UNIVERSITY *of* York

This is a repository copy of *Investigating variation in Arabic intonation : : the case for a multi-level corpus approach*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/102522/

Version: Accepted Version

Book Section:

Hellmuth, Sam orcid.org/0000-0002-0062-904X (2014) Investigating variation in Arabic intonation : : the case for a multi-level corpus approach. In: Farwaneh, Samira and Ouali, Hamid, (eds.) Perspectives on Arabic Linguistics XX1V-XXV. Studies in Arabic Linguistics . John Benjamins , pp. 63-89.

https://doi.org/10.1075/sal.1.06hel

White Rose
university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Chapter title	Investigating variation in Arabic intonation: the case for a multi-level corpus approach
Running head	Investigating variation in Arabic intonation
Author	Sam Hellmuth
Affiliation	University of York, UK
Address	Department of Language and Linguistic Science
	University of York
	Heslington, York YO10 5DD, UK
email	sam.hellmuth@york.ac.uk

Investigating variation in Arabic intonation: the case for a multi-level corpus approach

Sam Hellmuth, University of York, UK

This paper provides a first description of the intonational patterns of San'aani Arabic (SA, the dialect of Arabic spoken in the capital of Yemen) and a comparison of these patterns with those observed in Cairene Arabic (CA), revealing differences between the two varieties which mirror cross-linguistic prosodic variation. The SA analysis is based on qualitative transcription of portions of a multi-level corpus, including read speech sentences, a narrative retold from memory and a sociolinguistic data collection tool which yields free conversation data in the desired variety as well as information that can be used to confirm which variety is being used. The corpus design and methodology serve as a prototype for larger data collection to document intonational variation in Arabic.

1   Introduction

Spoken Arabic dialects vary from each other in a range of linguistically interesting ways, both in morphosyntax and in their phonetics and phonology. Analysis of the fine-grained differences observed between dialects - when examined in the context of the properties that are shared -

2

has been instrumental in the advancement of linguistic theory in a number of areas, not least in the field of metrical phonology (cf. Broselow, 1992; Watson, 2011).

The aim of this paper is to demonstrate that variation in the intonation patterns of spoken Arabic dialects is likely to provide similarly rich input to theories of intonational phonology, and to make a case for a particular methodological approach to the analysis of intonation, namely use of a multi-level corpus comprising a range of speech styles. This is achieved by presenting a first analysis of the intonation patterns of Sanaani Arabic (SA), based on a pilot set of multi-level corpus data elicitation tools. The resulting data is analysed within the framework of Autosegmental-Metrical (AM) theory (Gussenhoven, 2004; Ladd, 2008) which allows data from languages of different prosodic types to be directly compared. In this paper, the findings for SA are compared with published descriptions of Cairene Arabic (CA), and reveal that the two dialects differ in their intonational phonology, along known parameters of cross-linguistic structural variation in intonation (Jun, 2005).

The outline of the paper is as follows: section 2 sets out the key tenets of AM theory and how prosodic typology is expressed in the theory, summarises previous work on Arabic intonation and discusses the methodological approaches used in that work alongside approaches used to

investigate intonational variation in other languages; section 3 describes the data instruments used to elicit the multi-level corpus of SA analysed in this paper, and the essential properties of the intonational transcription system adopted during analysis; section 4 presents the inventory of pitch accents and boundary tones observed in SA, in different contexts, and demonstrates how the different types of speech style elicited in the corpus contribute to building up a full picture of the intonational phonology of the language; in section 5 the intonational phonology of SA is briefly compared with that of CA, and the paper closes by arguing for the need for parallel multi-level data collection in a range of spoken Arabic dialects.

2   Background

2.1   Formal analysis of intonational phonology

A number of models have been proposed to account for the patterns observed in the intonation of human languages. These models aim to provide either a functional account of intonation (how speakers use intonational patterns to achieve communicative goals), or a formal account (what is the range of possible intonational patterns in human languages), or both. This paper pursues a primarily formal account, working within the AM framework for analysis of intonation (Gussenhoven, 2004; Ladd, 2008), which developed out of pioneering work on the intonation of American English (Pierrehumbert, 1980), itself influenced by work on tonal patterns in

Swedish (Bruce, 1977). Analysis of the form of intonational patterns cannot be entirely divorced from their use and function, but any arguments made in the present paper about function will be essentially distributional in nature (Gussenhoven, 2007); that is, based on the observed occurrences of particular phonological patterns in particular semantic or pragmatic situations. As a result, a statement in this paper to the effect that a tonal pattern is frequently observed in a certain situation or sentence type, cannot – and should not – be interpreted as a prediction that the tonal pattern in question will always occur in that context, or that that context will always bear that tonal pattern. Nonetheless, it is helpful in the early stages of development of an intonational analysis, particularly for languages whose intonation patterns have not previously been described, to document observed regularities of sound-meaning co-occurrence.

Regarding the *form* of intonation, a fundamental claim of AM theory is that a continuous intonation contour is best modelled in terms of a linear sequence of level tones or targets, rather than as contours, such as falls, rises or rise-falls (see Ladd 2008 chapters 1-2 for arguments surrounding this approach). AM proposes two types of phonologically relevant tonal event – pitch accents and boundary tones – each of which may comprise a single tone (L, low or H, high) or more than one tone (e.g. LH or HL). Pitch accents are tonal events that are phonologically constrained to be realised on or near a metrical *head* (usually the primary stressed syllable within a word

or phrase), whereas boundary tones are phonologically constrained to be realised at or near the *edge* of some metrical domain (such as an intonational phrase). Tonal events thus mark the heads and/or edges of prosodic constituents. A hierarchy of prosodic constituents of increasing size is generally assumed, from the syllable upwards, through the Prosodic Word 'PWd', Phonological Phrase or Intermediate Phrase 'iP' and Intonational Phrase 'IP' (Selkirk, 2011). A distinction is commonly made between the nuclear accent, which in English is the rightmost pitch accent in an utterance and thus bears its main prominence, and pre-nuclear pitch accents occurring earlier in the utterance.[1] A distinction is observed in some languages in the inventory of accents observed in nuclear vs. pre-nuclear positions.

Cross-linguistic variation in intonation is modelled in AM theory in terms of differences in i) the inventory of possible types of pitch accents and boundary tones, and ii) the range of permitted associations of tonal events to metrical structure. Just as the phonotactics of a language dictate which features or segments may occur in which positions of the syllable or word, the intonational phonotactics of a language dictate what tonal events may occur in which positions of metrical structure. For example, English requires the head of every phonological phrase to be marked with a pitch accent (Selkirk, 2000), whereas the distribution of tonal events is rather different in other languages, which may require the head of every constituent at some other level of the prosodic hierarchy to bear a pitch

accent, such as the Prosodic Word (Hellmuth, 2007). Similarly, languages vary in which constituents of the prosodic hierarchy must be tonally marked. Languages may even vary in whether both heads and edges of constituents are tonally marked (as in English) or only the edges (as in French or Korean). To date no language has been reported which marks only prosodic heads, though it is to be noted that the number of intonation languages for which a detailed intonational description exists is as yet very limited, compared to, say, the number of tone languages which have been described and formally analysed (Beckman & Venditti, 2011).

2.2   *Arabic intonation*

Linguistic use of pitch is standardly analysed as falling somewhere along a continuum from tone languages, in which tone is fully lexical, to intonation languages, in which tone is fully postlexical (Hyman, 2006). Even though this oversimplifies things somewhat in the case of tone languages (which generally use pitch both lexically and postlexically, Beckman & Venditti, 2011), the consensus is that virtually all varieties of Arabic use pitch for purely postlexical purposes; that is, Arabic, in all its many variants, is an intonation language.[2]

Prior literature on the intonation of spoken Arabic dialects is still relatively limited, but in recent years a growing number of descriptions of individual

dialects have been published. For some varieties there are a range of competing AM analyses (e.g. Cairene: Rastegar-El Zarka, 1997; Rifaat, 2005; Hellmuth, 2006; El Zarka, 2011), but for others there is only a preliminary description (e.g. Syrian: Kulk, Odé, & Woidich, 2003; or Emirati: Blodgett, Owens, & Rockwood, 2007), or a description in a very different theoretical framework (e.g. Moroccan: Benkirane, 1998), or no description at all (e.g. Iraqi). A good summary and overview of the literature is provided by Chahal (2009).

Very few studies have made *direct* comparison of the intonational phonology of different Arabic dialects. A notable exception is Ghazali et al. (2007), which analysed a very small amount of data in six dialects.[3] Even the preliminary findings of Ghazali et al (2007) indicate, however, that there is variation across Arabic dialects along parameters of prosodic variation which have been observed to hold cross-linguistically. This includes variation in which level of the prosodic hierarchy displays obligatory association of pitch accents to its metrical head, or whether both heads and edges or just edges of prosodic constituents are tonally marked (see Hellmuth 2013 for discussion). Further work on Arabic intonation is thus likely to yield fruitful results, which will in turn provide insights into the as yet largely unexplored interface of intonation with syntax and semantics in Arabic.[4]

Cross-dialectal work on Arabic intonation is an important goal, therefore, and will be most effective if analyses of different dialects rely on shared (or at least explicit) theoretical assumptions, and are based on parallel data. This paper argues for a particular methodology for use in cross-varietal work on Arabic intonation: a multi-level corpus approach. The present case study is developed within the AM framework, but the case for working with multi-level corpora to develop an intonational analysis is independent of arguments for or against the choice of theoretical framework used to analyse the resulting data.

2.3   Data collection for intonational analysis

The studies of Arabic intonation surveyed briefly in the above section have been based on a variety of different types of data, ranging from fully controlled read speech sentences to recordings of spontaneous free conversation. Controlled read-speech data are an essential starting point in any study of intonation, and some preliminary studies report results from this type of data only (Norlin, 1989; Kulk et al., 2003; Ghazali, Hamidi, & Knis, 2007). An alternative approach, at the other end of the continuum, is to extract utterances of a certain type (questions, statements etc) from recordings of free conversations or monologues (Rifaat, 2005; Blodgett et al., 2007). A number of studies work from analysis of data from both of these two types (Kulk et al., 2003; El Zarka, 2011).

The current state of the art in data collection for investigation of cross-varietal intonation patterns is to use a multi-level corpus, which includes read speech and spontaneous conversation alongside a number of intermediate types of elicited speech. These intermediate types may include dialogue completion tasks, to elicit semi-spontaneous utterances in a variety of focus contexts (Skopeteas et al., 2006; Prieto & Roseano, 2010), map tasks (Anderson et al., 1991) and narratives read and/or re-told from memory (Grabe, 2004). For Arabic, some studies have moved in this direction and are based on data elicited from a mix of tasks (Chahal, 2001; Hellmuth, 2006), but no work to date on Arabic has benefited from availability of a fully multi-layered corpus, as is now available for a number of other languages, such as British English, Spanish and Portuguese.[5]

An additional issue in data collection, specific to Arabic, is which register of the language is to be elicited, and how. In general, studies on Arabic intonation have focussed either on the intonation of the formal variety (Modern Standard Arabic) when produced by speakers of a particular dialectal speech community (Rastegar-El Zarka, 1997; Rifaat, 2005), or directly on the intonation of a colloquial variety (Chahal, 2001; Hellmuth, 2006; Blodgett et al., 2007). Some studies combine analysis of data from both registers (El Zarka, 2011), a justifiable approach given that few differences have been observed to date between the intonational patterns of

10

speakers when their formal and colloquial speech is compared (El Zarka & Hellmuth, 2009). Reliable elicitation of controlled speech in the colloquial register is not an easy task (Siemund et al., 2002), and this in itself motivates multi-layered data elicitation techniques for work on Arabic.

In addition, all studies of Arabic intonation necessarily present analysis of data from a small sample of speakers to represent a larger speech community, such as that of a particular city (e.g. Cairo). Sociolinguistic studies confirm however - as is clear even to the casual observer - that within a single urban centre in the Middle East a wide range of socially determined varieties co-exist, with speakers able to function effectively in a number of varieties alongside their own 'mother tongue' dialect (Bassiouney, 2009). It is useful, therefore, to have some indication of what a speaker's mother tongue dialect actually is, and thus whether the speech data elicited is spoken as a first, second, or perhaps third dialect (if such clear distinctions can even be made). These issues are not exclusive to Arabic, and the present paper presents the results of first use - for intonational analysis purposes - of a sociolinguistic tool originally designed for work in regional British English dialects, the Sense Relation Network (Llamas, 2007). This tool is designed to elicit free conversation in speakers' first dialect, while also providing a record of local vocabulary items used by the speaker, as an aid to classification of the dialect under examination.

Finally, adaptation for use in Arabic of any data elicitation materials used for other languages must also take account of the segmental and metrical phonology of the language. Intonational pitch accents are associated with stressed syllables (see 2.1), and it is well-known that factors in the prosodic environment will affect the fine-grained realisation of pitch accents; these factors include the position of stress within a word (Prieto, van Santen, & Hirschberg, 1995), and the type of syllable (open or closed, heavy or light) which bears stress (Post, D'Imperio, & Gussenhoven, 2007). Prior work confirms that these prosodic environments affect the realisation of pitch accents in spoken Arabic dialects (position of accented syllable in the word affects peak alignment in Lebanese Arabic, Chahal, 2003; accented syllable type affects peak alignment in Egyptian Arabic, Hellmuth, 2007). In Arabic, the position of stress in the word co-varies to a large extent with the syllable type bearing stress. For example, in most dialects word-final stress is only ever realised on a 'super-heavy' CVVC or CVCC syllable (van der Hulst & Hellmuth, 2010); as a result, controlling for position of stress in target words also, to some extent, controls syllable type. If the effects on the fine-grained realisation of pitch accents of prosodic environment vs. utterance meaning are to be disambiguated, both factors must be controlled, in at least some portion of the data on which an analysis is based. In contrast to the corpus instruments which have been used for work on English and Spanish therefore, which do not control for position of stress in the word, in the

12

present corpus, accent position in the last lexical item is controlled in the read speech sentences.

In the sections of the paper that follow, section 3 presents the details of the multi-layered corpus instruments used in a pilot study to elicit data in Sanaani Arabic, together with the transcription system used to frame the analysis in this paper; section 4 shows, firstly, how the multi-layered corpus data can be used to identify the dialect under investigation, and, secondly, how the different levels of the data combine to facilitate proposal of a complete analysis.

3 Methods

3.1 Data collection

This section outlines a set of stimuli and tasks designed to collect a multi-level corpus of Arabic data. The data were collected in SA in a pilot study, to generate a first analysis of SA intonation, and to evaluate the effectiveness of the tools and determine whether additional tools are needed. The pilot study tools used in the present study are set out in Table 1.

	tool:	*yields:*
1.	read speech sentences	Lab speech realisations of different utterance types, parallel utterances produced by each speaker: 8 x broad focus declaratives (dec), 6 x wh-questions (whq), 6 x yes-no-questions (ynq), 6 x coordinated questions (coo) [N=26]. The position of the accented syllable in the last lexical item is systematically varied: final, penult, antepenult.

2.	narratives, read/retold	The read narrative yields data in which different speakers of the same dialect all produce the same sentences, in a narrative sequence. The retold narrative yields some identical or near-identical sentences, produced semi-spontaneously by different speakers of the same dialect.
3.	map task	The map task yields semi-spontaneous realisations of different utterance types; mismatches in map naturally generate questions. The names of landmarks on map contain mostly sonorant speech sounds.
4.	SRN[6]	The SRN yields naturally occurring conversation in dialect, and also facilitates dialect identification, by choice of lexical item.

Table 1 Pilot study multi-level corpus data elicitation tools

Items 1-3 in Table 1 are modelled on the data collection techniques used in the IViE project (Grabe, 2004). Development of an analysis of the intonational phonology of a language or dialect, whose prosody has not previously been described, is greatly facilitated by having multiple instances of the same set of utterances produced by different speakers. An innovation in these elicitation tools is systematic variation of the position of the accented syllable in the last lexical item in the read speech sentences, for the reasons set out in 2.3. In all, 26 isolated read speech utterances were elicited from each speaker, in an interactive role-play setting. Some sample sentences are provided in (1), illustrated in Figures 1-3 in the Appendix. Note that the syntactic form of the declarative and yes-no question is identical in this example, with the utterance-level semantic distinction between them realised by prosodic means alone (cf. Watson, 1993: 395).

(1) *dec* ˈhaːðaː r-raˈdʒaːl ˈjamani
           this   the-man   Yemeni
           'This man is Yemeni.'

     ynq ˈhaːðaː r-raˈdʒaːl ˈjamani
           this   the-man   Yemeni
           'Is this man Yemeni?'

whq	maː	ism	r-raˈdʒːaːl	l-ˈjamani		
	what	name	the-man	Yemeni		
	'What is the Yemeni man's name?'					
coo	ˈhaːðaː	r-raˈdʒːaːl	huː	ˈʔurduni	ˈwalla	ˈjamani
	this	the-man	he	Jordanian	or	Yemeni
	'Is this man Jordanian or Yemeni?'					

For the narrative task a Guha folk tale in Egyptian dialect (Abdel-Massih, 1975) was played as a prompt, and the participant then retells the story in SA from memory. The source story is 300 words in length and the re-telling analysed for this paper contains 283 words. The SRN task is modelled on Llamas (2007). The target lexical items were a set of words expected to vary across dialects spoken within Yemen. For example, the test sheet provided the MSA stimulus [ˈmɑtˤbɑx] 'kitchen', and participants were invited to provide the parallel word used in their dialect, e.g. [ˈdajmih], in SA (Watson, 1996). In the present study the SRN used MSA target words typed in Arabic script, laid out in a network format (see Llamas 2007). Use of MSA triggered some use of the spoken formal register of Arabic at the beginning of the recordings, but the tool was successful in encouraging participants to quickly settle into using their colloquial dialect (cf. Hellmuth, 2014). The SRN data analysed for this paper is the first 7 minutes of a 20 minute interview with two speakers. Examples of the retold narrative and SRN data are provided in Figures 4-6 in the Appendix.

Speech recordings were made by the author with 12 participants, during fieldwork in Sanaa, Yemen. Recruitment was initially through private educational institutes, then on a friends-of-friends basis. Digital recordings were made in quiet classrooms or living rooms, directly to .wav format at 44.1KHz 16bit, using a Marantz PMD660 and head-mounted Shure SM10 microphones. Conversational data was recorded in stereo, one speaker per channel, permitting analysis either of each speaker individually or of both.

Examination of the data revealed just three of the participants to be speakers of SA, the variety described in the work of Janet Watson (Watson, 1993; Watson, 2002; Watson & Asiri, 2009), rather than of other Yemeni dialects (see 4.1). The results of analysis of data from three speakers are thus reported, one male and two female: m1, f1 and f2. All were born and raised either in the Old City itself, or in the Al-Ga' district just west of the Old City (further background information is provided in Table 8 in the Appendix). None report any hearing difficulties. This paper presents the results of analysis of: 69 read speech utterances (produced by all three speakers), a narrative retold by f1, and the speech of speaker f2 in the SRN.

The design and testing of this stimulus set was carried out as a pilot study, and this paper forms part of the evaluation of the relative success of these tools. The work reported in this paper has informed the design of an

expanded stimuli set for use in data collection for intonational analysis purposes in a wider range of Arabic dialects.[7]

3.2   Transcription

The Tones and Break Indices (ToBI) labelling system is not an 'IPA for intonation' (Beckman, Hirschberg, & Shattuck-Hufnagel, 2005). Instead the analyst must determine what subset of labels is needed to describe the range of intonational patterns observed in a particular language or dialect. This equates more or less to the task of determining what the inventory of pitch accents and/or boundary tones is in the language. What is shared across ToBI-style labelling systems for different languages - at least in principle - is the set of assumptions about what the set of cross-linguistically *possible* labels is, and how those labels are to be defined.

Although most ToBI-style transcription systems do indeed adopt the key tenets of AM theory (described briefly in 2.1 above), analysts have varied greatly in how patterns in the f0 contour should be interpreted in terms of pitch accents. For example, in a bitonal accent, should the '*' be assigned to the most salient tone (Prieto, D'Imperio, & Gili Fivela, 2005), or to the one which best describes the shape of the pitch during the accented syllable, low or high (Prieto & Roseano, 2010)? Resolving these difficulties is beyond the scope of the present paper, but it is nonetheless vital for analysts to provide

a clear statement of the conventions on which a transcription is based, if the resulting analysis is to be interpretable (cf. International Phonetic Association, 1999, p. 29). The remainder of this section attempts to provide such a statement for the analysis which follows in section 4.

The data described in section 3.1 were transcribed using a set of 'ToBI-style' labels, and adopted, as a starting hypothesis, the definitions of how pitch accents map to f0 contour shapes that form the current consensus for labelling of intonation in work on Romance languages (cf. Prieto & Roseano, 2010; Prieto & Frotato appear). The stylised descriptions in Table 2 represent how each potential accent label is expected to be realised on a word in which the accented syllable is a heavy syllable (CVV or CVC) and is in penult position in the word. In Arabic this would be how the accent is realised on a word such as [lub.'naa.ni] 'Lebanese'. In this system the starred tone indicates the pitch contour shape during the accented syllable.

Trancription was carried out based on auditory impression, with reference to the waveform, spectrogram and pitch trace in Praat (Boersma & Weenink, 2013). It was assumed during transcription that the realisation of pitch accents would vary according to position of the accent in the utterance (nuclear/pre-nuclear), position of the accented syllable within the accented word (final/penult/antepenult) and type of accented syllable (heavy/light). Accents were analysed as being the 'same' pitch accent (or not) in different

places based on impressionistic analysis (that is, by ear), rather than purely based on phonetic realisation (due to expected variation in realisation due to prosodic context, see 2.3). The resulting auditory analysis was also checked for distributional consistency against the typical contexts in which that pitch accent was observed elsewhere in the data (Gussenhoven, 2007). Table 7 in the Appendix provides a summary of the observed variation in surface realisation of the phonological pitch accent categories proposed for SA.
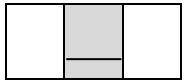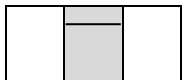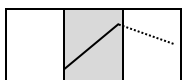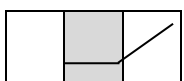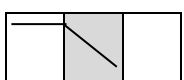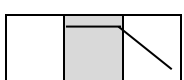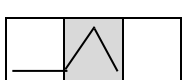
	σ^1	$'\sigma^2$	σ^3	*shape of f0 contour during accented syllable*
L*				A low plateau during the accented syllable.
H*				A high plateau during the accented syllable, no preceding f0 valley.
L+H*				A rising pitch accent during the accented syllable, with the f0 peak at the end of the accented syllable.
L*+H				A low plateau during the accented syllable with a subsequent rise in the postaccentual syllable.
H+L*				An f0 fall during the accented syllable.
H*+L				A high plateau during the accented syllable, followed by an f0 fall.
LH*L				A complex rise-fall within the accented syllable.

Table 2: Schematic representation of sample possible pitch accents using ToBI-style labels. The box representation shows a three syllable word, with the middle syllable shaded to indicate that it is the accented syllable.

The 69 read speech sentences were 'double-blind' transcribed by two transcribers in the first instance (the author and a second trained linguist who is a native speaker of Arabic). Inter-transcriber agreement in this initial analysis was 44%; the transcriptions were compared and used to generate a first working hypothesis as to the set of labels needed for transcription of SA. This label set was then used by the author to transcribe the retold narrative data (2min16s, 283 words), with additional use of an *alternative* tier for ambiguous cases (Brugos, Veilleux, Breen, & Shattuck-Hufnagel, 2008); and the resulting transcription was then checked for consistency with the proposed description of the model/labels, by the second transcriber. Inter-transcriber agreement in this second analysis was 80.7%, and rose to 93% when transcriptions on the alternatives tier were included. Inter-transcriber agreement for the SRN data is reported in Hellmuth (2014).

Variation in how pitch accents are realised when the accented syllable occurs in different positions of the word is discussed in section 4.2 below. Discussion of what subset of ToBI-style labels is needed to transcribe SA, in narrative and conversational data, as opposed to read speech data, follows in section 4.3. Before that, section 4.1 sets out the evidence in the dataset which identifies the dialect elicited as SA.

4   Results

4.1   Identification of variety used by speakers in the corpus

All of the speakers show marking of pausal forms using glottalisation and devoicing. Use of these laryngeal features at pause has been argued to be an areal feature found cross-linguistically in southern Arabia (Watson & Asiri, 2009; Watson & Bellem, 2011). Instances of these phenomena observed in the present data include those listed in 2).[8]

			target	*gloss*	*realised as*	*example*
(2)	a.	devoicing/glottalisation of obstruents	/ʒaˈdiːd/	'new'	[ʒaˈdiːd̥]	Fig. 6
	b.	glottalisation of long vowels/sonorants	/ˈjamaniː/	'Yemeni'	[jamaˈniːʔ][9]	Fig. 2
	c.	glottalisation/deletion of final nasal	/xamˈsiːn/	'fifty'	[xamˈsiːʔ]	Fig. 3

In order to classify speakers according to dialect variation within Yemen, there are salient features of different dialect groupings which are easy to identify. For example, all three speakers (f1, f2, m1) consistently use a [g] realisation of 'qaaf' <ق> and a [dʒ] realisation of 'jiim' <ج>. By using these features it was possible to differentiate these three speakers from other speakers in the corpus, whose families originate from beyond greater Sanaa, e.g. Ta'izz or Ibb, and who generally use [q] for 'qaf' and [g] for 'jiim'.

The SRN tool permits further fine-grained identification of the dialect spoken in the data. The vocabulary items preferred by the three speakers show agreement in many lexical items, independently identifiable as Sanaani (Watson, 1996), as shown in Table 3. In contrast, their choices

21

differ from those made by other speakers in the corpus (e.g. f4, who is from

Ba'dan, near Ibb), as shown in Table 4.[10]

children (s./pl.)	ja:hil ~ jihha:l	*give*	jiddi:	*above*	tˤɑ:luʕ
man (s./pl.)	radʒ:a:l~ ridʒa:l	*take*	jidʒirr ~ jibizz	*under*	na:zil
guest room	di:wa:n	*beautiful, nice*	ħa:li	*belongs to*	ħagg

Table 3: Vocabulary choices in which f1/f2/m1 agree

	f1/f2/m1	*f4*
kitchen	dajmih	ʃuhidih
sit	jigambir ~ jigawħiz	jugʕud
pregnant	wa:hima	mirzin

Table 4: Vocabulary choices where Sanaani f1/f2/m1 differ from Ba'dani f4

There is some variation among the group of three Sanaani speakers in their

vocabulary choices, in a small number of words. In some cases f2 uses a

different word from both f1 and m1, suggesting that, although she lives in

Al-Ga' (as f1 does), there may be some effect of her family originating from

a village outside Sanaa (Hamdan), rather than Sanaa itself. In a small

number of cases m1 diverges from both f1 and f2, suggesting that there may

be vocabulary differences between Old City Sanaani and the variety spoken

in Al-Ga', or between male and female speakers (cf. Naïm-Sambar, 1994).

	f1/m1 (parents from Sanaa)	*f2 (parents from Hamdan)*		*f1/f2 (live in Al Ga')*	*m1 (lives in Old City)*
sit	jigambir	jigawħiz	same	ʃibh	sa:ʕ ma:

Table 5: Variation in vocabulary choices among the three Sanaani speakers

In light of the above lexical choice data, the dialect under examination in the present study is identified as Sanaani Arabic.

4.2   Patterns observed in read speech

This section describes the inventory of pitch accents and boundary tones observed in the read speech sentences portion of the corpus, in pre-nuclear and nuclear positions. The varying phonetic realisation of these pitch accents, according to position of the accented syllable in the accented word, is also summarised, in order to flesh out the conventions of what each proposed phonological pitch accent maps to in surface realisation.

In pre-nuclear positions, in 'broad focus' read speech sentences of all types, both declarative and interrogative, one pitch accent type is predominantly observed. This typical pre-nuclear pitch accent is a rising accent, transcribed as L*+H, due to the fact that it is characterised by low level pitch during the accented syllable (L*), followed by a high turning point (+H). The trailing H peak of the L*+H accent is realised outside (after) the accented syllable if the accented syllable is in penult position. The L*+H accent may be followed by a H- phrase tone, or may occur phrase-internally. In phrase-internal cases there is a clear rise but no other obvious cues to phrase juncture (such as lengthening, pause or local pitch reset), suggesting that

analysis of the H peak as a trailing tone, that is, as part of the L*+H pitch accent, is appropriate. Two other pitch accents are observed in pre-nuclear position, in a handful of cases. Pre-nuclear L* accents were observed, after an early main prominence, in six tokens of wh-questions. A pre-nuclear H* was observed in three tokens, on words preceded by the particle [gad] (a particle denoting past aspect).

In nuclear position the read speech sentences allow a distinction to be established between two pitch accents. In declarative read speech sentences elicited in broad focus condition the most common nuclear pitch accent used is H*. In nuclear position the H* is characterised by a high plateau throughout the accented syllable in words with penult stress (e.g. [sˤanˈʕaːni] 'Sanaani', see Fig 1). In many cases the nuclear H* is realised somewhat downstepped in relation to preceding high turning points, in a typical 'final lowering' effect (Liberman & Pierrehumbert, 1984). This downstepped H* is transcribed as !H* and is argued to be a variant of H*, rather than a separate phonological category. This is because final lowering appears to be largely positionally determined, being observed only phrase-finally, and after a sequence of accents.[11]

In polar questions (ynq) an L+H* pitch accent is typically observed in nuclear position. The H peak is realised at the end of the accented syllable

on a word with penult stress. This is an earlier peak than in the typical pre-

nuclear L*+H accent. One could argue that the earlier peak is due to the

upcoming phrase boundary (cf. Prieto et al., 1995; Chahal, 2003), and thus

that the nuclear accent observed in polar questions is an allophonic variant

of prenuclear L*+H. Nonetheless an allophonic interpretation was rejected

at this stage of the analysis, in favour of a distinct nuclear L+H* accent, due

to the systematic appearance of the L+H* nuclear accent in a particular

context (polar questions). As will be seen in 4.3, this parallels the early peak

L+H* pre-nuclear accent observed in some focal contexts in narrative and

conversational data. For these sentence-initial focal accents there is no

option but to propose a separate pitch accent (L+H*), since they are not

close to a phrase boundary and are observed in a context with a parallel

semantic interpretation (invoking alternatives, see 5.2 below). A summary

of the pitch accents observed in the read speech sentences is provided in (3).

(3)   *where observed in read speech portion of corpus:*

pre-nuclear	L*+H	in most broad focus declarative sentences
	H*	only following [gad] in read speech (N=3)
	L*	only after early main prominence in wh-questions
nuclear	H*/!H*	in most broad focus declarative sentences
	L+H*	in yes-no questions
	L*+H	in utterances with a rising phrase-final boundary
	L*	after early main prominence in wh-questions

Turning to boundary tones, in the read speech sentences, nuclear H* is

always followed by a low boundary tone 'L%' (Fig 1-3). The L+H* nuclear

tone observed in polar questions is also always followed by a low boundary tone 'L%' (Fig 2). All cases of nuclear L*+H are followed by a high boundary tone 'H%' (Fig 3).

The position of the accented syllable was systematically varied in the last lexical item in the read speech sentences, making it possible to document observed variation in the realisation of pitch accents, in pre-nuclear and nuclear position, by accented syllable position. The patterns are summarised in schematised representation in Table 7 in the Appendix.

The trailing H peak of the most common pre-nuclear accent, L*+H, is realised outside an accented syllable in penult position, but inside, and towards the end of, an accented syllable in word-final position. For L*+H in nuclear position followed by a high boundary the position of the H peak is indeterminate.[12] The most common nuclear accent, H*, is characterised by a high turning point late in the accented syllable in words with antepenultimate stress, a high plateau throughout the accented syllable in words with penult stress, and a high turning point at the beginning of the accented syllable in words with final stress. For H* in pre-nuclear position, the peak is realised at the beginning of the accented syllable in words with antepenult stress, and towards the middle of the accented syllable in words with penult stress. In an L+H* nuclear accent produced on a word with final stress the peak is realised one-third to halfway through the accented

syllable. If the word has penult stress the peak is realised at the end of the accented syllable.

Based on analysis of short read speech utterances, an inventory would have to be proposed which comprises only those pitch accents in SA listed in (3), and which displays a distributional asymmetry between pre-nuclear and nuclear accents. The next section explores whether this inventory accurately represents the intonational patterns observed in longer, more spontaneous, stretches of SA speech.

4.3   Patterns observed in narratives and conversation

This section describes the inventory of pitch accents and boundary tones observed in analysis of the retold narrative and SRN. This semi-spontaneous and spontaneous data provides a more varied range of information structure and interactional contexts. The *additional* pitch accents required for transcription of this data are listed in (4) below, together with the situations in which they are observed (or, for accents already in the inventory, the newly observed contexts in which they occur).

(4)   *pre-nuclear*   H*   observed in 'flat hat' contours
                        L+H*   on early main prominence in focus contexts
      marginal cases LH*L   on early main prominence in focus contexts
                        H+H*   in 'speculative' contexts
      boundaries   L-   intermediate level of juncture, low tone

H-   intermediate level of juncture, high tone

The enlarged dataset confirms the status of L*+H as the most common pre-nuclear accent and of H* as the most common nuclear accent. The typical declarative pattern used in both narrative and conversation data is a rising falling pattern over the whole phrase (Rifaat, 2005), analysed in this paper as L*+H H* L-L%. The place of L+H* in the inventory is confirmed, with observation of its use in pre-nuclear position in contexts which appear to denote a contrastive topic or focus, alongside further examples of L+H* in nuclear position in polar questions (as previously observed in read speech).

There are seven instances in which the following fall after a L+H* is aligned tightly at the end of the accented syllable, rather than at the end of the word (where a boundary tone might plausibly be expected), for which an LH*L accent is tentatively proposed (cf.Prieto & Roseano, 2010). Similarly, an H+H* accent is proposed to account for four tokens in the SRN conversation with high pitch on the pre-accentual syllable; the contexts share the property of being lexical items which the speaker is not fully committed to as being 'Sanaani'.[13] The status of these marginal accents in the SA tonal inventory requires further investigation.

The longer stretches of speech also require proposal of an intermediate degree of juncture (cf. Prieto & Roseano, 2010, p. 4), yielding a model

which has both iP-final phrase tones and IP-final boundary tones.[14] Pitch range reset is used as the primary indicator of the start of a new IP; where the pitch range reset at the start of a new chunk of talk is at a lower level relative to that observed at the beginning of the previous chunk, this is annotated as an iP level juncture. Prieto & Roseano (2010) argue against labelling of iP-final phrase tones in IP-final position, despite their use to mark intermediate junctures, because in Spanish the nucleus obligatorily occurs on the last lexical item in the utterance, thus no stretches of unaccented material occur after the nucleus. This is not the case for SA, which allows non-final nuclei. Complex boundary tones are therefore proposed for SA.[15] The presence, or absence, of any mapping between these levels of prosodic phrasing and syntactic structure, remains a topic for future research.

Observed variation in the surface realisation of the core set of proposed pitch accents, by position in the utterance (pre-nuclear and nuclear) and by position of the accented syllable in the word, are summarised in schematised representation in Table 7 in the Appendix.

5 Discussion

5.1 Sanaani Arabic intonation

Based on a multi-level corpus of data from three speakers, a first model of

SA intonation is proposed, as summarised in the inventory of pitch accents

and boundary tones provided in Table 6.

pre-nuclear	nuclear	boundaries
L*+H	H*	L-L% [falling]
H*	L*	H-H% [rising]
L*	L+H*	H-L% [mid-level]
L+H*	L*+H	L-H% [fall-rise]
? LH*L		
? H+H*		

Table 6 Proposed model of Sanaani Arabic intonation

An inventory based on analysis of the read speech sentences alone (as in

4.2) would miss the symmetry of the core inventory, with four pitch accents

observed in all utterance positions.

The read speech sentences elicited for this paper were based on those used

in IViE (Grabe, 2004) in which changes in meaning were varied only in

terms of utterance type (e.g. declarative vs. question). The wider distribution

of the core pitch accents observed in connected speech data show that it

would be useful to have read speech data in which utterance-internal

information structure is also varied.[16]

5.2   Comparison with Cairene Arabic intonation

The intonational patterns of Sanaani Arabic (SA) differ from those observed

in Cairene Arabic (CA) in a number of ways. SA is compared with CA in

this paper because CA is the dialect which has been most thoroughly described in AM terms (Rifaat, 1991; 2005; Rastegar-El Zarka, 1997; Hellmuth, 2006). Despite differences in theoretical approach, and thus in the representations used, these prior studies agree with respect to key aspects of CA intonation, listed in (5) (see summary in Chahal & Hellmuth, 2013).

(5) a. small pitch accent inventory: one, or at most two, pitch accents
     b. rich pitch accent distribution: accent on almost every content word
     c. post-focal compression, but complete de-accenting is very rare

If these properties of CA are compared with those proposed for SA, in Table 6, a number of differences are apparent. Firstly, SA has a larger pitch accent inventory than CA. Even if a narrowly phonetic model of CA were adopted, in which variation by position in the utterance (pre-nuclear vs. nuclear) is modelled in terms of separate pitch accent categories, the overall size of inventory in the two dialects still seems to be different. Secondly, the distribution of pitch accents is sparser in SA than in CA. Although a pitch accent may be realised on every content word (e.g. in read speech), in the narrative and SRN data unaccented content words are observed. This can be modelled phonologically as a difference in which level of the prosodic hierarchy displays obligatory association of pitch accents to its metrical head (Hellmuth, 2007): in SA the head of every iP bears a pitch accent, whereas in CA the head of every PWd bears a pitch accent. Finally, in SA, words that are repeated in the discourse, and are thus given or old

information, are routinely de-accented (e.g. Fig 3); in contrast, in CA, no prosodic cues mark given status (Hellmuth, 2011). Variation in the degree of de-accenting is a known parameter of cross-linguistic prosodic variation (Jun, 2005; Ladd, 2008).

The multi-level corpus approach adopted in this paper for SA reveals a further potential difference between CA and SA, for which a prosodic analysis is not as yet available in CA, namely in the prosodic realisation of negation and wh-sentences. Watson (1997, p127) observes that in SA interrogatives "the question word generally occupies initial position… and almost invariably attracts the communicative focus". If the term 'communicative focus' is interpreted to denote the primary prominence of the utterance, the present data confirms Watson's observations, and permits a formal autosegmental-metrical analysis of them. In whole-sentence negation contexts (Fig 4), a L*+H accent is observed on the negative particle but appears to be obligatorily followed only by L* accents, resulting in the auditory impression of primary prominence on the negative particle. A similar pattern is observed on wh-words, in wh-questions (Fig 3). In contrast, primary prominence is not reported to be attracted to the negative particle in CA, nor to wh-words (Gary & Gamal-Eldin, 1981).

In SA, when the semantic context additionally invokes a sense of contrast (or 'alternatives': Rooth, 1996; Krifka, 2006), a different pitch accent is used, L+H* or LH*L. Differences in peak alignment depending on focus

32

context have been reported for CA in some studies (El Zarka, 2011), but not in others (Norlin, 1989; Hellmuth, 2009); parallel multi-level corpus data across dialects will allow uncertainty over such differences to be explored. In these contrastive cases, following items in the utterance will be realised with L* if new to the discourse, or unaccented if given (compare the first and second parts of Fig 3, in which speaker m1 provides two versions of the same wh-question), suggesting that future investigation of marking of accessibility, in SA and in other dialects, may merit future research (Baumann & Grice, 2006).

A pattern observed in the SA data which is definitely shared with CA is frequent use of a rise-fall pattern across whole phrases (Rifaat, 2005; El Zarka, 2011); Fig 1 provides an example across a sequence of intermediate phrases within a single IP. This pattern may well be shared by many spoken Arabic dialects, and use of a parallel multi-level corpus approach will allow generalisations to be made about such phenomena in future.

5.3   *Intonational variation in Arabic*

This study has shown that there is variation in the surface intonational patterns of SA and CA, which can be attributed in an AM analysis to differences in the inventory of pitch accents and boundary tones, and in the association of pitch accents to metrical structure.

This paper has argued that intonational analysis is facilitated by availability of a multi-level corpus, which includes a range of speech styles. Read speech data permit establishment of a first hypothesis about the most common pitch accents used in the dialect under investigation. It is invaluable to have tokens in which the position of the accented syllable is systematically varied, so as to be able to compare putatively distinct accents in the same prosodic context, and thus formulate transcription conventions which set out the surface realisation of the proposed inventory of pitch accents in different prosodic contexts. At the same time, however, this study has shown, as might be expected, that the full range of intonational patterns in a language is larger than that observed in read speech data. Complementing read speech data with narrative and conversational data provides a more complete picture, and, in the present study, a first model of SA intonation to test in further research.

Acknowledgements

References

Abdel-Massih, E. T. (1975). *An introduction to Egyptian Arabic*. Ann Arbor: University of Michigan.

Anderson, A., Bader, M., Bard, E. G., Boyle, E., Doherty, G., Garrod, S. et al. (1991). The HCRC Map Task Corpus. *Language and Speech, 34,* 351-366.

Bassiouney, R. (2009). *Arabic sociolinguistics*. Edinburgh: Edinburgh University Press.

Baumann, S. & Grice, M. (2006). The intonation of accessibility. *Journal of Pragmatics, 38,* 1636-1657.

Beckman, M., Hirschberg, J., & Shattuck-Hufnagel, S. (2005). The original ToBI system and the evolution of the ToBI framework. In S.-A.Jun (Ed.), *Prosodic Typology: The phonology of intonation and phrasing* (pp. 9-54). Oxford: Oxford University Press.

Beckman, M. & Venditti, J. (2011). Intonation. In J.Goldsmith, J. Riggle, & A. Yu (Eds.), *The Handbook of Phonological Theory* (pp. 485-532). Oxford: Wiley-Blackwell.

Benkirane, T. (1998). Intonation in Western Arabic (Morocco). In D.Hirst & A. Di Cristo (Eds.), *Intonation systems: a survey of twenty languages* (pp. 345-359). Cambridge, MA: CUP.

Blodgett, A., Owens, J., & Rockwood, T. (2007). An initial account of the intonation of Emirati Arabic. *Proceedings of the 16th ICPhS, Saarbruecken, Germany,* 1137-1140.

Boersma, P. & Weenink, D. (2013). Praat: doing phonetics by computer (Version 5.3.49) [http://www.praat.org] [Computer software].

Broselow, E. (1992). Parametric variation in Arabic dialect phonology. In E.Broselow, M. Eid, & J. McCarthy (Eds.), *Perspectives on Arabic Linguistics IV: papers from the 4th Annual Symposium on Arabic Linguistics.* (pp. 7-45). Amsterdam: John Benjamins.

Bruce, G. (1977). *Swedish word accents in sentence perspective.* Lund: Gleerup.

Brugos, A., Veilleux, N., Breen, M., & Shattuck-Hufnagel, S. (2008). The alternatives (Alt) tier for ToBI: advantages of capturing prosodic ambiguity. *Speech Prosody 2008 (Campinas, Brazil),* 273-276.

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

Chahal, D. (2003). Phonetic cues to prominence in Arabic. *Proceedings of the 15th International Congress of Phonetic Sciences,* 2067-2070.

Chahal, D. (2009). Intonation. In K.Versteegh (Ed.), *Encyclopedia of Arabic Language and Linguistics. Volume 2.* (pp. 395-400). The Netherlands: Brill Academic.

Chahal, D. & Hellmuth, S. (2013). Comparing the Intonational Phonology of Lebanese and Egyptian Arabic. In S.-A.Jun (Ed.), *Prosodic Typology. Volume 2.* (Oxford: Oxford University Press.

El Zarka, D. (2011). Leading, linking and closing tones and tunes in Egyptian Arabic - what a simple intonation system tells us about the nature of intonation. In E.Broselow & H. Ouali (Eds.), *Perspectives on Arabic Linguistics: Papers from the annual symposia on Arabic Linguistics Volume XXII-XXIII: College Park, Maryland, 2008 and Milwaukee, Wisconsin, 2009* (pp. 57-74). Amsterdam: John Benjamins.

El Zarka, D. & Hellmuth, S. (2009). Variation in the intonation of Egyptian Formal and Colloquial Arabic. *Langues et Linguistique, 22,* 73-92.

Gary, J. O. & Gamal-Eldin, S. (1981). *Cairene Egyptian colloquial Arabic.* London: Croom Helm.

Ghazali, S., Hamidi, R., & Knis, K. (2007). Intonation and rhythmic patterns across the Arabic dialects continuum. In E.Benmamoun (Ed.), *Perspectives on Arabic Linguistics: Papers from the annual symposium on Arabic Linguistics Volume XIX: Urbana, Illinois April 2005* (pp. 97-122). Amsterdam: John Benjamins.

Grabe, E. (2004). Intonational variation in urban dialects of English spoken in the British Isles. In P.Gilles & J. Peters (Eds.), *Regional Variation in Intonation* (pp. 9-31). Tuebingen: Niemeyer.

Grabe, E., Nolan, F., & Farrar, K. (1998). IViE - a comparative transcription system for intonational variation in English. *Proceedings of the 5th International Conference on Spoken Language Processing (Sydney, Australia).,* 1259-1262.

Gussenhoven, C. (2004). *The phonology of tone and intonation.* Cambridge: CUP.

Gussenhoven, C. (2006). Between stress and tone in Nubi word prosody. *Phonology, 23,* 192-223.

Gussenhoven, C. (2007). The Phonology of Intonation. In P.de Lacy (Ed.), *The Cambridge Handbook of Phonology* (pp. 253-280). Cambridge: Cambridge University Press.

Hellmuth, S. (2006). *Intonational pitch accent distribution in Egyptian Arabic.* PhD Unpublished PhD thesis, SOAS.

Hellmuth, S. (2007). The foot as the domain of tonal alignment of intonational pitch accents. *Proceedings of the 16th ICPhS, Saarbruecken, Germany*.

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

Hellmuth, S. (2009). The (absence of) prosodic reflexes of given/new information status in Egyptian Arabic. In J.Owens & A. Elgibali (Eds.), *Information Structure in Spoken Arabic* (pp. 165-188). Oxford: Routledge.

Hellmuth, S. (2011). Acoustic cues to focus and givenness in Egyptian Arabic. In B.Heselwood & Z. M. Hassan (Eds.), *Instrumental Studies in Arabic Phonetics. [CILT319]* (pp. 299-324). Amsterdam: John Benjamins.

Hellmuth, S. (2013). Phonology. In J.Owens (Ed.), *The Oxford Handbook of Arabic Linguistics* (pp. 45-70). Oxford: Oxford University Press.

Hellmuth, S. (2014). Evidence of distinct intonational properties in formal vs. colloquial Arabic? A case study. In E.Delais-Roussarie, M. Avanzi, & S. Herment (Eds.), *Prosody and languages in contact: L2 acquisition, attrition, languages in multilingual situations.* (Berlin: Springer Verlag.

Hoyt, F. (2014). Prosodic constituency and locality in Levantine Arabic Long-Distance Negative Concord. In R.Khamis-Dakwar & K. Fround (Eds.), *Perspective on Arabic Linguistics 26* (Amsterdam: John Benjamins.

Hyman, L. (2006). Word-prosodic typology. *Phonology, 23,* 225-257.

International Phonetic Association (1999). *Handbook of the International Phonetic Association.* Cambridge: Cambridge University Press.

Jun, S.-A. (2005). Prosodic typology. In S.-A.Jun (Ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing* (pp. 430-458). Oxford: OUP.

Krifka, M. (2006). Basic notions of information structure. In C.Féry, G. Fanselow, & M. Krifka (Eds.), *The notions of information structure* (Institutsverlag Universitaet Potsdam.

Kulk, F., Odé, C., & Woidich, M. (2003). The intonation of colloquial Damascene Arabic: a pilot study. *IFA Proceedings, 25,* 15-20.

Ladd, D. R. (2008). *Intonational Phonology*. (Second Edition ed.) Cambridge: Cambridge University Press.

Liberman, M. & Pierrehumbert, J. (1984). Intonational variance under changes in pitch range and length. In M.Aronoff & R. T. Oehrle (Eds.), *Language Sound Structure: Studies in Phonology present to Morris Hallé* (pp. 157-233). Cambridge, M.A.: MIT Press.

Llamas, C. (2007). A new methodology: data elicitation for regional and social language variation studies. *York Papers in Linguistics Series 2, 8,* 138-163.

Naïm-Sambar, S. (1994). Contribution á l'étude de l'accent yéménite: le parler des femmes de l'ancienne génération. *Zeitschrift fur arabische Linguistik/Journal of Arabic Linguistics /Journal de Linguistique Arabe,* 67-89.

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

Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation.* MIT.

Post, B., D'Imperio, M., & Gussenhoven, C. (2007). Fine phonetic detail and intonational meaning. *Proceedings of the 16th ICPhS, Saarbruecken, Germany,* 191-196.

Prieto, P., D'Imperio, M., & Gili Fivela, B. (2005). Pitch accent alignment in Romance: primary and secondary associations with metrical structure. *Language and Speech, 48,* 359-396.

Prieto, P. & Frota, S. (2014). *Transcription of Intonation in the Romance Languages*. Oxford: Oxford University Press.

Prieto, P. & Roseano, P. (2010). *Transcription of Intonation of the Spanish Language*. Muenchen: LINCOM.

Prieto, P., van Santen, J., & Hirschberg, J. (1995). Tonal alignment patterns in Spanish. *Journal of Phonetics, 23,* 429-451.

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

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

Rifaat, K. (2005). The structure of Arabic intonation: a preliminary investigation. In M.T.Alhawary & E. Benmamoun (Eds.), *Perspectives on Arabic Linguistics XVII-XVIII: Papers from the seventeenth and eighteenth annual*

symposia on Arabic linguistics. (pp. 49-67). Amsterdam/Philadelphia: John Benjamins.

Rooth, M. (1996). Focus. In S.Lappin (Ed.), *The Handbook of Contemporary Semantic Theory* (pp. 271-297). Oxford: Blackwell.

Selkirk, E. O. (2000). The interaction of constraints on prosodic phrasing. In M.Horne (Ed.), *Prosody: theory and experiment* (pp. 231-262). Dordrecht: Kluwer.

Selkirk, E. O. (2011). The Syntax-Phonology Interface. In J.Goldsmith, J. Riggle, & A. Yu (Eds.), *The Handbook of Phonological Theory* (Second Edition ed., pp. 435-484).

Siemund, R., Heuft, B., Choukri, K., Emam, O., Maragoudakis, E., Tropf, H. et al. (2002). OrienTel - Arabic speech resources for the IT market. *Proceedings of LREC 2002 (Arabic workshop)*.

Skopeteas, S., Fiedler, I., Hellmuth, S., Schwarz, A., Stoel, R., Fanselow, G. et al. (2006). Questionnaire on Information Structure (QUIS). *Interdisciplinary Studies on Information Structure (ISIS): Working Papers of SFB632 University of Potsdam., 4.*

van der Hulst, H. & Hellmuth, S. (2010). Word accent systems in the Middle East. In R.Goedmans & H. van der Hulst (Eds.), *A Survey of Word Accentual Patterns in the Languages of the World* (pp. 615-646). Berlin: Mouton de Gruyter.

Watson, J. C. E. (1993). *A syntax of San'ani Arabic*. Wiesbaden: Harrossowitz Verlag.

Watson, J. C. E. (1996). *Sbahtu! A course in Sana'ani Arabic*. Wiesbaden: Harrassowitz Verlag.

Watson, J. C. E. (2002). *The phonology and morphology of Arabic*. Oxford: OUP.

Watson, J. C. E. (2011). Word stress in Arabic. In M.Oostendorp, C. Ewen, E. Hume, & K. Rice (Eds.), *The Blackwell Companion to Phonology. Volume 5* (pp. 2990-3018). Oxford: Blackwell.

Watson, J. C. E. & Asiri, Y. (2009). Pre-pausal devoicing and glottalisation in varieties of the south-western Arabian Peninsula. *Langues et Linguistique, 22,* 17-38.

Watson, J. C. E. & Bellem, A. (2011). Glottalisation and neutralisation in Yemeni Arabic and Mehri: An acoustic study. In Z.M.Hassan & B. Heselwood (Eds.), *Instrumental Studies in Arabic Phonetics* (pp. 235-256). Amsterdam: John Benjamins.

Wellens, I. (2005). *The Nubi language of Uganda: an Arabic creole in Africa.* Amsterdam: BRILL.

Yeou, M. (2004). Effects of focus, position and syllable structure on F0 alignment patterns in Arabic. *Proceedings of JEP-TALN 2004*.

Yeou, M., Embarki, M., AlMaqtari, S., & Dodane, C. (2007). F0 alignment patterns in Arabic dialects. *Proceedings of the 16th ICPhS, Saarbruecken, Germany,* 1493-1496.

Appendix

Figure 1: Sample read speech declarative [ys-dec1-f1] ("This man is Yemeni"), with a typical pattern: a sequence of L*+H pre-nuclear accents then H* nuclear accent, realised in a compressed pitch range due to final lowering (!H*, see 4.2).
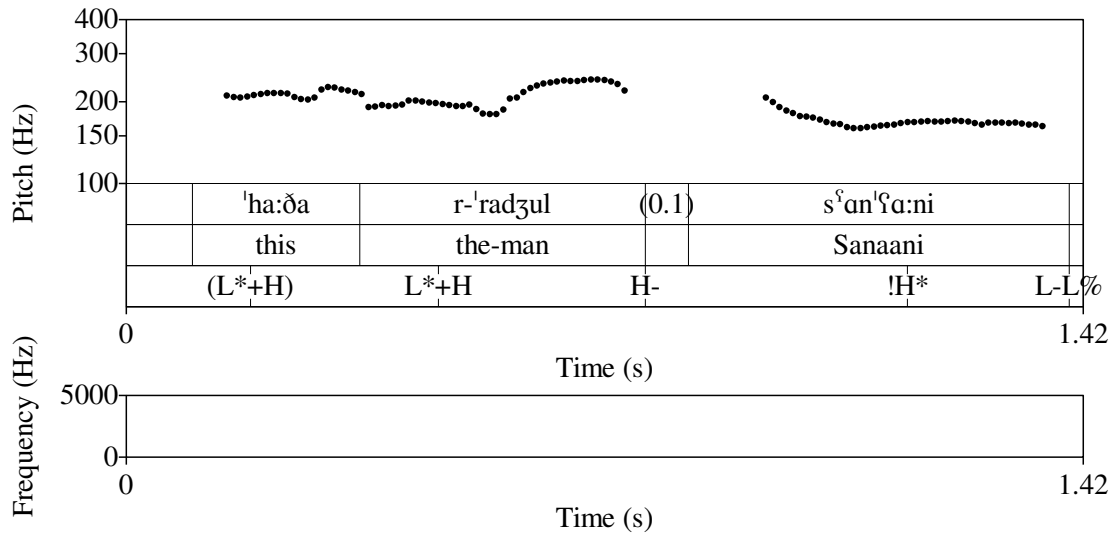
Figure 2: Sample read speech yes-no question [ys-ynq1-m1] ("Is this man Yemeni?") with a L+H* nuclear accent on the last lexical item; the speaker role-plays afelicitous response ("Yes"), produced with a typical declarative falling contour.

41
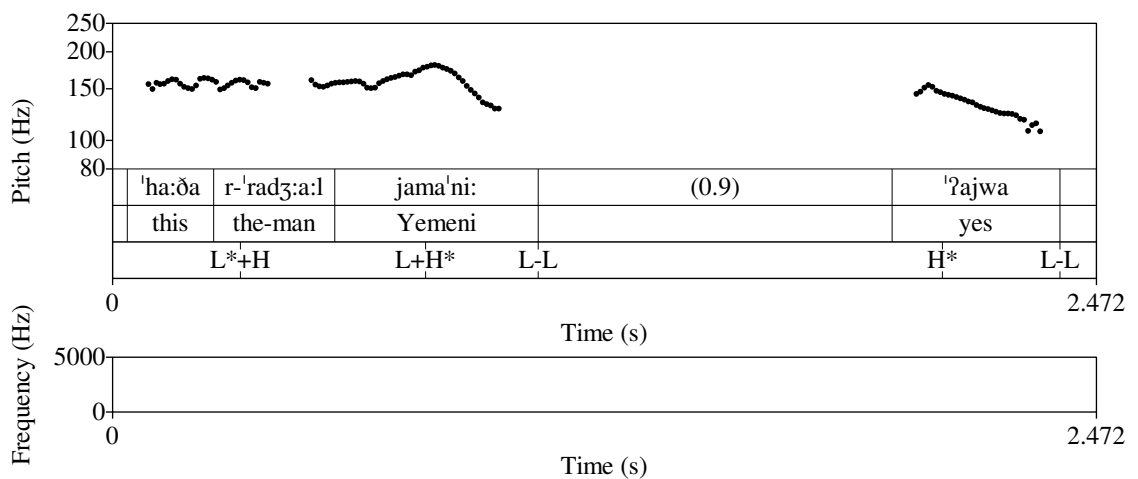
Figure 3: Sample read speech wh-question [ys-whq2-m1] ("What is the Yemeni man's name?"), with focus on the wh-word, followed by L* accents; the speaker then provides a reformulation in which repeated lexical items are fully de-accented.

ma:	ʔism	ar-raˈdʒ:a:l	l-ˈjamani	ˈʔiħna	nəˈgu:l-	0.08b-ism	ar-raˈdʒ:a:l	l-ˈjamani	
what	name	the-man	the-Yemeni	us	we-say	what-name	the-man	the-Yemeni	
L*+H		L*	(L*) L-L%		L*+H	H-H%L+H*		L-L%	
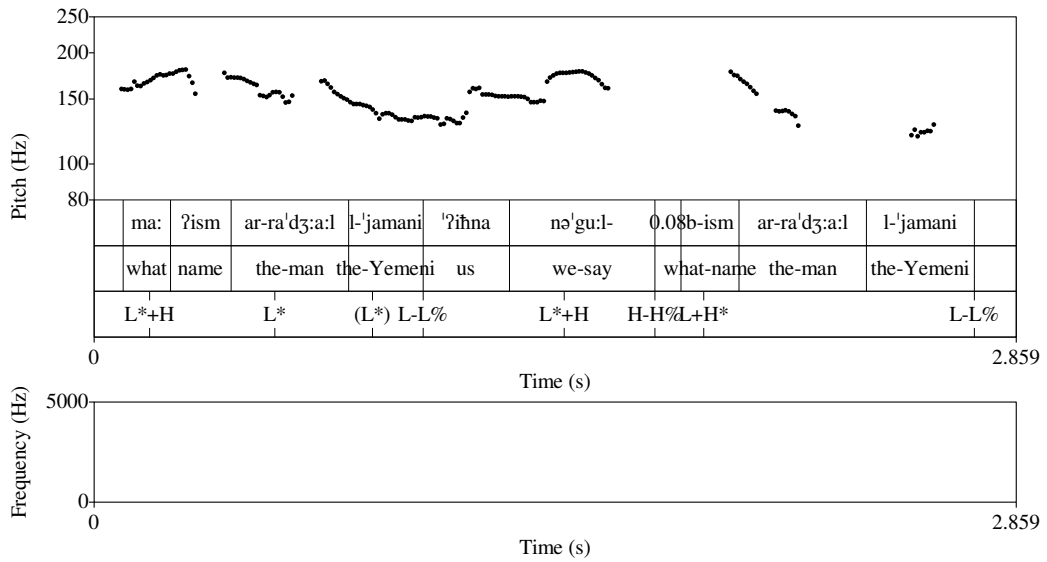
Figure 4: Excerpt from a retold narrative [ys-nar-f1_37-38] which includes advice about bartering techniques ("You don't want to give them fifty"), showing early focal prominence on the negative particle, followed by L* accents.

la:	tərˈdˤɑ:-ʃ	tiˈdʒib-lu-hum	al-xamˈsi:n		
NEG	you.want-NEG	you.give-to-them	the-fifty		
L*+H	L*	L*		L*	L-L%

42
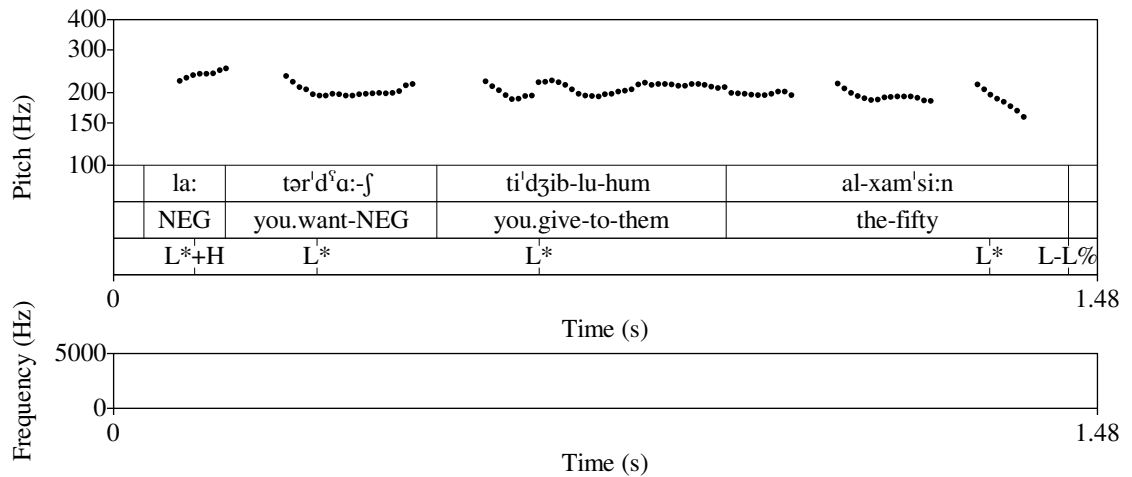
Figure 5: Excerpt from retold narrative [ys-nar-f1_37-38] ("So he arrived in Sanaa…"), showing an early peak L+H* accent on the utterance-initial word, at the start of a new discourse topic.

	ə::	dʒa:	ila	sˤanˈʕa:	
	um	he.came	to	Sanaa	
	L+H*			L*+H	H-H%
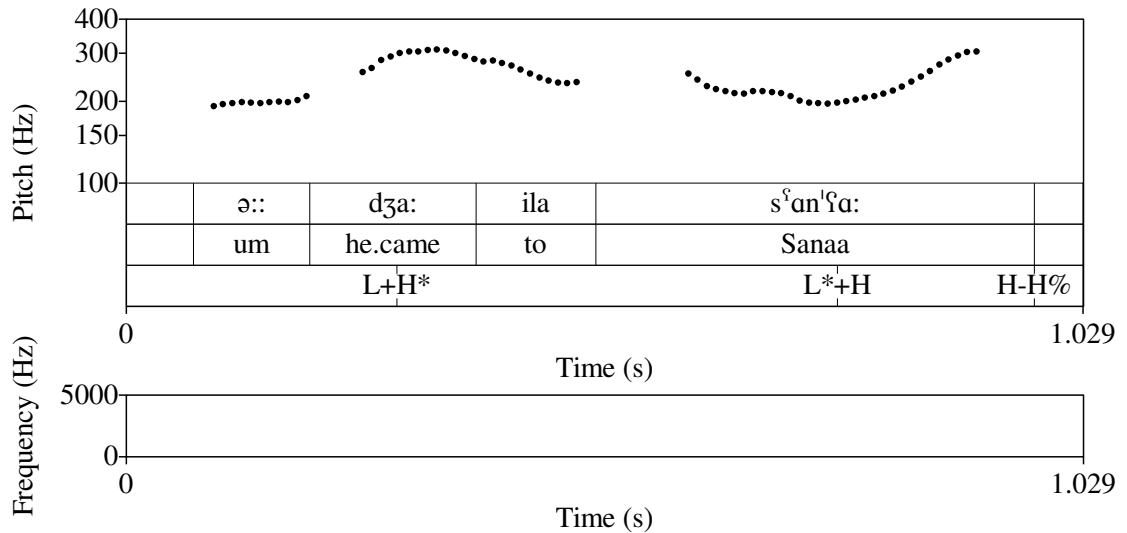
Figure 6: Excerpt from Sense Relation Network [ys-srn-f2_f3_271-276] ("We rarely use 'friend', because we've forgotten it over time, we use the new (word)."), showing LH*L focal accents (see 4.3) and intermediate junctures marked by L-.

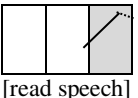| ˈnaːdiran	maː	nasˈtaxdam	xaˈbiːr-iː	0.1	laˈʔannuad	naˈsiːna-ha:	maʕa	z-ˈzaman	0.3	nasˈtaxdam	al-ʒadiːd
rarely	maː	we.use	"friend-my"	becauPART		we.forgot-it	with	the-time		we.use	the-new
LH*L		L*		L*	L-	H*	LH*L		L*+H L-	H*	L* L-L%
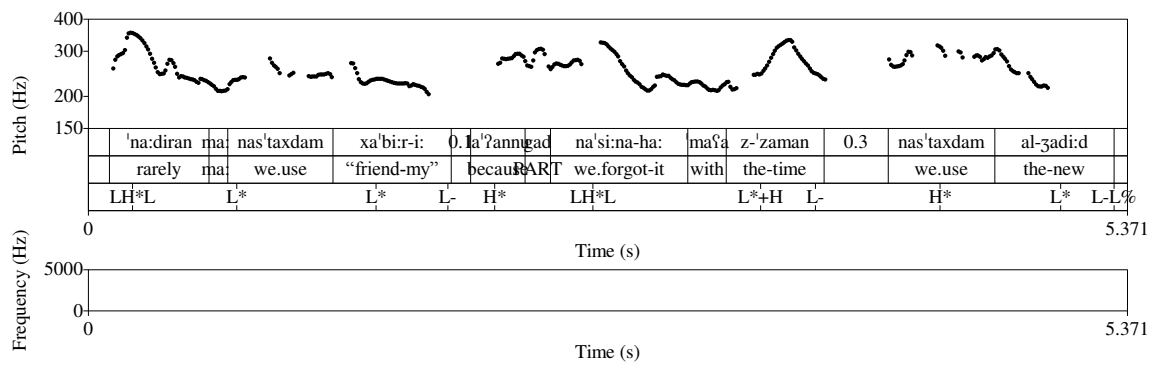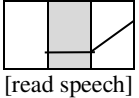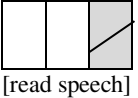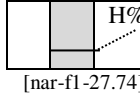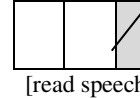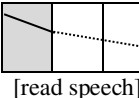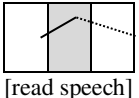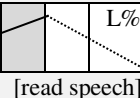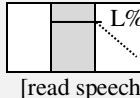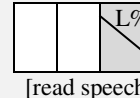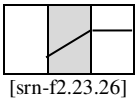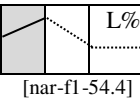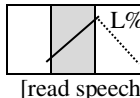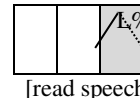
	as pre-nuclear accent, by stress position:			as nuclear accent, by stress position:		
	antepenult	penult	final	antepenult	penult	final
L*+H	no tokens	[read speech]	[read speech]	H% [nar-f1-24.36]	H% [nar-f1-27.74]	[read speech]
H*	[read speech]	[read speech]	no tokens	L% [read speech]	L% [read speech]	L% [read speech]
L+H*	no tokens	[srn-f2.23.26]	[read speech]	L% [nar-f1-54.4]	L% [read speech]	L% [read speech]

Table 7: Schematic representation of observed positional variation in surface realisation of the most common SA pitch accents.

	Age	Education	Occupation	Father Place of Birth	Mother Place of Birth
f1	20	Completed secondary school	Cleaner	Hamdaan, Greater Sanaa	Hamdaan, Greater Sanaa
f2	35	Completed primary school	Housewife	Al Ga Sanaa	Al Ga Sanaa
m1	29	University graduate	Administrator	Old City Sanaa	Old City Sanaa

Table 8: Background information about the SA participants.

[1] In British and American English, the main prominence of the utterance is almost always the last pitch accent in the utterance, with all following words realised either unaccented or in a highly compressed pitch range. As a result, in English 'main prominence' is generally assumed to equate to 'nuclear accent'. This cannot be assumed a priori for other languages, and thus on occasion a distinction is made in this paper between 'nuclear accent' (denoting the last accent in the utterance) and 'main prominence' denoting the most prominent accent in the utterance.

[2] The only proposed exceptions are at the margins of the Arabic language family: e.g. Nubi, an Arabic-based creole spoken in Uganda, in which stress appears to have been reinterpreted by speakers as H tones (Wellens, 2005; Gussenhoven, 2006).

[3] Some studies have explored issues in the phonetic realisation of intonational categories, such as peak alignment, using parallel data across dialects (Yeou, 2004; Yeou, Embarki, AlMaqtari, & Dodane, 2007).

[4] See Hoyt (2014) for a welcome contribution to this emerging field of research.

[5] The IViE corpus for British English (www.phon.ox.ac.uk/IViE/); cf. also the Interactive Atlas of Spanish Intonation (http://prosodia.upf.edu/atlasentonacion/index-english.html) and the Interactive Atlas of the Prosody of Portuguese (http://ww3.fl.ul.pt/LaboratorioFonetica/InAPoP/demo/index.htm).

[6] Sense Relation Network tool (see Llamas, 2007).

[7] The revised set of tools will be made available at www.york.ac.uk/res/ivar/.

[8] All figures are provided in the Appendix of the paper, and referred to from various points in the text as needed.

[9] This example also displays pre-pausal stress migration, another feature of SA reported by Watson (2002 ch.5).

[10] Comparison to one speaker from another dialect is reported here, as an example. Full analysis of the lexical variation observed among all 12 speakers is beyond the scope of the present paper, but will be reported in future work.

[11] In Sp_ToBI this pitch accent would be transcribed by convention as L* (Prieto & Roseano 2010:3); L* appears to have a different, wider distribution in SA. A non-lowered realisation of H* can be seen in a one word utterance, [ʔajwa] 'yes', in Fig 2.

[12] These L*+H H-H% cases are potentially open to re-analysis, either of the boundary tone as a zero boundary tone '0%' (Grabe, Nolan, & Farrar, 1998), or of the nuclear accent as L*, with the following rise due to the H-H% boundary.

[13] These cases are open to re-analysis as H* preceded by a high initial boundary tone '%H'.

[14] IP = Intonational Phrase; iP = Intermediate Phrase (see section 2.1).

[15] Since L-/H- are used in this paper to account for intermediate junctures, the analysis here also uses the original ToBI style notation at IP boundaries, combining a phrase tone + boundary tone at each IP edge (Beckman et al., 2005). An alternative approach would be to follow Prieto & Roseano (2010) in transcribing phrase tones only at IP-internal iP junctures, and marking IP edges with a single boundary tone. In the alternative scenario mid/complex boundary tones would have to be proposed, to capture the full range of observed contours: mid = M%, fall-rise LH%.

[16] Cf. Dialogue Completion Tasks (Prieto & Roseano, 2010), and/or read speech sentences elicited in a dialogue or frame-setting paragraph (cf. Norlin, 1989; Hellmuth, 2009).