown to be marked by (at least) f0, duration and F1/F2 (Hellmuth 2006b).

are realised in compressed pitch range but are not fully de-accented. It is thus plausible to analyse the +focus items in the present dataset (realised with increased f0 excursion) as bearing the main prominence of the utterance; here increased f0 excursion in EA equates to bearing the nuclear accent in LA (despite the fact that the pitch accent realised on the focussed item in EA is not final in the utterance). The present dataset has thus established that EA distinguishes levels ①and ② prosodically, though by means of melodic cues only (and not dynamic also, as in LA). Cross-linguistic variation in use of melodic vs. dynamic + melodic cues has already been established at the lexical vs. postlexical levels (Ladd 2008:165).

Turning to givenness then, let us suppose that this IS category is marked by means of level ② vs. level ③ prominence. This seems plausible in English, since a post-focal given item will be de-accented but will still show dynamic correlates of word-level primary stress. Similarly, in Chahal's (2001) study of LA, the words classified as bearing level ③ prominence were indeed found in post-focal positions and were also given in context (due to prior mention in the eliciting wh-question). Since every content word routinely bears an accent in EA, levels ② and ③ can be said to be conflated, and thus prosodically indistinguishable. Under this view it is to be expected that none of the expected acoustic correlates of prominence (f0, intensity and duration) will be found to show effects of givenness, as has is the case in the present study.

Acknowledgements

Thanks to the Egyptian speakers who participated in the original recordings, to John Local, Barry Heselwood and two anonymous reviewers for comments and suggestions, and to Bert Remijsen for use of the selective intensity measurements Praat script.

Appendix

A multivariate Analysis of Variance (MANOVA) was carried out with all ten measures as dependent variables, and with Focus Set and Gender as fixed factors using SPSS 16.0. To compensate for the risk of Type I error, due to the fact that measures of each of the various dependent variables were taken from data from the same set of participants, a more restrictive significance level ($\alpha = 0.005$) is applied, using a Bonferroni adjustment (α/ number of tests: 0.05/10 = 0.005) (Rietveld & van Hout 2005:177). The MANOVA table is provided below, and relevant details reported in the text (in section 3); statistically significant results are shown shaded ($\alpha = 0.005$).

Multivariate Tests

Effect		Value	F	Hypothesis df	Error df	Sig.
mf	Pillai's Trace	.677	26.617	10.000	127.000	.000
	Wilks' Lambda	.323	26.617	10.000	127.000	.000
	Hotelling's Trace	2.096	26.617	10.000	127.000	.000
focus_set	Pillai's Trace	.200	.920	30.000	387.000	.591
	Wilks' Lambda	.812	.916	30.000	373.446	.596

mf * focus_set	Hotelling's Trace	.218	.913	30.000	377.000	.602
	Pillai's Trace	.277	1.312	30.000	387.000	.130
	Wilks' Lambda	.741	1.338	30.000	373.446	.114
	Hotelling's Trace	.325	1.363	30.000	377.000	.100
Design: Intercept + mf + focus_set + mf * focus_set						

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
mf	xxn	42.771	1	42.771	40.474	.000
	yxn	53.735	1	53.735	39.735	.000
	xdur	2105.012	1	2105.012	2.547	.113
	ydur	1000.929	1	1000.929	2.635	.107
	ymp	21.556	1	21.556	2.344	.128
	yip	1.255	1	1.255	.112	.738
	yav	6.971	1	6.971	1.002	.319
	idav	107.680	1	107.680	32.513	.000
	h1a3mp	2722.949	1	2722.949	126.297	.000
	h1a3ip	1710.324	1	1710.324	68.767	.000
focus_set	xxn	2.010	3	.670	.634	.594
	yxn	1.348	3	.449	.332	.802
	xdur	167.634	3	55.878	.068	.977
	ydur	2930.636	3	976.879	2.572	.057
	ymp	9.267	3	3.089	.336	.799
	yip	12.112	3	4.037	.361	.782
	yav	10.637	3	3.546	.509	.676
	idav	1.965	3	.655	.198	.898
	h1a3mp	70.492	3	23.497	1.090	.356
	h1a3ip	13.767	3	4.589	.185	.907
mf * focus_set	xxn	5.418	3	1.806	1.709	.168
	yxn	6.198	3	2.066	1.528	.210
	xdur	394.982	3	131.661	.159	.923

	ydur	1377.798	3	459.266	1.209	.309
	ymp	7.949	3	2.650	.288	.834
	yip	4.935	3	1.645	.147	.932
	yav	3.573	3	1.191	.171	.916
	idav	5.010	3	1.670	.504	.680
	h1a3mp	51.873	3	17.291	.802	.495
	h1a3ip	122.257	3	40.752	1.639	.183

References

Abdalla, A. G. 1960. *An instrumental study of the intonation of Egyptian Colloquial Arabic.* University of Michigan.

Beckman, Mary. 1986. *Stress and non-stress accent*. Dordrecht: Foris.

Boersma, Paul & Weenink, David. *Praat: doing phonetics by computer (Version 4.6.10) [http://www.praat.org]*. 2007.

Campbell, W. N. & Beckman, Mary. 1997. "Stress, prominence and spectral tilt.". *Intonation: theory, models and applications.* eds. A.Botinis, G.Kourooupetroglou, & G.Carayannis 67-70. Athens: Proceedings of an ESCA Workshop.

Chahal, Dana. 2001. *Modeling the intonation of Lebanese Arabic using the autosegmental-metrical framework: a comparison with English.* Unpublished PhD thesis, University of Melbourne.

Chahal, Dana & Hellmuth, Sam. to appear. "Comparing the Intonational Phonology of Lebanese and Egyptian Arabic." *Prosodic Typology. Volume 2.* Oxford: Oxford University Press.

Cruttenden, Alan. 2001. *Intonation*. Cambridge: Cambridge University Press.

Cruttenden, Alan. 2006. "The de-accenting of old information: a cognitive universal". *Pragmatic organization of discourse in the languages of Europe.* eds. G.Bernini & M.L.Schwartz 311-355. Berlin: Mouton de Gruyter.

de Jong, K. & Zawaydeh, B. A. 1999. "Stress, duration, and intonation in Arabic word-level prosody". *Journal of Phonetics* 27:1. 3-22.

Halliday, M. A. K. 1967. *Intonation and grammar in British English*. The Hague: Mouton.

Hellmuth, Sam. 2005. "No De-accenting in (or of) Phrases: Evidence from Arabic for cross-linguistic and cross-dialectal prosodic variation". *Prosodies (with special reference to Iberian languages).* eds. S.Frota, M.Vigario, & M.J.Freitas 99-121. Mouton de Gruyter.

Hellmuth, Sam. 2006a. "Focus-related pitch range manipulation (and peak alignment effects) in Egyptian Arabic.". *Proceedings of Speech Prosody 2006.* eds. R.Hoffmann & H.Mixdorff 410-413. Dresden: TUD Press Verlag der Wissenschaften GmbH.

Hellmuth, Sam. 2006b. *Intonational pitch accent distribution in Egyptian Arabic.* Unpublished PhD thesis, SOAS.

Hellmuth, Sam. 2007. "The relationship between prosodic structure and pitch accent distribution: evidence from Egyptian Arabic.". *The Linguistic Review* 24:2. 289-314.

Hellmuth, Sam. 2009. "The (absence of) prosodic reflexes of given/new information status in Egyptian Arabic.". *Information Structure in Spoken Arabic*. eds. J.Owens & A.Elgibali Oxford: Routledge.

Hellmuth, Sam. 2010. "Functional complementarity is only skin deep: evidence from Egyptian Arabic for the autonomy of syntax and phonology in the expression of focus.". *The sound patterns of syntax*. eds. N.Erteschik-Shir & L.Rochman Cambridge, M.A.: Oxford University Press.

Jelinek, E. 2002. "Agreement, clitics and focus in Egyptian Arabic.". *Themes in Arabic and Hebrew Syntax*. eds. J.Ouhalla & U.Shlonsky 71-105. Berlin: Springer.

Kiss, Katalin. 1998. "Identificational focus versus informational focus". *Language* 74:2. 245-273.

Krifka, Manfred. 2006. "Basic notions of information structure". *The notions of information structure*. eds. C.Féry, G.Fanselow, & M.Krifka Institutsverlag Universitaet Potsdam.

Ladd, D. Robert. 2008. *Intonational Phonology*. Cambridge: Cambridge University Press.

Mitchell, T F. 1993. *Pronouncing Arabic 2*. Oxford: Clarendon Press.

Nolan, F. 2003. "Intonational equivalence: an experimental evaluation of pitch scales.". *Proceedings of the 15th ICPhS*. 771-774.

Norlin, K. 1989. "A preliminary description of Cairo Arabic intonation of statements and questions". *Speech Transmission Quarterly Progress and Status Report* 1. 47-49.

Rastegar-El Zarka, Dina. 1997. *Prosodische Phonologie des Arabischen.* Unpublished PhD thesis, Karl-Franzens-Universität Graz.

Rietveld, Toni & van Hout, Roeland. 2005. *Statistics in Language Research: Analysis of Variance*. Berlin: Mouton de Gruyter.

Rifaat, Khalid. 1991. *The intonation of Arabic: an experimental study.* Unpublished PhD thesis, University of Alexandria.

Rooth, Mats. 1996. "Focus". *The Handbook of Contemporary Semantic Theory*. ed. by S.Lappin 271-297. Oxford: Blackwell.

Schwarzschild, R. 1999. "Givenness, avoid F, and other constraints on the placement of accents.". *Natural Language Semantics* 7. 141-177.

Selkirk, Elisabeth O. 2000. "The interaction of constraints on prosodic phrasing". *Prosody: theory and experiment*. ed. by M.Horne 231-262. Dordrecht: Kluwer.

Siemund, R, Heuft, B, Choukri, K, Emam, O, Maragoudakis, E, Tropf, H, Gedge, O, Shammass, S, Moreno, A, Nogueiras Rodriguez, A, Zitouni, I, & Iskra, D. *OrienTel - Arabic speech resources for the IT market*. 2002. LREC 2002 (Arabic workshop).

Sluijter, A. & van Heuven, V. 1995. "Effects of focus distribution, pitch accent and lexcial stress on the temporal organisation of syllables in Dutch.". *Phonetica: International Journal of Speech Science* 52. 71-89.

Sluijter, Agaath, Shattuck-Hufnagel, Stefanie, Stevens, Kenneth, & van Heuven, Vincent. 1995. "Supralaryngeal resonance and glottal pulse shape as correlates of stress and accent in English.". *Proceedings of the 13th International Congress of Phonetic Sciences, Stockholm* 2. 630-633.

Soraya, H. 1966. *An intonational study of Egyptian colloquial Arabic*. UCL.

Swerts, Marc. 2007. "Contrast and accent in Dutch and Romanian.". *Journal of Phonetics* 35. 380-397.

Swerts, Marc, Krahmer, Emiel, & Avesani, Cinzia. 2002. "Prosodic marking of information status in Dutch and Italian: a comparative analysis". *Journal of Phonetics* 30:4. 629-654.

van Heuven, V. & Sluijter, A. 1996. "Notes on the phonetics of word prosody.". *Stress patterns of the world. Part 1: Background.* eds. R.Goedmans, H.van der Hulst, & E.Visch 233-269. The Hague: Holland Academic Graphics.

Vanhove, Martine. *Focalisation and deontic values in Beja. (Paper presented at the International Berlin Focus Conference 6-8th October 2005, ZAS, Berlin).* 2005.