Jeddah Arabic Intonation: An Autosegmental-metrical Approach

Hajar Ahmed Moussa

Doctor of Philosophy (Integrated): IPhD in Phonetics and Phonology

School of Education, Communication and Language sciences

Newcastle University

November 2019

Abstract

This thesis is a theoretical and instrumental investigation of intonation in Jeddah Arabic, an urban Arabic variety spoken in west Saudi Arabia. The study is carried out in an attempt to establish the dialect's prosodic properties and to widen the scope and volume of the literature on Arabic prosody that would in turn aid in the cross-dialectal comparison of prosodic and intonational patterns. The investigation is carried out in light of the *Auto-Segmental Metrical* theory of intonation- a theory that has been reported to account for the intonational patterns of many languages. In AM theory, intonation is manifested via prominent F0 behaviour in interaction with phonological structure, hence maintains a close relationship between accent distribution and phonological/metrical structure. This F0 behaviour is examined acoustically through pitch level, range and excursion size, in the form of increased peak height and excursion, pitch compression or absence thereof to mark intonational structure. In addition to pitch, other acoustic correlates such as duration and amplitude are examined as well. The thesis includes the examination of the different tunes, postlexical phrasing, and accent categories (contour shapes) that occur in the dialect. Moreover, and as an integral part of AM analysis, the thesis closely examines both theoretically and acoustically the concepts of tonal alignment and accentuation and information structure in this Arabic dialect. Data for the study were collected from 20 native male and female speakers of Jeddah Arabic. Data were then semi-automatically segmented and manually transcribed using a modified TOBI system for Arabic. It is found that JA speakers rely on both qualitative and quantitative detail to enhance intonationally important material that is conveyed prosodically. The results also point to that JA is a stress-accent language that is although similar to other languages in this group, contributes differently to the general cross-language prosodic variation. The dialect demonstrates prominent pitch accents that faithfully associate and align with stressed syllables and are distributed in two intonational levels above the prosodic word: the intermediate phrase and the intonational phrase. Those two intonational levels are found to be marked by both tonal and non-tonal correlates. Experimental evidence shows that contrary to the typical reported correlates of those prosodic constituents, in JA intermediate phrases boundaries demonstrate longer pre-boundary units than intonational phrases. This non-tonal pattern in intermediate phrase boundaries correlates with later alignment of the tone with respect to the onset of the stressed syllable.

To my mother…

Acknowledgments

First and always, I thank God for everything he gave me and for the blessing that is reaching this step. My profound gratitude and appreciation go to all the people whose encouragement, moral and academic support made this work possible. Among them is my family, friends and my supervisor. I dedicate this thesis to you and hope that this makes you and will always make you proud of me.

Table of contents

List of tables

List of figures

Chapter 1. General introduction

This thesis is a linguistic investigation into the Arabic variety spoken in Jeddah, Saudi Arabia. Jeddah Arabic is an urban Hijazi variety spoken in the commercial capital on the coast of the Red Sea. Alongside commerce, the city is considered a cultural and tourist destination hosting international book fairs, festivals, and home to a number of national official newspapers, TV and radio broadcasting corporations. As the city holds one of the major gateways in the country, the language spoken there is the main tool for the communication between residents and visitors for different reasons ranging from commerce, trading, and various services. More so, Jeddah Arabic is one of the two main varieties used by media outlets nationwide. It is the language used for advertisement, broadcasting and modern social media presence and outreach. It is also worth mentioning that the dialect also possesses its own informal writing style and notation that is used in the city across the different age and gender groups. This informal dialectal style mirrors the spoken variety and is used for informal means of communication, e.g. texting and chatting[1].

The dialect is well- recognised among both, the other Hijazi varieties and the other Arab dialects, and its unique properties can be informally recognised and described by naïve listeners. As a native speaker, the general remarks I get about the dialect are more often related to the suprasegmental aspect. Several studies have reported this effect, whereby listeners were able to discriminate dialects relying mainly on suprasegmental information. Among them, a language identification study conducted by Barakat, Ohala, and Pellegrino (1999) that reports by synthesising segmental information of utterances, native Arabic listeners were able to identify and distinguish between dialects relying on the prosodic cues of F0, and amplitude. This is to be expected since to some extent, the segmental and lexical aspects of the dialect are shared and intelligible among most Arabic dialects in the region. Therefore, it was the goal to linguistically investigate the ways and structures that make Jeddah Arabic prosodically unique. With this main goal in mind, the analyses were set to descriptively and theoretically establish the phonetic and phonological properties of Jeddah Arabic prosody regarding intonation that demonstrate its place in the cross-linguistic intonational typology. The Autosegmental-metrical theory of intonational phonology was employed for this purpose, as the theory expresses intonation via a mechanism that involves an interaction between the surface tones

[1] This information about the city is available at the Jeddah Municipality website. Online at: https://web.archive.org/web/20090717005415/http://www.jeddah.gov.sa/english/

and the language-specific underlying prosodic structure. Thus, the investigation would demonstrate both the surface and underlying properties essential to Jeddah Arabic intonation.

Traditions in cross-linguistic descriptions of intonation make reference to a suprasegmental melodic contour that is described in terms of tones. The contour is assumed to be comprised of a sequence of (H)igh and (L)ow tones representing the musical nature of spoken language. Those contours may be used to mark certain chunks in the discourse. That is, they may be used to group certain syntactic parts of the sentence together for an intended linguistic meaning. The tones within the contours may also be used to make a contrast between the elements in the speech signal making some different or more relevant than others depending on discourse. Ultimately, the intonation researcher would seek to figure out the function of these tones in the respective language they are studying, and whether or not those tones consistently coincide with certain locations or groups in the speech signal. This would include the analysis of the shapes/forms, locations and phonetic specifications of the tones, and what purpose these tones serve in the contour.

Tones then, seem to be regarded as the primitives on which intonational contours are constructed. The influential Autosegmental- Metrical theory (AM henceforth) adopted in this thesis proposes a formalisation that models the observed intonational patterns/contours. The model distinguishes the phonological notion of a Tone from its phonetic reality. The premise is that an underlying/abstract phonological Tone target: High or Low, translates into the surface variations in pitch/F0 observed in the contour. In this framework, intonational contours are analysed as a sequence of underlying target Tones as 'autosegments' that 'associate' with linguistically relevant units and edges of constituents in the metrical structure of an utterance. This phonological association gives rise to two types of pitch/F0 configurations differing in their function on the surface as either lending prominence: Pitch Accents, or demarcating boundaries: Boundary Accents or boundary tones. The framework also acknowledges the manner of the surface 'alignment' of F0 turning points to indicate their affiliation to the prosodic constituents.

It is proposed that linguistically relevant pitch configurations function by either adding value to the component they mark or by demarcating the speech chunks that the speakers may deem to be related. In terms of value, it is proposed that [T]ones demonstrate a direct relation to [P]rominence, i.e. more often, a tonal event occurring on a certain unit makes it more prominent to the listener. For example, a language may use a pitch accent to make one syllable more prominent than another, while other languages may use pitch accent to make two words different in meaning. In addition to F0, this prominence can further be enhanced by other suprasegmental properties like loudness and duration. In terms of demarcation, some pitch configurations occur at edges of certain units that signal the beginnings and ends of these units intonationally. In other words, those edge configurations signal the intonationally- relevant phrases used by the speaker for multiple discourse functions. In this case, too, alongside F0, other properties like pauses and duration can be used to signal the different junctures. What can be noticed here is that intonation description takes into account the tonal configurations in a language and how they function in the intonational structure of that language.

Regarding the phonetic forms of tonal events, it is proposed that pitch accents demonstrate contextual variation according to the pitch range and location of a tone. For example, an H accent is expected to be realised lower in the speaker range when it occurs finally in a sentence, while in comparison it is expected to be higher in F0 value if it occurs non-finally, thus showing variation in the level of the accent. The main speculation regarding this variation is whether a language marks those contextual differences in the phonology or phonetics, i.e. are the two accents categorically different or phonetically different? Moreover, differences in shape, e.g. L vs. H are assumed to mark distinct tonal categories in the underlying phonological structure of a language, whereby the L is observed to be at a low point in the contour, and an H is seen as a 'peak' high point in the contour. Depending on how they operate in the contour, categorical differences in tonal form also serve different pragmatic meanings in discourse. Thus the combinations thereof give rise to the different tunes in the language.

The previous aspects of intonation description have a number of implications in terms of the classification of a language within prosodic typology. The first implication regards how a language employs F0/pitch configurations. If the language employs pitch to mark lexical contrasts, it is considered a tone language where F0 is used to make grammatical contrasts. On the other hand a language may allocate a pitch event to make some syllables more prominent than others with reference to a metrical stress structure in the lexical level. This language then is considered a stress-accent language where 'accent' is used postlexically to make syntagmatic contrasts. However, the distinction is seldomly straightforward as in most intonational analyses- including AM- prosodic patterns interact with lexical features in the word level as mentioned above. It is thus reported that some prosodic structures may utilise pitch both lexically and postlexically. In Jun (2005) a survey of prosodic systems shows that in some languages postlexical pitch accents can be from lexical accents (Japanese), from lexical stress

(English) or postlexical stress (French). In Japanese for example, pitch is used to mark a prosodic constituent boundary as well as distinguishing two words in the lexical level. In French lexical units have invariant pitch patterns marking edges of prosodic domains in the postlexical level. In both languages there is a dense distribution of pitch accents whereby almost every prosodic word is denoted a pitch accent (ibid). In English pitch accents are used to make only a selected number of syllables more prominent than neighbouring ones postlexically. However those pitch accents in English only fall on syllables that are lexically stressed thus confounding the phonetic correlates for stress and intonational accent (Ladd, 2008). Therefore there was a need to distinguish between lexical prominence (stress) and intonational prominence (postlexical) in how and whether each level employs F0 configurations.

Beckman (1986) makes an elaborate effort to differentiate between word-level/lexical prominence and phrasal/postlexical prominence as both may employ pitch/F0 as a phonetic parameter in some systems. That is, her discussion was set to explore the acoustic correlates of stress and the correlates of accent cross-linguistically to establish a typology according to how a language makes use of pitch to convey prominence. Bearing in mind that both levels (stress and intonation) in Autosegmental-metrical theory make use of constituents in the prosodic hierarchy to express their properties. For this, she compares English (a stress-accent language) and Japanese (a non-stress-accent language), and finds that stress-accent languages typically use the acoustic correlates of duration, amplitude and vowel formants for 'Stress', while pitch is used to denote 'Pitch Accents' aligned to stressed syllables that make up an intonational contour. On the other hand, a non-stress-accent language like Japanese uses melodic pitch to convey lexical level contrasts and intonational pitch accents with no reference to a stress system. However, as later studies advise (Hellmuth, 2006a, De Jong & Zawaydeh, 1999, among others), there is indeed a variation in the extent of how languages at the two ends of the spectrum employ those acoustic cues.

The second implication would be regarding the unit that is considered prosodically relevant in a language, i.e. the phrasing levels that a language tonally marks above the prosodic word. It has been reported in the literature that part of the variation in intonational typology is due to the number of intonationally marked prosodic levels above the word. Languages are reported to show at least one level, or a combination of intonationally marked phrases. Among those reported domains are The Intonational Phrase, which is considered a high level of intonationally marked prosodic constituents, following is the intermediate phrase, reportedly equivalent to the syntactic phonological/ major phrase and then the Accentual Phrase, equivalent to a minor phrase. The variation in prosodic levels above the word has implications regarding the domain of pitch accent distribution. Parallel to the accepted notion of a stressed syllable being more prominent phonetically and phonologically than a neighbouring unstressed syllable in the lexical level, it is proposed that one accented word can be more prominent than a neighbouring unaccented word in the phrasal level. In other words, there is a structure that proceeds to distribute prominence among words in those intonationally marked phrases.

The third related implication is regarding prominence distribution in phrases. How speakers of a language intonationally distribute prominence has been suggested to demonstrate a rhythmic nature parallel to rhythmic organisation of stress intervals in the lexical level. Prosodic phrases above the word have been shown to demonstrate a rhythmic nature alternating sequences of prominent and non- prominent (accented vs. unaccented) words and syllables. The boundaries of those phrases may begin or end in prominent words or syllables on the left or on the right of the phrase, the sequencing may be broken down by junctures or breaks, and rhythm may also be derived by lengthening the end of a unit or strengthening the start of the unit. Also for the sake of interval rhythmicity, a language may choose to break a large constituent into smaller sized constituents with both independent tonal head and edge marking, and long-distance association with the main unit. Accordingly, after specifying the prominence intervals, the need is to specify the potential landing sites where tones are linked to those words. That is, whether the tones are linked to the designated head or edge of the level that a language chooses to distribute accents within. In which phrases do speakers distribute those accents in and by which means is where the cross-linguistic typological variation lies. The phrasing structure is also reported to be one of the means that indicate the information and intended meaning of an utterance. Several studies report that relative prosodic prominence aids in signalling the information structure and meaning of an utterance (Yeou, et al., 2007a), (Face, 2002), among others. This is particularly related to the concept of *Focus* ‘highlight or emphasis’, which is a type of discourse meaning signalled by prosody. Focus is manifested when a part of the sentence is allocated the most prominent pitch accent. As will be discussed in the upcoming chapter five, the intended type and location of focus may cause a change to the phrasing structure of a sentence, and subsequently influence the pitch accent distribution in a sentence.

The fourth implication regards the phonetic component whereby evaluating this link between the tone and the prosodic level it is linked to on the surface. It is assumed that pitch is an auditory quality that translates into the acoustic property of fundamental frequency (F0). It is also known that the course of an F0 contour can be affected by some segmental and contextual factors. Taking those sources of variability into account, an intonation researcher would want to evaluate whether the assumed link between a tone and a constituent can also be observed on the surface as produced by the speakers. For this matter, different experimental manipulations are carried out to observe how a language links the two phonetically. This connection between tone and segment has been shown to vary cross-linguistically, where in some languages an early vs. late alignment between the two results in two different tonal categories, whereby indicating that the phonetic reality of alignment does have an effect on the underlying tonal information of a language and how tonal form is perceived by listeners. Meanwhile in other languages tones may be shown to link to different components within the same landing site. For instance, in the case where two languages both link their tones to the accented word in a phrase, one may choose to realise the pitch accent on the head of the syllable rhyme, while the other structurally realises it on the second mora of a stressed foot. To date, the latter phenomenon of linking a tone to the same landing site but in relation to different levels in the prosodic constituent has been taken to generally indicate structural cross-linguistic or cross-dialectal differences in alignment (Hellmuth, 2006, 2019 among others).

The thesis therefore aims to research the tonal composition of the dialect, how intonational prominence is marked at the phrasal level, the essential phrasing constituents in the dialect, how intonational prominence affects accent distribution, and how the tones align to the speech units. A prosodic constituent hierarchy for the phrasal level will demonstrate the essential prosodic 'chunks' or intonational groups that are signalled prosodically in the dialect and how they collaborate to signal the meaning of an uttered unit. This aspect would also reflect the rhythmicity patterns essential to intonation expression in the dialect. A tonal and tune inventory will establish an intonational model for the dialect that in turn establishes how prominence is marked at this phrasal level in terms of accent shapes and their function as prominence- lending or edge marking. It will also demonstrate how prominence marking proceeds to distribute the relevant accents in a prosodic domain according to their function, unravelling the potential interaction between tones and prominence. Regarding the alignment of tones, investigation of this aspect ties in with the previous goals and would show us phonetically how the phonological tone associates with the prosodic structure, and by manipulating possible prosodic effects, how this variation is observed on the surface. The research questions regarding each aspect are presented in the following. Each chapter in the thesis is dedicated to investigating these questions accordingly. The structure of the chapters is also discussed thoroughly in the following.

Chapter two presents the fundamental properties and notions adopted by the Autosegmental- metrical theory of intonational phonology, which are basic notions assumed in the dialect to govern the interaction between text and tune. Additionally, the chapter reviews the assumed phonetic and phonological components of intonational description in the theory. In this chapter, it is explored how AM theory represents a pitch contour, the essential components of this contour, and how it is observed on the surface including the tonal implementation rules affecting the realisation of the contour. The chapter includes a presentation of how those notions and aspects are demonstrated cross-linguistically and cross-dialectally.

Chapter three is dedicated to the thorough description of the methodology employed in the thesis. The methodology concerns the corpus design, the data collection and analyses processes, and the recording and annotation processes. The chapter presents details on how the employed methods would aid in answering the research questions in the thesis.

Chapter four presents details on the tonal composition in the dialect. It presents the tonal inventory of JA intonational categories of pitch accents, phrase accents and boundary tones along with illustrations of their phonetic reality. It also phonetically demonstrates the pragmatically different tunes and melodies in the dialect and their makeup. Guiding the analyses in this chapter are the following questions: what are the accents used by JA speakers? What factors influence the realisation of an accent in a contour? And do those factors prompt a categorical or gradient difference? How are prominent words/syllables distributed within the sentence? Is there a ranking of those accents? Do they show consistent intervals that coincide with the grammatical grouping of the sentence? And how are the boundaries of those intervals marked phonetically and phonologically? The chapter then concludes with a model of JA intonational phonology, which includes the prosodic hierarchy and phrasal constituents used by this variety.

Chapter five is an experimental analysis of the within- phrase accent distribution mechanism suggested in the previous chapter. This includes testing of the hierarchical organisation of accents in a phrase, and what this organisation entails phonetically as a function of the information structure of the sentence. It also includes an investigation of the ways this dialect uses Focus to express this structure. Guiding the analyses here are the following speculations: How does rhythmicity interact with information structure in the dialect? What happens when we interrupt the expected rhythmicity of a phrase that begins or ends in a prominent nuclear syllable by placing the prominent word mid- sentence? Does this entail making changes to the phrasing levels used in the utterance? In other words, what if the focused constituent is anywhere else but final in a phrase, as the nuclear accent location dictates. How is focus realised then? Also, what if the uttered string is a neutral sentence; a broad focus sentence with no focus on a particular constituent, how is prominence realised then? How does the nuclear accent distribution take place? Finally, what correlates can be said to consistently mark narrow focus, and what correlates are used to mark broad focus in JA?

Chapter six is the second experimental analysis regarding the tonal alignment patterns in JA. It includes results on how the relevant tones align with the segmental tier, which in turn reflects the exact level 'tone bearing unit' in prosodic structure that is associated with this phonological tone. The following questions motivated the analyses in this chapter: From cross-linguistic evidence, contrasting tonal patterns of nuclear accents are expected to be correlates for the existence of the different prosodic levels observed in a language? Does JA mark those differences tonally, non-tonally, or both? In what manner? Moreover, it is known that a number of factors may effect tonal alignment in some languages? How does JA behave in this respect? Finally, the literature on Lebanese and Egyptian Arabic report a cross-dialectal variation in the exact location of tone relative to the Tone bearing unit? Is JA similar to Lebanese in aligning a tone with the stressed syllable or similar to Egyptian in aligning it with the stressed foot? Or does it constitute a new category?

Chapter seven is the concluding chapter summarising the findings in the thesis. It attempts a general analysis of the place JA holds in the prosodic typology. In this chapter a model of JA intonation is summarised in light of the AM theory. It is concluded that JA is a stress-accent language, demonstrating prominent pitch accent faithfully associating and aligning with stressed syllables and showing two intonational levels above the prosodic word: the intermediate phrase and the intonational phrase. Those two intonational levels are found to be marked by both tonal and non-tonal correlates. The chapter also reviews the contributions of the thesis both to cross-linguistic and cross-dialectal prosodic analysis. It then concludes with suggestions for future analysis on the dialect.

Chapter 2. Review of the literature

2.1 Introduction

Intonation is an important aspect of human communication. Intonation among other prosodic properties, has been shown to signal speakers' emotions, attitudes as well as provide linguistic information about the discourse, such as whether the part of speech is a question, statement, request, etc., as well as aid with word segmentation. Intonation is also used to signal turn taking during a conversation, in addition to signalling the beginning and end of the various parts that make up a conversation. It is also used to draw attention and highlight certain aspects to the listener, alongside its well-studied role in the syntactic disambiguation of grammatical material (Cole, 2015, Wagner & Watson, 2010). Intonation achieves those communication goals by employing suprasegmental features.

There is a general agreement regarding the relationship between the suprasegmental feature of fundamental frequency and intonation, as intonation is commonly viewed as the modulation of pitch in spoken language (Arvaniti, 2012). There is a long tradition of associating intonation with suprasegmentals where intonation is usually used to refer to "the systematic use of suprasegmental properties…to mark linguistic information beyond word identity" Cole (2015, p. 2). However, Jun (2005) explains that word lexical features also interact with intonation in such a way that firstly intonational (postlexical) features, such as pitch accents, phrasal and boundary tones all take place in syllables that are word-level components. She adds that it is due to this close interaction that the study of prosody needs to examine both word and sentence level features "postlexical prosody is constrained by lexical prosody, and postlexical prosodic information contains information about the lexical prosody" (Jun, 2005, p. 431, also Hellmuth, 2006). Additionally, other suprasegmental features such as amplitude, voice quality, spectral balance and duration can also be used to mark intonational categories, mark the beginning and end of various intonational groups, and to indicate the relative prominence relations that hold among them (Cole, 2015).

The relationship between lexical levels, intonation and the suprasegmentals is strong as can be noted. Another speculation to be added here concerns the relationship between these suprasegmental and lower level features and segments, i.e. the practical makeup of utterances. Via cross- linguistic examination, it has been proposed that phonological structure mediates between those levels and directs the interaction among them. Universal patterns of variation in pitch, duration, and acoustic cues of segments, were thus taken to reflect those phonological structures. This view of the interaction between the different phonological levels and the surface intonational form constitutes the base on which the Autosegmental-Metrical theory of intonational phonology models in (Pierrehumbert, 1980; Beckman & Pierrehumbert, 1986; Ladd, 2008; Arvaniti, 2017; Chahal, 2001; Hellmuth, 2006a) build on in order to analyse intonational structure.

This thesis aims to present a model for the intonation of Jeddah Arabic (JA) that is grounded in the laboratory phonology tradition of drawing conclusions based on the interaction between phonetic form and phonological structure. The model employs the tenets of Autosegmental-Metrical theory to analyse the intonational structure of this dialect via quantitatively and qualitatively identifying the observed tonal structure.

2.2 The theoretical framework of the Autosegmental- Metrical theory of intonational phonology

A formalisation that attempts to explain and explore the relationship between intonation and suprasegmentals gave rise to the Autosegmental-Metrical models of Intonational Phonology, and the Tones and Break Indices ToBI prosodic transcription notation discussed in the early literature in Pierrehumbert (1980), Beckman & Pierrehumbert (1986), and Ladd (2008), among others. This perspective regards intonation as the phonological association of tones and prosodic constituents, which paves the way for phonetic alignment that involves a timing coordination between segments and tones (Arvaniti, 2012). Studies employing a laboratory phonology approach are in particular concerned with these two notions of phonological association and phonetic alignment; their results and a number of speculations regarding this matter will be reported in the upcoming sections. The model assumes a link between surface accents and underlying phonological structure, as can be implied from the name. The model employs two levels of abstraction: Autosegmental and Metrical, as well as a phonetic component that is centred on the observation of the acoustic information as encoded in the F0, duration, and amplitude of the speech signal.

2.2.1 The Autosegmental element

The autosegmental element of the theory is related to the melody of the spoken utterance. It is the tonal tier where tones are represented as autosegments on an autonomous tier and independent from the segmental tier and representation. In the theory, a melody is composed of a sequence of abstract high and low tonal targets whose phonetic identity is represented by peaks (High) and valleys (Low) in the speaker's pitch range (Arvaniti, 2017), (Chahal, 2001), (Gussenhoven, 2004), (Jun & Fougeron, 2000). The linguistically relevant tones phonologically associate to specified landmarks in the metrical structure, and are of two types: those that mark prominence (pitch accents), and others that mark boundaries (edge tones). The remainder of the contour is unspecified, whereby the pitch transition between those important landmarks is derived by phonetic interpolation rules (Ladd, 2008), (Chahal, 2001). More on interpolation is discussed in the upcoming section 2.4.2.

2.2.2 The Metrical element

The metrical component of the theory uses principles from Metrical Phonology and Prosodic Phonology to account for the underlying representation level (Ladd, 2008). This component concerns the abstract association between text and tune. Cole (2015, pp. 2-3) summarises the notion behind such a proposition as one that accepts that a hierarchically organised phonological structure specifies the locations for the distribution of suprasegmental features (e.g. tones), and influences the timing and magnitude of a phone's articulatory gestures hence marking the contexts for variation in their phonetic realisations. In other words, the theory employs the metrical strength concept that distinguishes syllables based on relative strength/prominence (stressed vs. unstressed) in the lexical level, alongside a constituency hierarchy from prosodic phonology, which links tones on the tonal tier to relevant phonological units in the postlexical level to distinguish the relative prominence of a tonal event (accented vs. unaccented, prominent vs. non prominent)[2]. This indicates that the theory makes use of the hierarchically layered phonological structure (syllable, foot, word, phrase, utterance) to explain Tone- Text association (Pierrehumbert, 1980).

2.2.3 The Prominence hierarchy

The prominence hierarchy reflects the relative prominence relationships the theory assumes to phonologically hold between relevant units, and the observed surface realisation of this prominence. Accents within an utterance either signal lexical-level prominence or post-lexical-level prominence. Lexical level prominence is expressed through lexical stress assigned to prosodic words according to the language's phonological rules. In the post-lexical level prominence is signalled by pitch accents and nuclear accents. These are F0 configurations that encode pragmatic functions of some sort, whereby the composition of a tune takes into account the series of accents within an utterance leading to the boundary tone to compose a meaningful tune. Those add a further degree of prominence in addition to the lexical level prominence - word stress- (Al-Zaidi, 2014), (Hellmuth & Chahal, 2014), (Chahal, 2001) (Chahal, 2003). The hierarchy is ordered whereby the highest prominence is given to the *nuclear accent*: the last pitch accent in a phrase, followed by the *pitch accent* and lastly *lexical stress*. The relationship is presented in the following table:

---

[2] The theory uses prosodic constituency levels that have been reported to be marked by intonation.

Type/rank	Implication/relative prominence
3- Lexical stress	Stressed syllable more prominent than unstressed.
2- Pitch accent	Accented syllable more prominent than unaccented.
1- Nuclear accent	Last pitch accent before phrase boundary is most prominent. More prominent than pre-nuclear (pitch) accented, unaccented, and post-nuclear (pitch) accented (some languages).

Table 1:Prominence Hierarchy as adapted from literature.

Nuclear and pitch accents are prominent because they associate with the *head* of a foot (the metrically stressed syllable). Accordingly, languages that demonstrate a phonotactic constraint of only associating an accent with the stressed syllable are called 'stress-accent' languages (e.g. many Germanic and Arabic languages), while languages that realise pitch accents independently from stress are called 'non-stress accent' languages (e.g. Japanese and many Asian languages) (Ladd, 2008). The hierarchy then ranks accents whereby an accented syllable is more prominent (phonologically and acoustically) than an unaccented syllable, a nuclear accented syllable is more prominent than a prenuclear accent or postnuclear accent- if permitted in the language.

A number of studies have reported that relative prosodic prominence aids in signalling the information structure and meaning of an utterance (Yeou, et al., 2007a), (Face, 2002), among others. This is particularly related to the concept of *Focus* 'highlight or emphasis', which is a type of discourse meaning signalled by prosody. Focus is prosodically realised when a part of speech is assigned the most prominent pitch accent 'a nuclear accent' and thus becomes *focal, narrow- focused, contrastively- focused* or *in-focus* (ibid).

2.2.4 Focus in Autosegmental-metrical theory of intonation

In the AM theory of intonation, the notion of accentuation is viewed through a relative prominence hierarchy (Beckman & Pierrehumbert, 1986), (D'Imperio, 1997b). This hierarchy determines the phonological rank and status of accents within an utterance. Pitch accents occurring within an utterance are subject to the relative prominence hierarchy presented in table [1] above making some more prominent than others.

In this sense, it is said that nuclear accent denotes focus (Xu & Xu, 2005, Chahal, 2001, 2003, Hellmuth, 2011, Ladd, 2008). That is, the focused constituent is the most prominent and important constituent in a string. This high status of a nuclear accented target is signalled by phonetic detail, such as peak height, duration, the realisation of segments, and amplitude as will be shown in the upcoming chapter 5. Languages also report cross- linguistic variation in focus marking showing contrasting qualitative and gradient patterns to mark narrow and broad focus utterances.

Studies on broad and narrow focus (Wang & Xu, 2011, Chahal, 2001, 2003, Chahal & Hellmuth, 2014, Al-Zaidi, 2014, Jun & Fletcher, 2014, among others) generally report that in a broad focus string (neutral sentence) the last pitch accent before the phrasal boundary is promoted as a nuclear accent. This was shown to fall on the final word in the sentence that demonstrates enhanced phonetic detail compared to the pre-string. In a narrow focus sentence, the nuclear accent is placed on the focused constituent regardless of its location (initial-medial-final). Studies report that the differences between the two conditions are also signaled acoustically by changes in duration, intensity and F0 values (height and/or range), tone rise duration and rise speed in the target syllables, all as a way of enhancing the narrow focused target against the whole string (ibid). In the phonology, this realisational difference between the two focus conditions is captured in metrical structure. Ladd (2008) analyses this difference as a metrical strength alternation in the sentences, analogous to the notions of iambic and trochaic stress patterns. A recursive weak- strong (i.e. unaccented- accented) rhythmic alternation is interpreted as broad focus, while a strong- weak alternation encourages a narrow focus interpretation. However, a weak- strong alternation is ambiguous between a broad focus utterance and a narrow focus utterance with final focus.

On this matter, studies have reported that there is indeed minor variation between narrow focus utterances where a target is placed in final position, and broad focus utterances. For example, Chahal (2001) reports no major differences between narrow focus targets in final position and broad focus apart from higher peaks on the target in the former case. In broad focus, on the other hand, a final target has relatively similar pitch level as the pre-nuclear accents. This effect was also reported for English in Xu and Xu (2005). This suggests that in addition to other phonetic and phonological correlates, the relationship between peaks in target utterances is a signal to focus type in these languages. In the theory this is taken to be indicative of a close relationship between focus-related phonetic detail (F0 variations) and focus interpretation, which was also supported in perceptual studies (Among them are Gussenhoven, 1983 on English, Baumann & Winter, 2018 on German and van Donzel & Beinum's, 2000 study on the perception of prominence in Dutch).

2.2.5 The Constituency hierarchy

The aim of the theory is to establish that the pitch contour of a language reflects the phonological association to positions in the prosodic structure. Jun (2005) describes the theory as one that views prosody in terms of prosodic structure and tones. She adds that the theory is concerned with prominence relations within a word in lexical representation and among words in postlexical representation. Thus, prosodic structure is defined above and below the word level. Lexically, prominence is marked by tone, stress accent, or lexical pitch-accent, while postlexically it is marked by heads and edges of prosodic units (that are marked by special pitch configurations). Accordingly, the theory presents two types of tonal events: pitch accents that are phonologically associated to prominent syllables (prominence-lending tones), and edge (boundary) tones which are pitch movements aligned to the edges of prosodic constituents- the prosodically marked phrasing levels- (Hellmuth, 2006a).

Postlexically, the theory recognises a phonological organisational structure that interacts with the tonal tier. The constituents of this structure are considered as domains for a number of phonetic and phonological rules, including segmental and tonal phenomena. They are organised hierarchically whereby higher levels include one or more of the lower level constituents. Each constituent is delimited by a boundary and dictates the prominence strength relationships among the subordinates within. Those boundary/edge demarcating and prominence phenomena have been taken in the literature as empirical evidence for the existence and universality of those prosodic constituents. Note the following figure:

Postlexical level	Intonational phrase (IP)
	Intermediate phrase (ip)
	Accentual phrase (AP)
Lexical level	Prosodic word (PW)
	Foot
	Syllable
	Mora

Figure 1: Intonation- based Prosodic Hierarchy. Adapted from Frota (2012).

A high level of intonationally marked phrases is the Intonational phrase. This phrase is approximately equivalent to a syntactic clause and is obligatorily composed of at least one or more intermediate phrase(s) (Chahal & Hellmuth, 2014). Unless moved elsewhere (subject to focus interpretation), among the ips in this phrase, the last pitch accented word carries the nuclear accent and is followed by a boundary tone demarcating its right edge. This phrase is the domain for silent pause insertion, final lowering and phonetic declination (Chahal, 2001). The edge of an IP is marked by a boundary tone T% that marks the boundary irrespective of the prominence relations among constituents in a phrase (Beckman & Pierrehumbert, 1986). The next level is the intermediate phrase approximately equivalent to an NP, VP, or PP syntactic phrase. This phrase is made up of a number of prosodic words with at least one pitch accent. In this phrase too, the last pitch accent bears the nuclear accent status and is followed by a phrase tone demarcating its right edge. The edge of an ip is marked by a phrase tone T- that also takes place regardless of pitch accent or nuclear accent location (ibid). The ip is the domain for the distribution of accentuation (pitch accents), and the relative prominence relations among them. This phrase is also typically the domain for pitch reset where the pitch range for the phrase is sustained until the phrase tone and then a new pitch range is used for the following phrase. These phenomena mark an ip as well as a less strong sense of disjuncture (compared to an IP) and an optional silent pause (Chahal & Hellmuth, 2014).

Additionally, evidence from the phonetic *alignment* or location of accentual peaks occurring in these two phrasing levels (ip and IP) is reported as evidence for their existence (Ladd, 2008, among others). It has been shown that a peak occurring in the vicinity of an IP boundary is moved further to the left from its edge, more so than a peak occurring in the vicinity of an ip (more on this is discussed in detail in the alignment chapter **6**). These phrasing levels can be marked by both tonal and non-tonal cues. The tonal cues are in the form of tone configurations at the edge of the respective phrase, final F0 lowering and pitch reset, as well as tonal alignment patterns. For instance, the ip was found to be the domain of pitch reset whereby a new ip is marked by a new pitch range, and the domain of downstep whereby the ip boundary blocks downstepping of peaks in the new phrase (Beckman & Pierrehumbert, 1986). On the other hand, the IP was found to be the domain of phonetic declination, which involves the lowering of successive peaks up to an IP boundary- including nuclear accent (ibid). More on those tonal implementation rules are discussed in the phonetic element section 2. The non-tonal cues take the form of pre-boundary lengthening and silent pausing. It is reported that the effect size of these cues is highly correlated with the strength of the boundary, i.e. the stronger the boundary (IP), the longer the pause, the longer the duration of syllable in the boundary vicinity, the greater the tonal changes and tonal alignment effects, and vice versa (Chahal, 2001 among others). Regarding juncture, Beckman and Pierrehumbert (1986) resort to pre-boundary lengthening, as well as abrupt tonal transition as main evidence for the occurrence of an ip edge in the absence of a silent pause.

Finally, the existence of an Accentual phrase (AP) is not extensively reported, though a number of languages make use of this phrase in their intonational systems. The AP is a prosodic constituent lower than ip and above the PW and whose boundary is tonally ‘postlexically’ marked (Jun, 2005). The main characteristic reported for an AP is its invariant tonal pattern, and its juncture strength that is perceptually less prominent than an ip edge (Jun & Fougeron, 2000). This phrase has been shown to exist in Korean, Japanese (Beckman & Pierrehumbert, 1986, Jun, 2005, Gussenhoven, 2004), and French (Jun & Fougeron, 2000). It is more commonly found in languages with rich accent distribution where every prosodic word bears an obligatory pitch accent (or every small group of words in a fast speech rate, as discussed regarding French in Jun & Fougeron, 2000). This dense accentuation meant that in these systems each word has to tonally mark its boundaries, alongside bearing a pitch accent. In addition to marking the edge, it meant that each pitch accent has to be head of some phonological domain, similar to stressed syllables being heads of prosodic words, and nuclear accents being heads of intermediate phrases. Therefore it is reported that pitch accents are heads of Accentual phrases. In light of these discussions it can be sensed that there may exist cross-linguistic differences in the number of postlexical levels reported for each language. While most languages report the existence of an intonational phrase in their systems, not all are reported to have an accentual phrase, for example.

The current thesis adopts the following view from Cole (2015, p. 3) expressing the mechanisms involved in the interaction between prosodic structure and tonal phenomena:

> “With prosody based in phonological structure, its expression in both segmental and suprasegmental properties can be understood as arising through two mechanisms. First, prosodic structure defines the locations where tone features are linked (e.g, at the edge of a phonological word, on a stressed (prominent) syllable within the word, or a phrase-final syllable), giving rise to the pitch contours that carry lexical, grammatical or pragmatic meaning. Second, prosodic structure influences the timing and magnitude of articulatory gestures for consonants and vowels…gestures are lengthened and strengthened in certain prosodic positions…while they are shortened and reduced in other positions”

This quote expresses the role prosodic structure plays in the analysis of intonational structure. It has been reported in many studies that there is indeed empirical articulatory and acoustic phonetic implementation evidence for the marking of prosodic structure (Beckman & Edwards, 1990). For example, Chahal (2001) reports pre-boundary durational lengthening of syllables before an intonational phrase boundary, and less lengthening before intermediate phrase boundary. Wagner and Watson (2010) report the degree of pause duration and juncture was shown to be different in strength between IP and ip boundaries (IP has longer pauses and stronger degree of juncture than ip). Also, Hellmuth (2006a) reports declination of F0 level among accents to be delimited by an IP boundary and affects all accents in an IP domain until the boundary tone. Moreover, as discussed in Cole (2015) the more complex a tonal configuration is before a boundary, the stronger the boundary. This concept constitutes the difference in boundary marking notation in AM models (see the notation in Beckman & Hirschberg, 1994), whereby an IP boundary is a combination of a phrase tone and a boundary tone (e.g., L- L%) that is reflected in the acoustic signal as two movements before a pause, whereas a simple monotonous configuration (e.g., L-) before an ip boundary.

2.3 Tune to text association: Phonological association and phonetic alignment

Of the basic tenets of AM theory are the notions of phonological association and phonetic pitch alignment. As Hellmuth (2006a, p. 94) adds: "[…] AM theory, it is generally assumed that the surface alignment of the pitch contour is a reliable indication of the underlying phonological association of tones to prosodic targets". The underlying motive behind these notions in AM was to uncover the relationship between text and tune (Arvaniti, 2012). In the theory, tones associate with rhythmically strong positions in the metrical structure (prominent and boundary-adjacent syllables in mainly stressed feet). It follows then, those elements (tones and sonorant segments) that are phonologically associated need to be temporally synchronised (aligned) as well (ibid). Alignment patterns are expected to therefore demonstrate a temporal coordination between tones and segmental structure. The principles governing this coordination generated much research in the realms of both phonetics and phonology. The main goal was to uncover how much of the variation of alignment patterns are due to phonetic implementation and how much is due to phonological representation. As tonal alignment demonstrates the temporal alignment of fundamental frequency F0 with respect to the segmental string as observed in the F0 contour, it is thus considered as the phonetic/acoustic component of the theory of intonation, a concept that is crucial in any analysis of intonation as discussed in the following (Prieto, 2011).

2.3.1 Models of tonal alignment

A number of models have attempted to express both the association between autosegmental information and metrical structure, as well as the surface coordination between autosegments, the metrical structure and the segmental information. In AM, these units and pitch movements are strictly related, whereby a pitch movement is *anchored* to specific segments- onsets and offsets- of a phonological *tone-bearing unit* (TBU, such as stressed syllables) (Ladd et al., 1999); that is, within the confines of the TBU. Moreover, the duration of the pitch movement and location is expected to be highly correlated with the duration of the TBU. This constitutes the concept of a "starred" and "unstarred" tone in bitonal accents in earlier accounts of AM (Pierrrehumbert, 1980). A starred tone is phonologically associated and aligned with the accented syllable, and differences in alignment would entail differences in phonological category. For example, an L+H* tone is phonologically distinct from an L*+H in that the starred tone would occur in time with the accented syllable and the unstarred tone would lead or trail.

However, it is well known that phonetic factors, such as: tonal crowding, speech rate, segmental composition, syllable structure, affect how a tone is aligned with the phonological units. Alignment modification has been observed in many languages whereby a peak is located earlier/ later than expected relative to its phonological unit (D'Imperio, 1997a, among others). Hence, the location of the target accented syllable has its effects on tonal alignment. These effects of location or 'time-dependent' re-adjustments are recognised as stress (accent) clash, tonal crowding and prosodic boundary effects in the literature. Stress clash (which refers to accent clash and not word-level stress) is defined as the occurrence of successive pitch accents across words, but within the same phrase (IP, ip) whereby accents following each other by zero syllables are aligned earlier than when followed by an increasing number of syllables (Prieto, 2011, Chahal, 2001). This behaviour is also accompanied by lengthening the first accented syllable in some languages (Prieto et al., 1995 for Mexican Spanish, Frota, 2012 for European Portuguese). Tonal crowding is defined as the occurrence of both word stress and pitch accent on a phrase-final syllable in finally stressed syllables, interfering with a boundary tone marking an IP. Studies report that in this case a re-alignment or re-adjusment of the peak takes place, as opposed to a non-final syllable (D'Imperio, 2012). As such, proximity to a prosodic boundary (right-edge boundaries) has been reported to affect peak alignment and location exerting a leftward push on the tone cross-linguistically. That is, a tone occurring in the vicinity of a strong prosodic boundary (IP) is reported to be placed earlier than it occurring in the vicinity of (ip) in a regressive manner (Wichmann et al., 2000, Chahal & Hellmuth, 2014, Prieto et al., 1995). This tonal alignment behaviour is also provided as evidence for the existence of these prosodic constituents in the languages under investigation. The effects of segmental and syllabic composition are discussed thoroughly in section 0. Speech rate differences have also been shown in some studies to affect peak placement- although the study in Ladd et al. (1999) does not report such an effect of speech rate on tonal alignment.

While AM treats differences in alignment as differences in phonological category and association through the "starred" and "unstarred" concept, it still does not explain the within-language and cross-linguistic variation in segmental anchoring where the same category is aligned only slightly differently with respect to the segmental string, and sometimes surpassing the designated TBU. Therefore, there have been phonetic attempts to explain the tonal alignment behavior taking into consideration the phonetic conditioning factors in a unified model (Atterer & Ladd, 2004, Ladd at al., 1999). The model named- *Segmental Anchoring Approach hypothesis* (SAH)- maintains that while association is fixed to the TBU, tones within a bitonal accent are free to align (be anchored) with any of the segments in the TBU (within the TBU or in relation to its boundaries) taking into account the phonetic factors (Atterer & Ladd, 2004). For example, Ladd et al. (1999) found one of the cases in their study where the speaker varied the location of the peak according to speech rate, while in all the other cases, the alignment of L and H were constant with respect to the TBU boundary. That for them was not enough evidence to show that this speaker was consciously producing a different/contrastive accent type (thus not a different accent in the phonological representation).[3] Therefore, this approach regards non-contrastive alignment patterns as due to language-specific rules (and inter-speaker phonetic variability as mentioned in Atterer & Ladd, 2004).

Other attempts to define the anchor points for bitonals that aim to explain the surface variation in alignment- or absence thereof- have been discussed in the literature. There are the *Constant duration* and *Constant slope* models. Constant slope is motivated by Ashby (1978). Constant duration was motivated by Fujisaki (1983) to explain Japanese prosody. In the former, individual tones in rising accents follow each other at a constant, predictable duration, resulting in accents with fixed durations that are freely aligned with respect to the accented syllable boundary. The latter approach, somewhat similar, regards accents to have fixed slopes correlating with their durations. In this approach, accents are thought to be aiming at a constant slope, shape or movement that is independent of where individual tones are aligned with respect to the accented syllable. The approach advocates a constant shape for the accent that remains constant under different segmental contexts. Such approaches would essentially mean that a given language would contain numerous accent categories according to their duration and

[3] They explain that this one speaker produced utterances with an extremely low speech rate.

according to how the phonetic factors influence these durations, and would fail to explain cases in languages where there are two accents with the same shape but that differ in how the individual tones align with the TBU (e.g. L*+H vs. L+H*). As a consequence, these approaches seem the least economical with a possibility of leading to *overspecification* of F0 configurations.

Another view of alignment patterns is the *Structural Anchoring Hypothesis* mentioned in Hellmuth (2006a) to account for tonal alignment of Egyptian peaks. This alignment pattern she contends is structural, whereby tone targets in rising accents align to a stressed foot, in particular the second mora of the foot- be it a second vowel or a coda. This hypothesis maintains that tones align with reference to edges-rather than heads-[4] of TBUs. Hellmuth adds- following this hypothesis- that the surface alignment patterns would differ according to syllable structure, foot duration and word duration. Her study reports that in Egyptian, while the L is consistently aligned with the onset of the syllable, the H in CV syllables is located outside the syllable but inside the foot in bisyllabic feet. In CVC and CVV syllables (in monosyllabic feet) it is located in the rhyme (coda or second vowel) (p. 240). In her data this shows as consistent alignment of H to the second mora in a foot. She therefore concludes that the foot rather than the syllable is the TBU in Egyptian. Indeed, on the other hand, if the peak location is insensitive to the syllable type, hence number of moras, we would not expect drastic surface variation, and would be able to clearly define a segmental anchor for the H tone within the target syllable.

The alignment patterns reveal a great deal about the language's true tonal behaviour and about the cross-linguistic and dialectal variation. It also somewhat indicates that the timing properties of segments and syllables influence/direct tonal alignment. This has also been shown in perceptual studies to considerably affect lexical and sentence meaning and identification (Gussenhoven, 2004, D'Imperio, 2012). Arvaniti (2012) - an advocate for AM- argues that this variation in alignment is consistent with the theory's notion of *underspecified* tonal targets that can be freely associated with their choice of strong metrical or boundary constituents and temporally synchronised relative to any of their various segmental events. Nevertheless, the mechanisms underlying those choices must be clearly stated and further studied in AM models. In a similar vein, Prieto (1995, 2011) argues that there are changes in alignment that provoke changes in meaning and are due to phonological effects/categories/representation. However, alignment patterns brought about by the phonetic influences such as tonal modification, speech

[4] In bimoraic systems, edges are the second 'weak' moras in a foot, while heads are the first 'strong' moras (Hellmuth, 2006).

rate or segmental and syllabic structure mentioned above are generally due to differences in language-specific phonetic implementation rules. If these do not induce differences in meaning, it follows that they are not phonological or categorical. Therefore studying the phonetic factors affecting tonal alignment (stress clash conditions for example) helps to show how unique phonetic implementation is to each language and variety. In the current thesis, this aspect would potentially shed light on the implementation rules in this Arabic variety.

2.4 Phonetic element: Phonetic implementation

The theory embraces a number of tonal implementation rules that transform underlying tones into F0 values in the acoustic signal. These rules affect the level Scaling of the tone, as well as the transition between tones in a sequence 'Interpolation'. Those implementation rules also show both universal and language-specific patterns.

2.4.1 Scaling of tones

Scaling affects the level of a particular tone with reference to the pitch range in the language. There are four reported factors that affect pitch range: Upstep, Downstep, Final lowering and Declination. As Pierrehumbert (1980) and Ladd (2008) assert, F0 values are scaled relative to the bottom of the speaker's pitch range: the baseline.

Declination is a rule that triggers a gradual lowering of pitch range throughout an intonational phrase (Pierrehumbert, 1980). This affects both the peaks H and valleys L of accents (Hellmuth, 2006a). Upstep is a Tonal spreading rule that affects High boundary tones following High phrase accents in an Intonational phrase. This typically occurs in boundary combinations of the type H-L% and H-H%. In H-L%, the L% boundary tone is locally raised to a level similar to the H- phrase accent preceding it, which is essentially higher than what would occur if the boundary was preceded by an L- phrase tone in L-L% boundaries for example. This results in plateau tunes, where the L% never falls to a true low F0 value. In H-H% combinations, the H% is considerably scaled to a level even higher than the high phrase accent preceeding it, which is in turn higher than an H% occurring after an L- phrase accent in L-H% (Pierrehumbert, 1980, Chahal, 2001). This boundary combination results in high rising question tunes (detailed illustrations are presented in chapter 4). In contrast to upstep, downstep is a rule compressing the pitch range of peak H* that lowers their F0 value !H* compared to preceding peaks. In Pierrehumbert (1980) this is a purely phonetic rule triggered by a preceding bitonal accent of the type H+L where subsequent peaks are downstepped to the same or lower level than this pitch accent. However, as suggested in Ladd (2008) and Chahal (2001) downstep can be a conscious phonological choice not triggered by any phonetic factors and can affect edge tones as well as pitch accents. This is shown in Chahal (2001) as downstepped tones give rise to stylised tunes with specific meanings 'Plateau tunes'. Downstepped accents have an even lower F0 value than those that are affected by declination, and once a downstepped accent occurs in an ip, all subsequent accents within the same domain are downstepped as well (ibid). Final lowering is a marked rule affecting the last downstepped peak, which is further lowered in value than a previous downstepped accent (Hellmuth, 2006a, Chahal, 2001).

2.4.2 Phonetic interpolation

Phonetic interpolation is concerned with the transition in between accents in a contour. This rule makes reference to the value (level) of the two accents, as well the time between them. The resulting transition is implemented as a ratio between the two values (Pierrehumbert, 1980, Ladd, 2008). Thus, interpolation is considered to be monotonic and linear between all L and H targets and with reference to the trajectory between the two tonal events (Arvaniti, 2017). As a result of many experimental studies, interpolation between two H* peaks on the other hand is said to be 'Sagging' with sagginess enhancing as a result of further temporal distance between the two (ibid). Note the sagging transition between peaks in the following illustration:

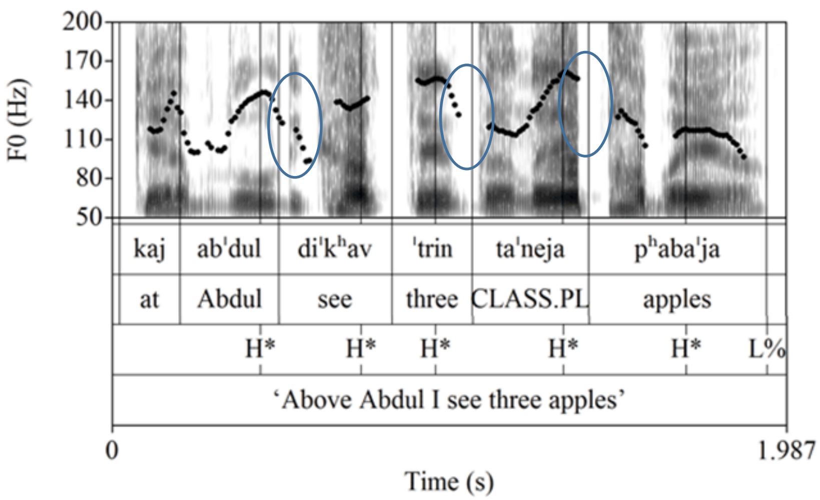


Figure 2: Sagging interpolation between H* accents (blue circles indicate transitions). From Arvaniti (2016)

That is, the theory maintains that transition between L and H is linear: rising from low, transition from L to L is also linear: level and low, while transition from H to H is sagging: dip then rise for the next H (Chahal, 2001). Additionally, phonetic evidence shows that the larger the space/distance between two peaks (H accents), the shallower the sag and vice versa (ibid). The theory also recognises one exception to the sagging interpolation in what is known by a "flat hat" transition, whereby pitch between two H accents remains high but level and showing no dip in the contour. This is taken as a result of phonetic pitch range manipulation.

2.5 Arabic studies on prosody and prominence

2.5.1 Stress

Prosody has long been regarded as a structural and hierarchical tool for organising utterances (Beckman & Edwards, 1994). It is regarded as the means of chunking utterances (words or phrases) into units, generally by reference to a number of properties. Stress is one of the properties for expressing the rhythmic structure or alternation- in languages that possess them- holding between the syllables in a lexical word. As was mentioned earlier in the thesis, the classification of a language as stress-accent or non-stress accent as related to intonation is influenced by whether it associates its intonational accents with lexically stressed prosodic constituents (Beckman, 1986, Chahal, 2001). In stress-accent languages stressed syllables are associated with prominent pitch movements whose distribution in phrases reflects the rhythmicity in the intonational level.

Descriptions of Arabic seem to agree that stress is an essential part of the language's prosody. Arabic dialects are reported to be stress languages and quantity sensitive languages, where stress is generally predictable depending on the syllabic structure of words, for example, the weight and position of the stress-bearing syllables (Broselow et al., 1997, Chahal, 2001, Hellmuth, 2006a, Watson, 2002, De Jong & Zawaydeh, 1999, Al-Zaidi, 2014). Thus in Arabic, heavy CVV, CVC and superheavy CVVC, CVCC typically attract stress by default. In terms of stress systems, it has been established in the literature that Arabic exhibits bounded feet- feet maximally contain two syllables/moras - (McCarthy & Prince, 1990, Watson, 2002, among others). Hayes (1995) argues that the stress pattern derived by bounded feet is captured by the law of rhythmic structure, called the *Iambic/Trochaic law* (for a discussion of this law please refer to Watson 2002, pp. 86-87). The foot inventory constructed under this law is: syllabic trochee, iamb, and moraic trochee. The concept of 'heaviness' plays a crucial role in stress assignment in the moraic trochee and iambic languages, whereby feet are constructed on moras. The differences in the surface realisations of stress in Arabic dialects are explained by the notions of extrametricality and direction of foot construction (Hayes, 1995).

In metrical theory, stress patterns are regarded as reflexes of rhythmic structure with rhythmic beats as headed feet (Hayes, 1995, Hulst van der, 2014b, Beckman & Edwards, 1994). A development of this foot-based metrical theory of stress is that of Hayes (1995). He emphasises that the metrical units of 'feet' constitute the basic elements in the expression of stress systems. He also adds mechanisms of 'moraic theory' for the purpose of distinguishing light vs. heavy syllables in stress systems that rely on this weight notion to construct and group their feet. The following is a summary of the stress mechanism in metrical theory:

Word stress parameters
Foot formation
Feet are left-headed/right-headed
Feet are assigned from right-to-left/ left-to-right
Feet are bounded/ unbounded
Word formation
Feet are grouped into a left-heded/ right-headed word tree
Extrametricality
The final syllable is ignored (yes/no)
Weight-sensitivity
A syllable with internal weight must be a head (yes/no)

The Stress mechanism (as adapted from Hulst van der, 2014a, p. 17)

The summary provides information on the interaction between feet and syllables. Feet are constructed on syllables and are in some systems sensitive to their internal structure and weight. Hayes' (1995) theory contends that stress rules are conditioned by the properties of these syllables, regarding the syllable as the stress-bearing unit. Furthermore, the grouping of the syllables is subsequently used to construct a foot of one of the types: syllabic trochee, iamb, or moraic trochee.

As for the mechanism formalising moraic distribution and structure, in systems that employ them, can be sought in Moraic theory (Hyman, 1985, Hayes, 1995, Watson, 2002, McCarthy and Prince, 1990 and others). In this theory syllable weight is represented by moras μ, which are also the basis for the representation of metrical structure. Onsets are non-moraic, light syllables and geminates are allocated one mora, and heavy syllables are allocated two. Long vowels are assigned two moras by default, whereas non-final heavy CVC syllables assign a mora to the final consonant by the *Weight-by-Position Condition* (WBP) to make this syllable type visible to stress rules. Concerning the number of moras a syllable may contain, it is reported that Arabic dialects typically set an upper limit of two moras for each syllable. Al-Mohanna (2009) states that syllable bimoraicity is ranked undominated in a considerable number of Arabic dialects. This constraint is captured in Broselow's (1992, p. 10):

Bimoraicity constraints:

Syllables are maximally and optimality bimoraic.

Mccarthy and Prince (1990) maintain that reference to the mora is inevitable in the description of stress in Arabic and provide arguments supporting the claim of its importance. In their view, it is due to the reference to moras that the last consonant in 'final' CVC syllables in Arabic is considered light and skipped by stress, i.e. 'extrametrical' and 'non-moraic', because it is a weightless onset when followed by vowel-initial suffixes (note that onsets are weightless in Arabic). Also, Broselow et al. (1997) maintain that final CVCs pattern with light CVs in response to a constraint prohibiting final consonants to occupy a mora (by WBP), thus adding unwanted weight to a final syllable. Similarly, the final consonant in superheavy syllables (CVVC, CVCC), is either regarded as an extrametrical syllable (McCarthy & Prince, 1990); a degenerate syllable for being 'extrasyllabic: not belonging to the syllable node' (Watson, 2002), or an unsyllabified extrametrical consonant (Hayes, 1995). When these superheavy syllables occur non-finally, the last segment shares a mora with the immediately preceding segment (the long vowel in nonfinal CVVC syllables). This 'mora-sharing' notion prohibits the final consonant from occupying a mora, thus making the syllable 'trimoraic' which is unfavourable by Arabic and some other languages. This notion was evident through the timing patterns of these syllables analysed and discussed by Broselow et al. and reported subsequently.

Phonetic studies report moraic structure to be reflected in the surface segmental duration, i.e. the phonetic timing patterns in Arabic and other languages (Broselow et al., 1997, Khattab & Al-Tamimi, 2014). Broselow et al. provide cross linguistic, as well as empirical evidence that segments occupying two moras are durationally longer in production than those occupying one, and those occupying one mora are longer than those sharing a mora. These conclusions may suggest that since the mora has an effect on timing, it would have an effect on the realisation of prominence. An example for this is Japanese, a reportedly mora-timed language and accordingly a lexical-pitch accent language (accent on the second mora in word level) with pitch accents functioning postlexically only to mark boundaries around prosodic constituents (Jun, 2005, Venditti, 2005). One can also observe similar findings in Arabic where a mora played a significant role in tone realisation. An example is in Hellmuth (2006a) who finds that in Egyptian (a reportedly stress-timed language) where every word bears a pitch accent, the pitch peak aligns with the second mora of a stressed foot. However, Jun (2005) states that AM models and subsequently ToBI notations do not contain clear links between timing units smaller than the word and intonation. She suggests that it is hard to say how a rhythm unit influences the prosodic units above the word that are essential for intonation expression in a linguistic system, as a language can demonstrate an intonational phrase IP whether it is mora-timed, syllable-timed or stress-timed. The examples above both from Japanese and Egyptian further support this observation as both languages contain IPs in their intonation structure. At this stage, it can be said that the differences in the function of pitch relative to the timing unit and how it is reflected in prominence realisation are likely to be due to a structural variation among languages. Hence, the acoustic correlates they show for tones in the surface are reflective of language- specific structure. However, a thorough description of why and how rhythmic units are involved is indirectly related to the scope of an intonation-based analysis. Though it must be stated that rhythmic alternation is present in intonation and is expressed via the sequencing of prominent and non-prominent prosodic units larger than the word. Accordingly, prosodic phrases may begin and end in prominent units, demonstrate an alternating sequence of accented and unaccented words, and lengthening may be observed at the ends of those rhythmic units (Jun, 2005).

All in all, studies on Arabic report that with the consideration of a few exceptions, the language is typically demonstrates a moraic trochee stress system, where feet are constructed with reference to the heaviness of the syllables in terms of moras. Stress is applied cyclically every two moras in Arabic and syllable types predict stress location. Bearing in mind that the dialects still show variation in the underlying foot application that contributes to some of the surface differences. These conclusions are based on evidence from varieties of Egyptian Arabic (Hayes, 1995, Hellmuth, 2006a, Watson, 2002), Yemeni Arabic (Watson, 2002), Bedouin Hijazi Arabic (Al-Mohanna, 2009), Lebanese Arabic (Chahal, 2001), Jordanian Arabic (De Jong & Zawaydeh, 1999, 2002) and Palestinian Arabic (Hayes, 1995).

According to Hayes (1995), Hellmuth (2013) the variation in stress patterns among the dialects relates to a number of properties that include foot type, the direction of construction of feet in a word, differences in units subject to extrametricality in the stress application process (segment, foot or syllable), and finally in terms of the syllables that are left out in the process due to the inability to be grouped into the preferred foot type. In the following is a table summarising some of the Arabic dialects' stress derivation rules in relation to the aforementioned properties in Hayes (1995) as referenced by Hellmuth (2013, p. 13):

Dialect	**Foot type**	**Foot construction**	**Extrametricality**	**Degenerate feet**
Classical	unbounded	left-headed	consonant	N/A
Bani Hassan[5]	moraic trochee	left-to-right	foot	permitted
Palestinian	moraic trochee	left-to-right	foot	absolute ban
Cairene	moraic trochee	left-to-right	consonant	absolute ban
Lebanese	moraic trochee	right-to-left	syllable	absolute ban
Bedouin Hijazi	moraic trochee	right-to-left	syllable	absolute ban

[5] Bani Hassan Arabic is a Bedouin dialect spoken in northern Jordan (Irshied & Kenstowicz 1984).

Negev Bedouin	iamb	left-to-right	foot	permitted
Cyrenaican Bedouin	iamb	left-to-right	foot	absolute ban
Jeddah Arabic [6]	moraic trochee	left-to-right	consonant and foot	absolute ban

Table 2: Foot construction in some Arabic stress systems from Hellmuth (2013, p.13).

Following the criteria in Hayes (1995), Watson (2002) and Hellmuth (2013), here is a presentation of the stress derivation rules in Jeddah Arabic. It can be noticed that the dialect in Jeddah also assigns primary stresses according to the moraic weight of the feet.

(1) JA word-stress algorithm:

a. Stress a final CVVC/CVCC, or CVV:

/ʒii.ˈbii(h)/ ‘you fem. bring it’ */ʕi.ˈtaab/* ‘reproach’ */ ʃa.ˈbakt /*‘I connected’

b. Otherwise, stress a non-final CVC, CVV, CVVC:

/ˈraab.ʕa/ ‘fourth fem.’ */ˈmas.ʒid/* ‘mosque’ */za.ˈbaa.di/* ‘yogurt’ */ˈgaa.lu/* ‘they said’ */ˈmak.ta.ba/* ‘Library’

c. Otherwise, stress the first light syllable:

/ˈku.tu.bi/ ‘my books’ */ˈba.la.du/* ‘his country’ */ˈxa.ṭab/* ‘he has proposed (for marriage)’ */ˈʃur.ṭa/* ‘Police’ */ˈsa.ma.ka/* ‘Fish, Singular’

d. In three/four light-syllable words, stress the penult:

/dˤa.ra.’ba.tu/ ‘she hit it/him’ */ʔa.ka.’la.tu/* ‘she ate it’ */ma.sa.`Ha.tu/* ‘she wiped it’

Based on those patterns, I now formulate the rules for JA stress in the following that situates Jeddah Arabic in the stress patterns in table [2] above:

(2) JA Stress Rule (adapted from Hayes stress rule, 1995):

a. **Consonant Extrametricality** C ➔ (C) / _____]WORD

b. **Foot Construction** Form moraic trochees from left to right.
Degenerate feet are forbidden.

c. **Foot Extrametricality** Foot ➔ (Foot) / ____]WORD

d. **Word Layer Construction** End Rule Right (ERR): assign stress to the head of the rightmost visible foot/or the only visible foot.

I propose that JA has a moraic trochaic stress system constructing feet on moras. Foot parsing operates from left to right. The dialect shows aspects of extrametricality on both levels: the consonant level, and the foot level. Degenerate feet are absolutely banned as a left over light syllable cannot form a foot and remains unfooted. Extrametricality operates when elements that

[6] Jeddah Arabic patterns are added in the table by the author of this thesis for comparison and illustration purposes.

fall outside the permitted foot size are rendered invisible to rule application, and *Labeling Rules*: which foot is assigned main stress (Hayes, 1995). In JA both consonants and feet can be rendered extrametrical. Foot extrametricality is observed in word such as */mak((taba))/* with initially stressed heavy CVC syllables. In this word two feet are created; one on bimoraic (*mak*) and the second foot is marked extrametrical (doubled parentheses thus invisible). Consonant extrametricality is observed in Arabic when the last consonant in 'final' CVC syllables is considred as light and skipped by stress because it's non-moraic and because it becomes a non-moraic onset if followed by vowel-initial suffixes, therefore makes it extrametrical (Mahfoudhi 2005, McCarthy & Prince, among others). The word */'mas.ʒid/* is an example for this case. Also in superheavy syllables, the final consonant is regarded as an extrametrical syllable (McCarthy & Prince, 1990a), a degenerate syllable for being (extrasyllabic) (Watson, 2002), or an unsyllabified extrametrical consonant (Hayes, 1995)- the latter definition is followed here. The core difference between final C in heavy syllables and superheavy syllables is that the former syllables acknowledge this C as part of the syllable node and because of it being extrametrical, it cannot be assigned a mora through Weight By Position condition, and thus the whole syllable is regarded light. Superheavy syllables, however, do not acknowledge the final C as part of their syllable, thus by treating it as extra it intervenes between the final 'heavy' syllable and the edge of the word rendering this syllable visible to stress (Watson, 2002, pp.90-93). Finally, in words with a number of light syllables such as the examples in (d) above, stress is on a light penult. The words show suffixation of the vowel-initial object suffix */-u/* to the third person feminine singular words */ˋʔakalat/* 'she ate', */ˋdˤarabat/* 'she hit' and */ˋmasaHat/* 'she wiped'. Following the usual foot construction direction from left to right in JA, the words will each contain two left-headed feet, e.g. (*ʔa.ka)(la.tu),* foot extrametricality would render *(latu)* as extrametrical and ERR would incorrectly stress the only visible foot (*ʔa.ka)*. Watson suggests reversing the direction of stress assignment for Cairene to right-to-left in similar cases of stressed light penults (p.97). I follow this for JA and consequently, foot parse operates from right-to-left forming left headed feet protecting *(latu)* from extrametricality, then ERR assigns stress to the rightmost penult.

2.5.2 Cross- linguistic intonational structures and Arabic intonation

In this section is a review of some of the cross- linguistic intonation systems and Arabic intonation systems. Prior to reviewing the intonation structure of Arabic intonation, it may be relevant to discuss some cross- linguistic variations regarding prosodic structures as observed in the literature by Jun (2005), Jun and Fletcher (2014), Arvaniti (2016), Hellmuth (2019) and

Wagner and Watson (2010). According to those publications, typological variation is observed relative to a number of prosodic aspects, although some aspects are regarded universal. In anticipation of the discussion in chapter [4], the universal trends have been observed mainly regarding melodies and meaning. Among those melodies is the intonational expression of questions or interrogatives for example. The cross- linguistic evidence presented points towards the fact that most languages express polar question tunes via the use of a rising pitch configuration observed on the final syllable in the sentence. Variation in this case would be regarding the language- specific pitch range of the observed contour, and the specific boundary combination used by the language according to its tonal inventory (ibid).

Whilst acknowledging such universalities, the research proceeds to explore the other parameters that are used to mark prosodic variation. Among those is the function of a pitch event in a linguistic system, the tonal inventory size, the number of prosodic phrasing constituents and prominence distribution therein, the phonetic correlates used to mark prominence, the variation in the prosodic marking of focus and information structure and variation in the alignment patterns of tones with prosodic structure (Hellmuth, 2019). In this regard languages are said to use a pitch event to mark lexical contrasts (Japanese, Chinese, Yoruba), or mark postlexical structure (most West Germanic languages). In addition to this, a pitch event may be prominence- lending (the pitch accent in English, German), edge marking (Japanese, French), or both as in most languages with Intonational phrases. Regarding phrasing levels above the word, languages may choose to group a particular number of words/ syllables in APs with a marked edge configuration (French, Japanese), or according to syntactic function in an ip or an IP with varying edge configurations and boundary strength phenomena (Lebanese, Egyptian, English, Italian) (Jun, 2005, Jun & Fletcher, 2014, Arvaniti, 2016, Hellmuth, 2019, Wagner & Watson, 2010). Related to that, languages may choose to distribute their accents on every word in a domain showing a dense accent distribution (Egyptian,), or show a sparse distribution of accents only marking the nuclear constituent in a specified domain while deaccenting or compressing other pitch events in the same domain (Lebanese, Germanic languages). Hence showing cross- linguistic variation in the domain of accent distribution. The tonal inventory size of a language may be small with gradient manipulation according to pitch range (1 accent shape in Egyptian, and a constant tonal pattern in Japanese), or big with categorical/ phonologised variation according to pitch range (5-6 accent types in Lebanese, English, and Greek). Regarding phonetic correlates that mark prominence, the languages vary according to the number of correlates they use to mark prominence. Languages may use a combination of stress, accent, duration, segment realisation, and amplitude to lend prominence to words, or choose two or three of those. Similarly, languages may choose to mark edges with tonal (tone height, complex F0 configuration/ segment realisation), non-tonal configurations (duration, lengthening, juncture and pausing) or both. As for the variation in the marking of prosodic focus, languages may choose to mark new/focused items with prominent accents while realising the rest of the old/given information in the utterance with compressed pitch accents- or sometimes deaccnted all together. Other languages like Egyptian as reported in Hellmuth (2006) for example clearly resists deaccentuation of old/given information and consequently every word in the phrase is realised with a pitch accent. Finally, there are also instrumental reports on the location of a specified tone with regard to the Tone Bearing Unit in a language. As was discussed earlier in the models on alignment section, there is wide consensus that cross- linguistically, the syllable is the tone bearing unit, and prosodic variation is related to how the tone is aligned relative to this unit. According to the tonal implementation rules in the language as observed via the different experimental manipulations, some systems align it within the confines of the syllable, relative to syllable and syllabic composition, or outside the TBU but within a consistent distance from it (ibid).

In light of those typological differences, it is suggested that Arabic is an intonational language, where tone plays no role in differentiating words, but has a major role in postlexical representation with reference to a metrical structure for the studied dialects (Hellmuth, 2006a, Chahal, 1999, Jun, 2005, Hellmuth & Chahal, 2014, Hellmuth, 2014 among others). In her description of Lebanese Arabic intonation that adopts the AM theoretical framework, Chahal (2001, 1999) states that the dialect displays tonal events that lend prominence (6 pitch accents: H*, L*, L+H*, !H*, L+!H*, H+!H*) to certain stressed syllables and others that mark phrase edges (5 edge tones: L-, H-, !H-, L%, H%). In regards to prominence lending tones, they associate with metrically strong (stressed) syllables, upgrading them as prominent for conversational reasons. Regarding pitch movements marking edge tones, it displays tonal configurations showing two phonologically and tonally-marked levels of intonational phrasing: the intonational phrase (IP) and the intermediate phrase (ip). The 6 pitch accent types and 4 edge tones inventories are categorically (phonologically) contrastive and the combinations yield pragmatically different meanings and tunes. Evidence for the two levels of phrasing are found according to the tonal configurations and excursions at the edges of these phrases, in addition to further phonetic cues (E.g. pre-boundary lengthening). The syllables in words following the accented syllable are de-accented: realised with extremely compressed or no pitch accent, ruling out the Word as the domain of pitch accent distribution, as opposed to Egyptian Arabic discussed in the upcoming paragraph. The tonal implementation rules observed in Lebanese are Upstep, phonological Downstep, interpolation and declination. She concludes by stating that the dialect is an intonational, stress-accent language, similar to English, which displays tonal (F0) and non-tonal (duration, amplitude) cues to prominence (phrasal and lexical).

In her (2014) analysis of Sanaani Arabic intonation, Hellmuth reports that this dialect spoken in the capital city of Yemen, demonstrates 7 accents of the shapes: L*, H*, L+H*, L*+H, H+L*, H*+L, as well as tentatively proposing an LH*L accent. The edge tones inventory is comprised of 4 tones of the shapes: L-, H-, L% and H%. The study is conducted on a sample of two females and one male and includes analysis of intonational patterns in read and conversation styles mainly dedicated to the analysis of tunes and pitch accent shapes. Polar question tunes show the choice of the L+H* nuclear accent followed by a L% boundary tone, while declarative tunes show a falling F0 contour marked as *H L-L%. There is a brief discussion regarding the prosodic phrasing levels Sanaani demonstrates, and it is concluded that there is evidence for intermediate (ip) and intonational phrases (IP). The evidence is primarily regarding the pitch range reset at the start of new phrases, while no discussion is included regarding the mapping of the prosodic phrases with syntactic structure or the other tonal and non-tonal correlates of those constituents. Regarding the distribution of accents, the author observes that in the Sanaani corpus of read speech, every content word is pitch accented, while in narratives a sparser accentuation is observed. Also, that old/given information is de-accented similar to Lebanese for example. Unfortunately, the analysis in this study does not benefit from the different formal experimental designs employed to observe the true nature of intonational patterns. The analysis hence does not include carefully constructed/controlled focus and tonal alignment experiments that eliminate the different confounding factors, nor does it include the analysis of the tonal implementation rules or phonetic correlates. This restricts Sanaani's comparability with other dialects in aspects beyond surface generlisations.

For Egyptian Arabic analysis conducted under the AM theory, Hellmuth (2006a, 2014) presents evidence showing that it is both similar and different from Lebanese Arabic. In terms of intonational phrasing, the dialect exhibits IPs and MaPs (Major phonological phrase equivalent to an ip). As for pitch accents, focused and given words are realised with one pitch accent of the shape L+H*. Hellmuth (2006a) argues that this rich distribution is to mark prominence within the prosodic word, in addition to the other cues of duration and intensity. Egyptian is similar in this respect to Tunisian and Jordanian in Bouchhioua (2008) and De Jong and Zawaydeh (1999) who report that pitch is a correlate of lexical prominence. The dialect is analysed as possessing only one pitch accent type, as opposed to six in Lebanese. As for the edge tones, Hellmuth states that the dialect observes the two boundary tones H% and L% marking the right edge of IPs, in addition to a combination of phrase and boundary tones marking ips. The tonal implementation rules reported in that analysis for Egyptian Arabic also include Upstep, Downstep, interpolation and declination.

Those previous studies on Arabic (Hellmuth, 2006a, 2014, Chahal, 2001) confirm Arabic as a stress-accent language with pitch used in the postlexical level to denote pitch accent. The intricate phonetic details and correlates used in the different prosodic phenomena studied in those analyses are reviewed accordingly in the respective chapters in this thesis. The results of these experimental studies were taken as working hypotheses that guided the research questions. The main hypothesis regarding intonation here is that the Arabic dialects that belong to the same language would still differ somewhat significantly in their intonational systems. They would differ in the distribution and type of intonational features, as well as in the level where prominence is marked in the prosodic structure. These typological differences are as Hellmuth and Chahal (2014) conclude, crucial for further research on Arabic intonation. Also, these reviews show that in controlled experiments conducted within the laboratory phonology framework, a great deal can be revealed about intonational features that may sometimes encourage further progress in theoretical frameworks.

2.5.3 Jeddah Arabic

The current thesis aims to propose a model of Jeddah Arabic intonation comparable to the models presented above. The main goals are to investigate the tonal make up and behaviour, the phonological levels and the focus realisation in this variety.

The dialect spoken in Jeddah belongs to the Hijazi group of dialects. Hijazi is a term that is assigned to the varieties spoken in the western region of the Kingdom of Saudi Arabia. This dialect has two types: the Bedouin Hijazi, spoken in the city of Taif and its suburbs, and the Urban Hijazi dialect, spoken in Makkah (Mecca), Jeddah, Medina, and Yanbu (Moussa, 2012 and references therein). The two types display some phonological and morphological differences, for many reasons, among which is the influence of a number of languages on the urban dialects (Abu-Mansour, 2008). This influence is related to the population and ethnicities settling there due to historical and modern migration, such as the Bukharies: referring to people from central Asia; Nigerians, Caucasians, Moroccans and Yemenies among others (Ingham, 2009, Al-Essa, 2009).

To my knowledge, there exists a limited number of phonetic or phonological studies on the Arabic dialect spoken in Jeddah (hereafter JA). Most references to the dialect in the literature are generally found in sociolinguistic studies. Meanwhile a very limited number of references exist for phonological accounts. Alshehri's (1993) study describes the effects of urbanisation on the dialects of rural immigrants to Jeddah. Another sociolinguistic by Al-Essa (2009), describes the dialect contact between JA and Najdi Arabic in regard to the de-affrication of Najdi affricates in the speech of the Najdi community settled in Jeddah. Regarding phonological studies, a few analyses aim to present an account of the general characteristics defining the Hijazi dialects- stress, syllabification and information structure- for example, Al-Zaidi (2014) on the Bedouin Hijazi dialect spoken in Taif, Al-Mohanna (2009) also on the Bedouin Hijazi varieties; Abdoh (2010) and Moussa (2013) on the other hand are conducted on the JA dialect and their findings will benefit the current study. The dissertation in Moussa (2013), is dedicated to the phonology of the dialect analysing its segmental phonology and phonological rules including syllabification. It concludes that Jeddah Arabic shares differences as well as similarities to the Arabic studies regarding segmental level phenomena. Within the L1 acquisition framework is Abdoh's (2010) Doctoral thesis that phonologically analyses the representation of first word in JA within the framework of prosodic theory in McCarthy and Prince (1986, 1990). The thesis is dedicated to the description and analysis of early words in the JA grammar, in addition to the word-level prosody in terms of the phonological acquisition of stress. It concludes that L1 acquisition of stress shows evidence for a mora and weight affect on preferred stress patterns in language development. The children show bias towards trochaic strong-weak patterns that preserve strong syllables in cases of deletion, e.g. preserving /`taan/ in /fus.`taan/ 'Dress'. This was taken by Abdoh as evidence that the children were faithful in observing internal foot structure of prosodic units (words).

The current study is therefore different both in scope and focus from the previously mentioned contributions on Hijazi, since this specific dialect's prosodic features and structure above the word level have not yet been described or analysed. The study of Arabic prosody and intonation in general is in its infancy, however there is a growing body of studies dedicated to the analysis of the language prosody. And although there exists no thorough study of this dialect's prosody, the thesis will benefit from a number of studies on the other well-studied Arabic dialects in general. The thesis will primarily benefit from Chahal's (2001, 2003, 1999) studies on Lebanese prosody and Hellmuth's (2002, 2004, 2005, 2006, 2007, 2011) series of studies on Egyptian prosody. The research questions, instrumental and theoretical analyses are directly influenced by the LA and EA models of intonation due to the theoretical focus of those studies that are parallel to the current thesis.

Though it must be noted here that some studies highlight a restricted comparability among the dialects due to the differences in the vocalic and consonantal durations and syllabic complexity affecting the phonetic implementation of the dialects' stress and rhythmic patterns (Ghazali et al., 2002). Thomas (2011) adds that dialectal variation in intonation for example can induce differences in the phonological inventory of tones, as well as their meaning and phonetic realisation. These points must be taken into account while analysing prosody and the display and distribution of the suprasegmental features in JA. It must be highlighted that the researcher is aware of such intricacies and in the absence of a corpus of comparative data or database on Arabic dialects' prosodic features, only general observations can be made. Detailed information about the methodology employed in this study is discussed in the following chapter.

Finally in this chapter is a conclusion with the main research questions to be investigated in the thesis. In light of the components of AM theory, here are the questions that will help determine where JA fits within the prosodic typology:

- In terms of the realisation of intonational prominence, what features does JA demonstrate? What is the inventory, density, distribution, and the tonal and non-tonal correlates for intonational prominence in JA?
- In terms of intonational range, how many tunes/meanings does the dialect express intonationally?
- In terms of the realisation of tonal forms, what tonal shapes does JA demonstrate and what is their function in an intonational contour?
- How is intonational prominence regarding Focus realised in JA? How are prominent accents distributed and ranked as they interact with the focal meaning of an utterance? What phonological and phonetic patterns does JA demonstrate as a result of this interaction?
- In terms of post-lexical prosodic levels, how many does JA demonstrate, and how do their correlates map with the syntactic grouping of an utterance? Also, what are the tonal and non-tonal boundary correlates that aid in arguing for the existence of those prosodic levels?
- With regards to the phonological Association and phonetic Alignment components of the AM theory, in what manner does a tone associate with the prosodic level, and in what manner does a pitch accent align on the surface?

Chapter 3. Methodology

3.1 Introduction

This chapter presents a detailed account of the methodology employed in the study. It includes the methods, corpus design, the segmentation and transcription processes, the recording process, as well as the theoretical motivation behind the choices in each.

3.2 Corpus Design

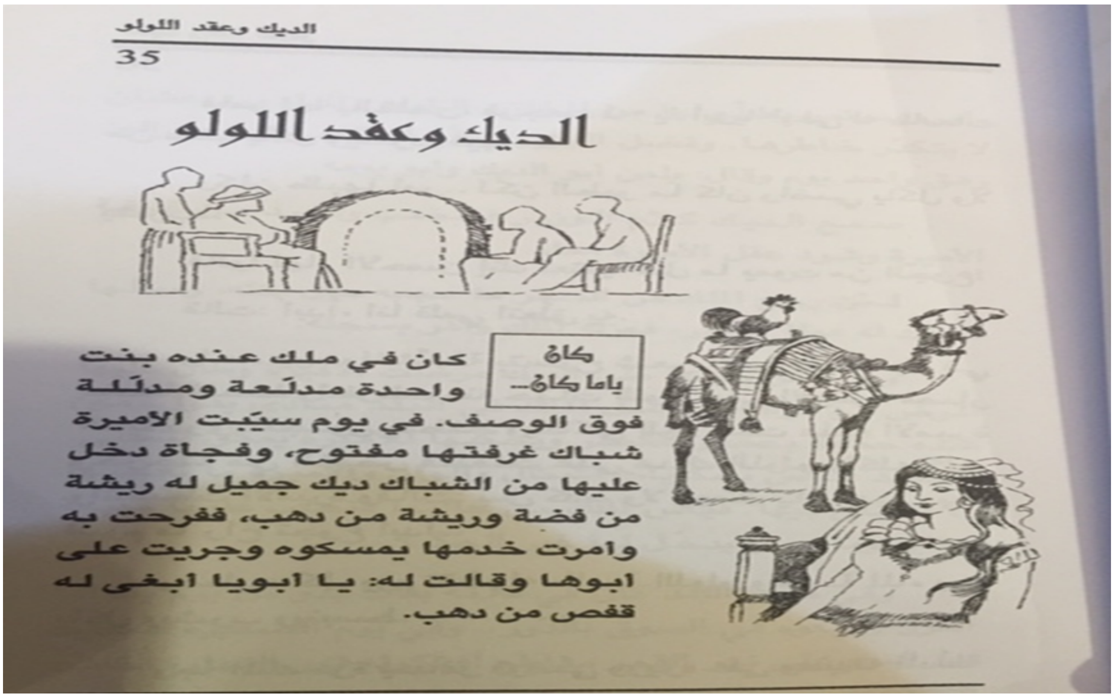
The corpus for this research study was designed to aid the investigation of the intonation structure in the Arabic dialect of Jeddah. As it may be known that no thorough account exists of the dialect to date, especially that of prosody. Therefore the corpus was designed to ensure the data is as representative as possible of the dialect and its patterns and system not only for the sake of the current study, but for future investigations as well. The aim was to collect data that can be viewed as a representative sample of the processes in the dialect, the structure and behaviour that will also serve as documentation of this variety. This will subsequently allow for future classification of the dialect in linguistic research.

Following previous studies on Arabic intonation (Al-Zaidi, 2014, Chahal, 2001, Hellmuth, 2006a) and according to the criteria for intonation research in Hellmuth et al's ongoing efforts for establishing a corpus of Intonational variation in Arabic (IVAr- A project taking place in association with the University of York and funded by the ESRC), material for the study consisted of the following controlled and semi-controlled variables that took place in four stages in the recording:

- Narrative: folk tale or story. Read then retold from memory.
- Scripted speech: variety of syntactic and semantic constructions with systematically controlled prominence positions.
- Short dialogue: information structure realisation.

3.2.1 Narrative

The narrative for this study is a tale from the folklore of Jeddah named "The story of the rooster and the pearl necklace". It is a well-known anecdote that was fortunately scripted and documented in the city's library archive. The folk tale is one from a series of tales in the form of more than ten scripted volumes in an attempt to document the city's oral culture. The tales were collected and validated by a group of academics led by the editor Dr. Lamia Baeshen, in association with the city's municipality council (Baeshen, 2008). Only part of the story was used in this study due to its long narrative that contains approximately 90-100 sentences. The story was obtained from the Library of Jeddah in old town and is written using JA in the colloquial form. It is worth noting here that to ensure participants used the informal and colloquial forms, some of the words were further reduced in spelling on the PowerPoint slides presented to participants to make them different from the MSA forms and thus encourage colloquial forms. In the following is an excerpt from the story followed by the full text transcription and translation:



الديك وعقد اللولو
35

الديك وعقد اللولو

كان
ياما كان...

كان في ملك عنده بنت واحدة مدلَعة ومدلَلة فوق الوصف. في يوم سيَبت الاميرة شباك غرفتها مفتوح، وفجاة دخل عليها من الشباك ديك جميل له ريشة من فضة وريشة من دهب، ففرحت به وامرت خدمها يمسكوه وجريت على ابوها وقالت له: يا ابويا ابغى له قفص من دهب.

Figure 3: The first page of The Story of the rooster in JA (Baeshen, 2008).

قصة الديك وعقد اللولو.

\gɪsˤsˤat addiik wu ʕʊgd alluulu\

كان يا مكان. كان في ملك عندو بنت وحدة مدلعة ومدللة فوق الوصف.

\kaan ja makaan\ \ kaan fii malik ʕindu bint waħda mdallaʕaw mdallala fɔɔg alwasˤf\

في يوم سّيّبت الاميرة شباك غرفتها مفتوح، وفجأة دخل عليها من الشباك ديك جميل لوه ريشة من فضة وريشة من دهب .

\fi jɔɔm sajjabat alʔamiira ʃʊbbaak ɣʊrfatha maftuuħ wʊ faʒʔa daxal ʕaleeha min aʃʃʊbbaak diik ʒamiil luu riiʃa min fidˤdˤaw riiʃa min dahab\

ففرحت بيه وأمرت خدمها يمسكوه وجريت على ابوها وقالتلو: يا أبويا ابغاك تجيبلو قفص من دهب.

\fafirħat bii wu ʔamarat xadamha jimsakuu wʊ ʒirjat ʕala ʔabuuhaw gaalatlu jaabuuja ʔabɣaak tiʒiblu gafasˤ min dahab\

فأَمَرلها بالقفص وجابوه. بعدين طلبتو وقالت: يا أبويا ابغالو طَاَسَة من دهب ياكل ويشرب فيها. وكانت طلباتها أوامر.

\faʔamarlaha bilgafasˤu jaabɔɔ. baʕdeen tˤalabatu gaalat: jaabuuja ʔabɣaalu tˤaasˤa min dahab jaakulu jiʃrab fiiha\ wu kaanat tˤalabaataha ʔawaamir\

لكن العجيب انو الديك ما كان راضي ياكل ولا يشرب.

\laakin alʕaʒiib ʔinnu addiik maa kaan radˤi jaakul wala jiʃrab\

قالو لها: الاحسن انك تعتقيه قبل مايموت من الجوع، قالت: ابداابداً، انا قلبي اتعلق بو مره.

\gaalɔɔlaha alʔaħsan ʔinnik tiʕtigii gabil ma jɪmuut min aʒʒuuʕ. gaalat ʔabadan ʔabadan ʔana galbi tʕallagbu marra\

وفي يوم من الايام والمّشّاطه بتسرّح شعر الاميرة انشبك في عقدها المشط وانفرطت منو حّبّات اللولو على الارض، فصار الديك زي المجنون ينطنط ويكاكي الين حّنّت عليه الاميرة وفتحتلو القفص وخّرّجتو وانفلت الديك على هداك اللولو واكلو اكل. وانبسطت الاميرة وقالت: بس كده، ولا يهمك ياحبيبي.. بعد اليوم ماراح تجوع ابداابداً.

\wʊf jɔɔm minalʔajjaam walmaʃʃaatˤa bitsarriħ ʃaʕr alʔamiira anʃabak fi ʕugdahal muʃʊt wanfaratˤat minnu ħabbaat alluulu 3ala alʔardˤ. fasˤaar addiik zajj almaʒnuun jinatˤnitˤu jikaaki ʔileen ħannat ʕalee alʔamiira wufataħatlul gafasˤ wuxarraʒatu wanfalat addiik ʕala haadaak alluulu wu ʔakalu ʔakil. wanbasˤatˤat alʔamiiraw gaalat bas kida wala jihimmak ja ħabiibi baʕd aljɔɔm maa raaħ tiʒuuʕ ʔabadan ʔabadan\

وصارت كل يوم تقدملو طاسة اللولو وطاسة موية ورد ياكل ويشرب وينبسط.

\ wʊ sˤaarat koll jɔɔm tigaddimlu tˤaasˤat alluulu w tˤaasˤat mɔɔjat ward jaakulu jiʃrabu jinbasˤitˤ\.

The story of the rooster and the pearl necklace.

Once upon a time. There was a king who had one daughter who was very spoiled beyond description.

One day the princess left her room window open, and suddenly a nice-looking rooster with one feather of gold and another of silver entered through the window.

This made her very happy and so she ordered the guards to capture it and she ran to her father and said: “oh father, I want you to get him a golden cage”.

So her father ordered for a cage and they got it. The princess then asked her father and said: “oh father, I want you to get him a golden bowl so he can use it to eat and drink”. All her wishes were granted/fulfilled”

But, surprisingly, the rooster refused to eat or drink.

People told her: “it’s best if you set him free before he dies of starvation”. To which she replied: “never! Never! I really love him”.

One day while the help was combing the princess’ hair, the comb got tangled in the pearl necklace the princess was wearing and shattered. All its pearls fell off on the floor. Seeing this, the rooster started jumping relentlessly and crowing loudly like a fool in his cage. Until the princess sympathised with him and opened the cage and let him out. The rooster ran towards the pearls shattered on the floor and started to eat the pearls. The princess was so glad she said: “is that it? Don’t you worry my dear. From this day forward you will not be hungry ever again”. Ever since, the princess gave the rooster pearls in a bowl to eat and a bowl of rose water to drink everyday tomake him happy.

The tale is used in an attempt to yield both controlled and semi-controlled forms of data. The controlled form is the read version, whereby all the participants were presented with the same version to read. The semi-controlled or spontaneous form is the re-told version whereby the participants were asked to re-tell the story from memory only aided by prompts containing key incidents and objects in the study. The prompts were in the form of pictures of a rooster, a pearl necklace, a king and princess, a golden cage and a golden bowl, all of which are important parts of the tale.

The use of narratives in intonation studies serves two simultaneous purposes. First, the read (controlled) form demonstrates the "reading intonation" style, as opposed to the "conversational intonation" style, which has been shown to be livelier and richer in terms of intonational markings (Jun & Fletcher, 2014). Conversely, the read form regarded as "laboratory speech" yields very slow, clear and careful productions (Xu, 2010). Accordingly, and by virtue of being systematically 'controlled', this allows the researcher to control for speech rate, occurrence of prosodic features and their type, desired/target tunes and their composition, as well as exact locations/distribution of targets. In addition, in the read form, the phrases are controlled for by the use of commas and full stops that cue sentence beginning/end and continuation, therefore cuing syntactic phrasing/grouping. Ideally, all of these controlled aspects would be comparable across speakers. Moreover, and as Xu (2010) argues that controlling these factors helps distinguish prosodic features variation as a result of style (read vs. conversational) from variation as a result of other phonetic and phonological factors. Second, by including a re-told (semi-controlled/spontaneous) form we observe the naturally occurring intonational markings and their distribution parallel to what they would occur as in real life and daily conversations.

The tale itself contains approximately 17-25 utterances in the form of declarative statements and 'indirect' reported speech. It is therefore expected that the tale would demonstrate mainly declarative tunes, including the reported speech style that would demonstrate declarative or imperative tunes. These are expected to show a falling contour/intonation towards the end of the utterance, as reported in all the previously mentioned Arabic dialects. The tale would also demonstrate- in addition to the tonal inventory of the dialect- other tunes such as the continuation and declarative plateau tunes-especially in the re-told version. These are used when a participant wishes to show that the utterance is not yet finished and will be continued, or that s/he is trying to remember relevant information or is thinking about it while speaking- all of which are cued by special edge configurations (Chahal, 1999). These additional tunes may thus increase or decrease the number of sentences in the re-told version.

One of the main differences, as mentioned previously, between read and re-told data is the phrasing levels and their occurrence, i.e. the number of inserted boundaries marking the prosodic constituents: Intonational Phrases (IPs) and Intermediate phrases (ips) in a string of sentences. The working hypothesis is that phrasing levels (IP, ip) would potentially differ in their number or occurrence across the two narrative types with retold versions containing less phrases. This has been shown in studies to be related to the occurrence of pauses in the utterance. A slow and careful production yields more pauses and therefore more phrases and vice versa (Jun & Fletcher, 2014). Accordingly, the narrative was used in this study to figure out the tunes that exist in the dialect, their configuration, the tones in the dialect and the phrasing levels. Details and examples from the narrative are presented in chapter [4].

3.2.2 Scripted Speech

In this part of the corpus, a broader range of sentences was used that would yield more tunes in addition to the narrative tunes. A total of 45 sentences constitute the scripted speech material. This total includes the tonal alignment experiment stimuli, the focus experiment stimuli, the intonational tunes stimuli, filler sentences, but excludes the 3 repetitions. Please refer to appendix [D] for details of the target sentences. This part contains question tunes- 3 yes/no and 2 WH types, 5 in total excluding repitition s. 2 imperative sentences, and 1 request sentence to yield imperative and request tunes, 3 in total excluding repetitions. The question tunes were elicited by using question words in WH questions, and question marks. The imperative tunes were elicited by using exclamation marks ending the target sentences. As for the requests, in anticipation of the the discussion regarding this tune in chapter [4], those tunes were elicited by the use of the polite marker /mumkin/ ‘please’, which immediately prompted the participants to use a polite tune. The polite marker and ending the sentence in a question mark was enough for the participants to produce the correct request tune with no problems faced by the researcher during the recording. Filler sentences include longer stretches of sentences that are helpful in observing the phrasing patterns in the dialect needed to establish prosodic constituents. Those 7 sentences varied in syntactic constructions to include: coordination, negation, nominal, listing (order), construct state and mathematical constructions. The tonal experiment stimuli consist of 20 target sentences in total excluding repetitions and their elicitation process is discussed below. Finally, the focus experiment stimuli consist of 6 target sentences, and 4 filler sentences. All in all, 2700 utterances including 3 repetitions of each by

20 speakers constitute the full corpus for the current thesis; details are reported in appendix [D].

3.2.2.1 Focus dialogue

The goal of including focus in this study is to foremost investigate whether or not focus is marked intonationally in the dialect of Jeddah. Moreover, it is included to examine the effect of focus on accent distribution, the prominence hierarchy in the respective sentence, and the phonetic and phonological patterns distinguishing broad/narrow focus. The phonetic details include, F0 height and excursion size, rise time, rise speed, duration, and intensity discussed thoroughly in the respective chapter. The phonological details include, accent type, and sentential position of focus, all of which are examined in the focused constituent /`laa.na/, as well as in the pre and post focal word /ri.`taal/. Accordingly, the design of the focus material constitutes of declarative sentences as answers to questions in the form of a short dialogue. The use of a dialogue paradigm to elicit information focus is encouraged in many studies (Al-Zaidi, 2014, Jun & Fletcher, 2014, Chahal, 2011, Xu, 2011). Material for this experiment is presented in detail in Appendix [D]. The stimuli is designed to yield comparable and controlled targets- controlled in sentence length and segmental composition- in different sentential positions and under different focus conditions. Each identical target sentence: /[laana] xarajat ma3 [ritaal] gabil 2ams/ 'Laana went out with Ritaal two days ago' under (broad/narrow) is examined to observe the phonetic manifestation of the categorical differences that occur between the two conditions, as well as according to sentential position (bracketed names alternate in position). The analysis of the target material includes the measurement of: vowel and syllable duration to assess the quantitative effect of focus type narrow/ broad on length. Vowel/syllable/word/sentence F0 Hz/ST values are measured to assess peak height and range as a result of focus type. Syllable/word and sentence Pitch range Hz/ST are measured to assess the effect of focus on excursion size of target word and pre/post string. Syllable Rise speed (ST/S) and rise duration (MS) are measured to assess the effect of focus on the rise speed in rising tones L+ H*. And finally, vowel, pre/post focal streng Intensity dB is measured to assess the effect of focus on amplitude.

3.2.2.2 Tonal Alignment material

Along with the previously mentioned sentence types and tunes were the target tonal alignment sentences uttered with a declarative tune. Tonal alignment is the temporal alignment of fundamental frequency F0 with respect to the segmental string and as observed in the F0 contour (Prieto, 2011). Tonal alignment is thus considered as the phonetic/acoustic component of the theory of intonation, a concept that is crucial in any analysis of intonation as discussed earlier in section 0.

Tonal alignment patterns and behaviour is reported to: 1) cue and mark prosodic constituents, 2) cue and demonstrate the canonical TBU *Tone Bearing Unit* in a given language or variety, and 3) at the level of perception, encode intonational contrasts affecting the pragmatics of the uttered tune. For the purposes of the current thesis' aims and research questions, the focus will be on investigating the first two aspects; the third aspect is beyond the scope of the current study and will be investigated in the future.

Accordingly, to investigate how tonal alignment patterns cue prosodic constituents and TBUs, the design of the experiments must control a number of contextual and microprosodic factors. Studies have shown that the segmental composition and identity of the string (manner, voicing, length) affect the peak location (Wichmann et al., 2000). Sonorants are best candidates for alignment studies as they maintain the F0 contour and minimise tracking errors (Prieto et al., 1995). The duration of the segments have been reported to show differences in alignment between long and short segments- late vs. early alignment, respectively (Chahal, 2001, Prieto et al., 1995, D'Imperio et al., 2007). The duration of the syllable as a whole was also reported to be highly correlated with alignment location (D'Imperio et al., 2007 for Neapolitan Italian, Chahal, 2001 for Lebanese Arabic). Syllable structure (open/closed) and syllable weight has been reported to show differences in alignment even in dialects of the same language (Yeou et al., 2007a, D'Imperio et al., 2007, for discussions on syllable weight cf. Hayes, 1995, Watson, 2002, Broselow et al., 1997).

As was discussed earlier in the previous chapter, the position of the target accented syllable also affects tonal alignment. These effects of location or 'time-dependent' re-adjustments are recognised as stress (accent) clash, tonal crowding and prosodic boundary effects in the literature. Studies report that in these cases a re-alignment or re-adjusment of the peak takes place. This behaviour is also accompanied by lengthening the target accented syllable in some languages (Prieto et al., 1995 for Mexican Spanish, Frota, 2012 for European Portuguese). And this tonal alignment behaviour is also provided as evidence for the existence of some prosodic constituents in the languages under investigation.

Observing how JA resolves or behaves in the aforementioned environments is one of the main goals of the experiment. Therefore, the scripted material to investigate tonal alignment was designed to include four tri-syllabic target words with varied target stressed syllable locations (initial-medial-final) and constant segmental compositions. All the syllables had the sonorant /l/ as onset, long and short /a/ vowel as nucleus and, /t/ as coda in closed syllable type.

Yielding target syllables of the shapes: /la/, /laa/, /lat/ and /laat/, that vary in syllable type and mora count respectively. The words were placed in five structural positions in a declarative sentence. Material for this experiment is presented in detail in chapter [6] and Appendix [D]. Target stimuli were examined for the occurrence of pitch accents, phrase and boundary tones, in addition to the shape/ type of pitch accent that occurred on target words. Accordingly, target word, syllable, and sentence were the domain for acoustic analysis. Maxima H and minima L, F0 turning points were examined in the target word and syllable. Quantitatively, relative and absolute alignment measurements were carried out. Following previous studies, the potential segmental landmarks in target words were identified as 1) for the alignment of valley L: the anchor points measured are the onset of stressed syllable, and onset of stressed vowel, 2) for the alignment of peak H: the anchor points are the onset and offset of the stressed vowel, onset of the stressed syllable, and offset of the stressed rhyme.

3.3 Recordings

Recordings for the study took place in Jeddah during a field trip to the city. As per the Newcastle University ethical regulations, the recordings proceeded after obtaining permission from the participants in the form of consent. Additionally and in accordance with university regulations, the participants were informed about the nature of the conducted study, general aims and confidentiality maintenance procedures. Each participant was recorded individually in a quiet room and a total of 20 recordings were collected- 10 male and 10 female files. The recordings were carried out using an Edirol digital voice recorder, model number (R09HR). Participants were seated comfortably in front of a Macbook monitor and the recorder was placed nearby on a stack of books. An attachable (lapel) microphone was placed approximately 15 cm away from speaker's mouth and wirelessly connected to the Edirol recorder. On the monitor were the PowerPoint slides containing target material. Each slide contained maximally 5 sentences and was ensured to not contain the repetition of a sentence, as well as ensuring the inclusion of filler sentences. Before the actual recording took place, the participants were talked through the instructions and practiced a couple of utterances to familiarise themselves with the procedure. Then, the recording proceeded in four stages with 1-2 minutes break in between if the participant asked for one. The first stage was the read folk tale, the second was the retold folk tale with prompts, the third was the scripted speech material, and the fourth stage was the focus dialogue between researcher and participant. The researcher believes that starting off the recordings by the folk tale sets the desired tone for the experiment and encourages subjects to use dialectal forms throughout. The subjects were instructed to read material as normal as possible at normal speech rate. For the focus material, they were specifically instructed to emphasise target in narrow focus (Xu & Xu, 2005, Wang & Xu, 2011). In the event an utterance was perceived by the experimenter not to have been produced correctly, the participant was asked to repeat the whole utterance again. The length of the recordings was 13 minutes on average including filler sentences and three randomised repetitions of target material. The digital voice recorder was set to save audio files in WAV format that were then digitised and saved on the computer and hard drive for backup. The sampling rate was 44 kHz and the sampling format was 24-bit Stereo channel (2 channels).

3.4 Annotation, Segmentation, Transcription

The 20 collected audio files were subsequently ready for annotation. This process was generally carried out by the careful examination of the acoustic waveform of the recordings on PRAAT (Boersma & Weenink, 2019). The analysis proceeded in four stages, namely: MarkPause segmentation, orthography labelling, Forced-Alignment segmentation, Prosogram for phonetic annotation, and phonological ToBI transcription. As may be known it is hard to obtain and use fully automatic and reliable annotation systems- and especially for Arabic-, therefore the systems used in this research study are semi-automatic that required further extensive manual labour.

3.4.1 Mark Pause

The first process took place in the aim of creating a textgrid for each soundfile and insert boundaries around pauses and speech streams. This was obtained by using a PRAAT script named MarkPause, found via The Speech Corpus Toolkit for Praat *SpeCT*, available at: https://lennes.github.io/spect/ . The use of the script was as a first-pass for subsequent ease of annotation. As mentioned in the description, this script utilises Praat's silence detection criteria to mark boundaries around speech sounds and silences/pauses based on intensity values in a long soundfile. One of the advantages of this script is that it allows the user to define the upper and lower intensity thresholds as well as the duration of a pause according to the research goals. For this study, the standard configurations in the script were used as they are within the reported margins advised for measuring silences in connected speech. For pause duration, this was set to be minimally 6 seconds or less. For pause intensity, this was set to be maximally 59 dB or less. Maintaining the same threshold was found to mark pauses differently across speakers (errors detected) and so the criteria was modified accordingly for each sound file, as advised by the description of the script.

3.4.2 Orthography- labelling

The next stage is the manual orthography-labelling segmentation. After obtaining a blank textgrid with boundaries only around pauses, this stage was performed to label the speech intervals. An orthography-labelled textgrid is further needed for the next Forced-Alignment stage. To label speech, the Arabic transliteration notation developed by Al-Tamimi and Khattab (2015) was used (please refer to Appendix A: Transliteration) .It is a notation based on the Semtalk tool that adapts and converts the CHILDES system to be compatible with Arabic and Hebrew transcription. This was done to write out full sentences on an orthography tier in the textgrid. The current stage resulted in a fully annotated textgrid and sound file that was ready for the next stage: segmentation of individual words and sounds within the sentences.

3.4.3 Forced Alignment

The third stage is the Forced-Alignment semi-automatic segmentation. As previously mentioned it is to segment words and phonemes, and it relies on speech detection to mark boundaries around the targets. This Forced alignment system is developed by Jalal Al-Tamimi for Arabic and based on the plugin PraatAlign version 1.9 (Lubbers and Torreira, 2016). The system uses the acoustic models used in MAUS forced alignment system (Schiel, 1999) and more specifically the "universal" acoustic model. To allow for Praat align to recognise the transliteration used in the current thesis, a phonetiser was created that matched each transliterated symbol to SAMPA. The forced Alignment system then used the HTK toolkit to perform the acoustic segmentation and the matching between grapheme and the acoustic signal. After running this, the resulting textgrid was manually checked and examined for accuracy. Some issues were detected and modified in the output textgrid, which is expected as this is a semi-automatic process and manual input is indeed needed. The issues arising were mainly in relation to the segmentation of specific phonemes. For instance, long vowels failed to be correctly segmented, unlike short vowels. Also, plain/pharyngealised fricatives, nasals and glides were most frequently inaccurately segmented and the system tends to cut in the middle of these segments substantially reducing their duration. Nonetheless, the segmentation and boundaries around whole words was almost perfect with a 90% accuracy rate. Generally, the use of the Forced-Alignment system helped tremendously speed up the phonetic annotation process, in comparison to a 100% manual annotation. To recap, this stage resulted in a textgrid with two further tiers (three overall) for phoneme, word, and orthography.

3.4.4 Phoneme segmentation

Manual modification of inaccurate phonemes and overall segmentation of all phonemes in a sound file was performed following the guidelines of prosody segmentation in Turk et al. (2006). This segmenting criteria is defined relative to the constriction onset and release as observed in the spectrogram and waveform. For oral closures, the onset was defined as accompanied by a low overall amplitude and absence of first 3 higher formants. And the release is defined as where the burst takes place. If multiple bursts occur, the first is taken as release onset. If no burst is evident, following F2 onset is taken as stop release. The duration of vowels following voiced and voiceless stops start at previous consonant release, to thus maintain constant criteria relative to consonant constriction. For sibilants, the preceding and following vowel F2 was used to determine frication energy offset and onset. In addition, there was a lookout for offset/onset frication energy in spectrogram and accompanying silent gaps that were also used as criteria for defining sibilants. Nasal stops were marked according to abrupt spectral changes at onset/offset, as well as the occurrence of sharp dips and rises in the waveform. For glides and /h/ that prove to be the most difficult to segment cross-linguistically, the mid-point of preceding vowel to glide was taken as a reference point to observe formant structure, and the mid-point of transition from a glide to a subsequent vowel as reference point. Lastly, for lateral /l/, the constriction onset and release was defined as the point where spectral discontinuity occurred.

3.4.5 ToBI Transcription

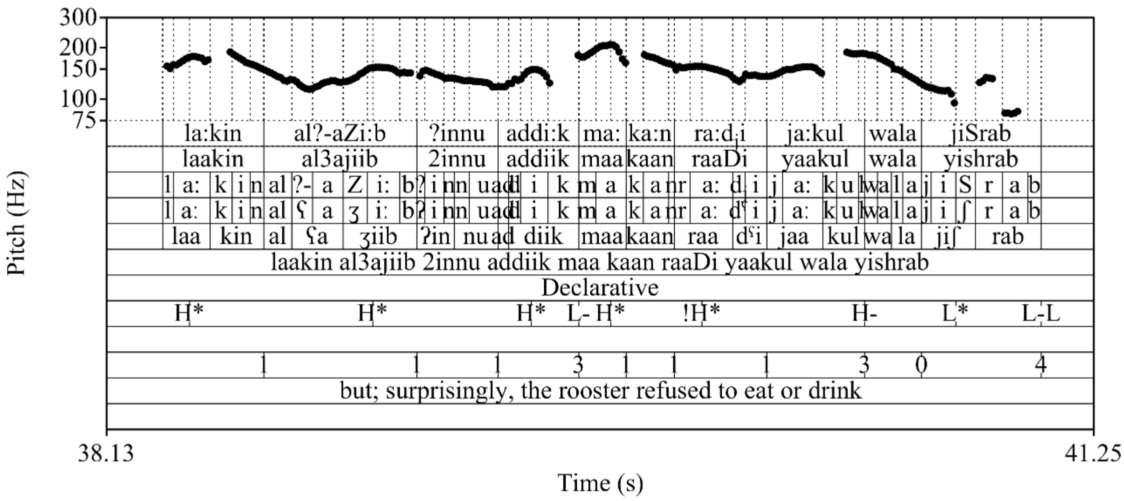


Figure 4: ToBI annotation of a declarative tune. Speaker 01_M

The first stage of prosodic annotation is the phonological ToBI transcription stage. For this process the thesis follows the guidelines and textgrid format for prosodic annotation provided by Beckman and Hirschberg (1994) via The Ohio State University website to label the F0 contour in the target textgrid. In this level, in addition to the phoneme, word, and orthography tiers, there is a break-index tier, tone tier and miscellaneous tier. A syllable and a translation tier were also added to mark boundaries around syllables and translate the Arabic utterances into English for ease of illustration. The break-index tier is one of the important elements in intonation transcription as it is the tier specifying lexical and phrasal junctures in connected speech. Each juncture/boundary is marked at the right edge and is assigned a value from zero to 4. Zero marks clitics, 1 marks word boundaries, 2 marks pauses with no significant tonal marks, 3 marks intermediate phrases, and 4 marks full intonational phrases. Next, the tonal tier- not less important than the break-index- is a tier for specifying the tonal targets and shapes observed in the contour. The AM theory being a two-tone system only recognises two tone targets H, L and allows for combinations of the two. Finally, the miscellaneous tier is for specifying any comments and uncertainties about the utterance, such as laughter, hesitations, prolongation, etc. Of course, there exist language and dialect-specific accent shapes and categories according to pitch range, alignment and scaling, therefore the transcribed tones in this thesis were based on how they were observed in the speakers' contours. The following paragraph provides information on the starting assumptions regarding the reported tonal categories for JA.

The identification of the intonational categories of pitch accents, boundary tones and prosodic structure (phrasing levels) in the current thesis is carried out in light the reported categories in Chahal (1999, 2001), Hellmuth (2006) and Hellmuth and Chahal (2014). The intonational categories in the language are based on the observed F0 contour. According to the speaker's local pitch range, the phonological categories of pitch accents and edge tones are defined. That is, for every marked phrase, the H and L targets are defined in relation to the local pitch range for the phrase. For example, we mark an H* pitch accent occurring in an intermediate phrase where we observe the F0 peak to be. This implies that the F0 value of this accent be in the mid to upper part of the speaker's range for that phrase. Another example, we mark an H- phrase tone demarcating the edge of the intermediate phrase boundary where we observe an elevated trend in the contour, again, the value would be in mid-upper speaker range. One of the functions of phrase tones is controlling the pitch configuration in the stretch between the last pitch accent in the phrase and the phrase tone. That is, a high phrase tone is expected to be preceded by a plateau or rising configuration in the upper speaker range. This is defined as *interpolation* in the theory. As Arvaniti (2012, 2017) explains, in AM, the F0 contour is phonetically realised as a series of target tones and interpolation in between the targets. The target tones are the salient turning points in the F0 contour (peaks and valleys) while the remainder of the contour is generated by interpolation between the targets. Hence the theory combines the phonetics and phonology to realise intonational categories. As for the prosodic phrases, both the formally mentioned tonal correlates were examined as well as non-tonal correlates. Following the analysis in Chahal and Hellmuth's studies, the main non-tonal correlate investigated is pre-boundary lengthening of units before intermediate and intonational phrase boundaries. For the identification of those boundaries, the tonal alignment material was used, as the target words are placed in intermediate phrase boundary and in intonational phrase boundary. Subsequently, the duration of the target stressed/accented syllable, and vowel was extracted. The stressed/accented syllable location in the tonal alignment experiment varies in distance to the prosodic boundary in question. The stressed/accented syllable can be located two syllables from the boundary /`la.ha.fu/, one syllable from the boundary /mu.`laa.zim/and /riH.`lat.hum/, or zero syllables from the boundary /lil.Haf.`laat/. For the full list of sentences please refer to Appendix [D]. The results of this analysis are discussed in chapter [4].

The following table summarises the phonetic specification for the reported Arabic pitch accents in Hellmuth and Chahal (2014), and Hellmuth (2014). The surface realisation is with regards to the accented stressed syllable in the prosodic word. These are the criteria employed throughout this study in order to identify the JA tonal categories.

Pitch accent	**Phonetic realisation**
H*	Monotonal high peak realised in speaker's middle upper range. It occurrs within the span of the accented syllable.
L*	Monotonal low valley realised in speaker's lower pitch range. Occurrs within the span of the accented syllable.
L+H*	Bitonal pitch accent. The starred peak is realised within the rhyme of accented syllable, whereas the leading L aligns exactly with the onset of the stressed syllable.
L*+H, H*+L and H+L* [7]	L*+H, Bitonal accent. The starred valley occurs during the accented syllable that is followed by a sharp H rise that occurrs outside the accented syllable (postaccentually). H*+L bitonal accent. The starred H is realised as a high plateau within

[7] Those pitch accents are proposed for Sanaani Arabic in Hellmuth (2014). They are not reported for Lebanese or Egyptian.

	the span of the accented syllable that is followed by a fall in F0 outside it. The opposite counterpart H+L* where the starred tone is realised as a fall from a preceding high during the accented syllable.
L+!H*	A rising bitonal accent where the peak is realised in a lower level (downstepped) than preceding high peaks during the accented syllable.
!H*	A monotonal downstepped peak that is realised in a lower level than a high peak. Realised in the mid pitch range.
H+!H*[8]	A bitonal accent where the starred high is downstepped to a lower level compared to the preceding high tone occurring within the span of the accented syllable.

Table 3: Pitch accent realisation in Arabic. Adapted from the literature.

In this stage of prosodic annotation, we also define the tunes and their composition. A tune is an intonational pattern with a distinct pragmatic meaning. A tune is defined relative to the tonal configuration at the edge of an IP in combination with the last ip phrase tone. Cross-linguistically, common configurations have been found to accompany major tune types (Chahal & Hellmuth, 2014, Chahal, 1999, Beckman & Hirschberg, 1994). For example, a falling contour with L- phrase tone and L% boundary tone (shape: L-L%)- both very low points in speaker's range- has been found to usually mark declarative tunes and sentences. An opposite configuration H-H% has been found to mark polar questions. An L-H% configuration was found to mark continuation tunes that indicate incompleteness, among other tune types. As will be discussed in chapter [4], those edge configurations are also reported for Arabic albeit with reported variation in the combination of nuclear accents, edge configurations and pitch range manipulation that mark the particular intonational tunes. In addition to L-L% to mark declaratives, and H-H% to mark interrogatives, Lebanese and Egyptian Arabic dialects show a number of other intonational tunes. Among them is the L-H% edge sequence to express incompleteness of an utterance for example. Overall, the guidelines in this section form the working hypotheses for identifying and characterising the tonal composition for the Jeddah Arabic dialect.

[8] This pitch accent is only reported for Lebanese in Chahal (2001, 2014). Not reported for Sanaani or Egyptian.

3.4.6 Prosogram[9]

Prosogram (version 2.18e) is a prosody analysis tool that helps transcribe pitch variations in spoken languages. This tool works on the stylisation of the pitch movements in perceptual Semitones (1 ST relative to 1 Hz) taking into consideration the speaker's pitch profile. Pitch stylisation is based on a model of tonal perception as proposed by early works of d'Alessandro and Mertens (1995). Prosogram, detects the F0 direction and size, intensity, and durations of voiced segments in a given utterance. The script also detects silent pauses in longer stretches of speech, as well as F0 discontinuities (typical of octave jumps). The tool also has the option of automatic, semi-automatic or manual segmentation of corpus using alignment methods such as Easy Align that was successfully applied to many European languages (a list of the publications are included in the user guide on the author's page). Unfortunately, Arabic is not supported by these methods and therefore manual phoneme segmentation of the data was carried out prior to this step as was discussed previously. It is worth noting that the intonational structure regarding tonal configuration and tunes in this thesis is based on the F0 contour as observed in Praat. The Prosogram tool is thus only used for pitch profiling as confirmatory analysis for the ToBI transcription of intonational categories. For illustrative purposes, Appendix [E] contains prosograms counterparts of all the textgrids reported in this thesis.

Prosogram was used in order to second pass the ToBI annotation reported above. At times, the researcher has faced a great deal of difficulty determining certain surface F0 patterns during ToBI analysis on PRAAT. The main difficulty faced was in regards to the scaling (level) of a tone with respect to the observed pitch range. The scaling (downstepping marked with an !) of the High tone as can be inferred from table [3] above, matters in the categorical difference between high accents in Arabic dialects. Additionally, it was shown in section 2.4.1 above that high edge tones are also subject to scaling differences. In the Lebanese Arabic dialect, Hellmuth and Chahal (2014) report edge configurations of the type: !H-L% that mark mild reproach tunes, and !H-H% that are used to label some unmarked continuation tunes. As the current corpus contains long stretches of material across 20 different speakers, the labelling of the ToBI categories needed to take into account the speaker and utterance pitch range for each labelled sentence. In PRAAT, pitch range is not a readily available option in the textgrid window, and so having to recall the pitch information for speaker and utterance or go back earlier in the textgrid to determine speaker pitch range proved a tedious task. In Prosogram, the

[9] I would like to thank Prosogram author Piet Mertens for his valuable input and troubleshooting cooperation via email during this stage.

pitch range is included in the utterance window and shows the upper, mid and lower ranges, which facilitates the identification of the target tone. In the following paragraphs, I provide examples of the difficulties I faced while annotating the speech signal in PRAAT. Only for the purposes of this section, I provide screenshots of the examples as they would occur in PRAAT, and as I would have seen them at the time of transcription.

The WH question tune in Jeddah Arabic is reported in chapter [4] to be of the following contour: H* nuclear accent and !H-H% edge configuration. Over the course of transcription, it was unclear to me as a novice intonation researcher how to label this edge configuration, while bearing in mind that in Arabic question tunes are typically realised with a high rising boundary. The speakers in the corpus produced WH edge configurations that compared to declarative tunes, are not low enough to be realised as such- even when the declarative tune ends in a high nuclear accent. On the other hand, compared to a typical high rising yes/no question tune, they are not high enough to be realised as such. The following is an example of the labelling of a WH question tune in PRAAT, and the identical tune in Prosogram:

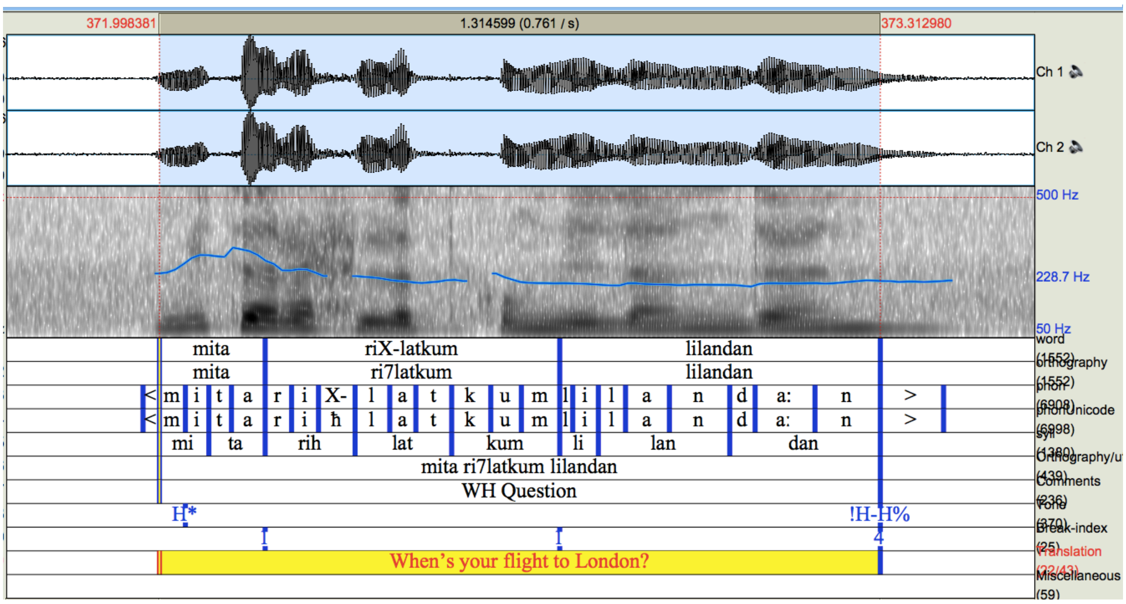


Figure 5: A WH question tune with ToBI annotation in PRAAT. Speaker 06_F

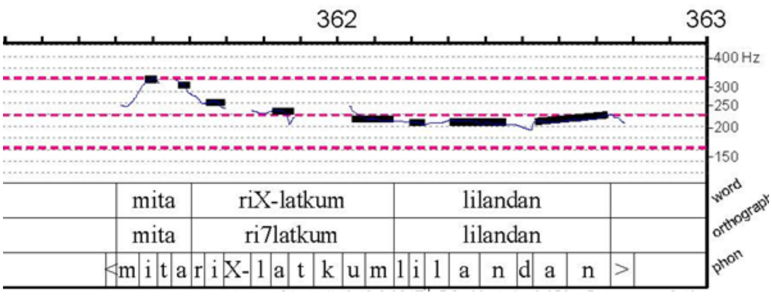


Figure 6: A WH question tune in Prosogram Manal WH question. Speaker 06_F

In Prosogram, I could see that the edge falls within the mid pitch range, and a low is not targeted. The rapid turning point following the H* nuclear accent on the question word /`mita/ ‘When’ is analysed as tone spreading of the downstepped -!H phrase tone until the end of the phrase.

Another common difficulty I faced was the F0 discontinuty in contours. Those discontinuities at times lead to uncertainty and ambiguity concerning whether a pitch configuration is present or not, whether it is a target accent, a pitch transition, etc. Prosogram was useful here as the phonetic tool only recognises and reports the salient pitch configurations in Semitones. This was faced particularly in the labelling of the nuclear accent and edge tones of yes/no questions. In chapter [4] yes/no question tunes are labelled as follows: L+H* nuclear accent and high rising H-H% as edge configuration. In PRAAT, some contours looked like the following:

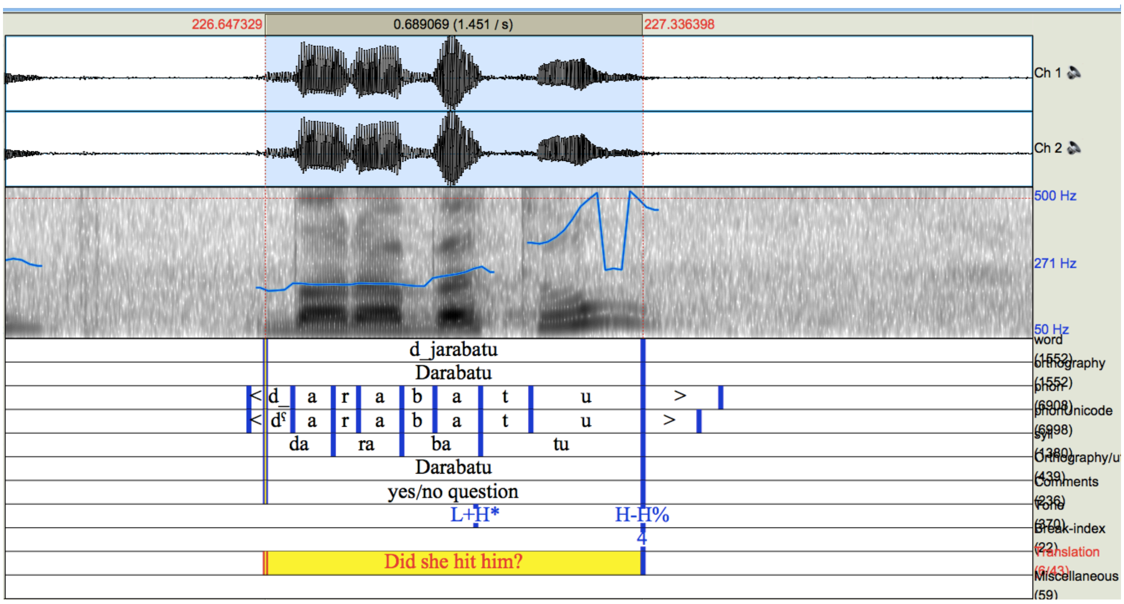


Figure 7: A yes/no question with ToBI annotation in PRAAT. Speaker 10_F

The figure shows a high edge, but it is also showing some other movement. Additionally, there is a seemingly high pitch movement on the stressed/accented syllable /ba/, which is also preceded by a brief low turning point. In prosogram the pitch contour demonstrates that the accent on the stressed syllable is L+H* as the high occurs within the upper mid pitch range and is aligned with the nucleus of the syllable, while the leading tone starts from a low point and is aligned with the onset of the syllable- typical of reported Arabic L+H* accents. This manner is also used to rule out an L* nuclear accent interpretation. As for the edge movement, we can see in the following that Prosogram retains the most important pitch movement and cleans up the unnecessary information.

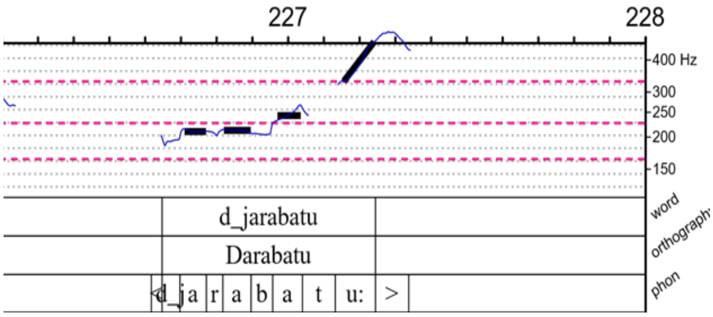


Figure 8: A yes/no question tune in Prosogram. Speaker 10_F

Those were the main aspects where ToBI and Prosogram were used together to inform the phonological and phonetic analysis of the F0 contour for the Jeddah Arabic dialect.

This discussion concludes this part on the phonetic and phonological annotation of the recorded material. We have seen that in this stage, the recorded files are annotated word-by-word, syllable-by-syllable and phrase-by-phrase. We also specify in this stage the intonational categories and their composition. Evidence was presented for the workflow followed in the transcription process and the rigorous examination of the textgrids phonetically and phonologically. In order to obtain acoustic information and measurements for each of the experiment chapters [5] and [6], specially tailored Praat scripts were designed with the help of the supervisor. The details of these scripts are discussed in each respective chapter.

Chapter 4. JA Tunes and Tones

4.1 Introduction

This chapter is an analysis and discussion of the intonational system in Jeddah Arabic. The aim is to propose an intonational account of the dialect combining the phonetic and phonological levels. A successful study and analysis of intonation within the Autosegmental-Metrical Approach, or intonation analysis in general must include all or some of the following components according to the goals and interests of the study (Jun & Fletcher, 2014):

1. The tonal inventory and tunes.
2. The intonational categories (pitch accents, phrasal tone, boundary tone): Prosodic structure and its marking (phrasing).
3. Focus prosody: Prominence hierarchy.
4. Tonal alignment patterns.

These will help give a thorough account of the intonation of the language or dialect under study as they encompass the major areas essential to intonation. The tonal inventory of the dialect is crucial to observe what tones exist in the language and their specification. The intonational categories' importance stem from their conveyance of prominence and prominence relationships in the language. Lastly, the prosodic structure, information structure and alignment behaviour also serve an important role in cuing intonational categories and prominence as will be seen shortly. Essentially, in order to study intonation we need to identify intonational features, their specification, role and behaviour.

In the following discussion is a review of aspects 1 and 2 above that are relevant to the analyses in this chapter, aspects 3 and 4 are discussed thoroughly in the subsequent chapters. Regarding the tonal and melodic inventory in a language, there is cross- linguistic literature reporting some universal tendencies for what pragmatic contours are expected to demonstrate. Among those are interrogative and declarative contours that show semi-consistent trends across languages. Declarative contours are reported to generally exhibit a falling configuration throughout the utterance accompanied with a low boundary marking the edge. On the other hand, polar questions for example are reported to show a rising trend ending with a high boundary on the edge of the phrase. The cross- linguistic variation in the marking of those two general patterns lies in the choice of the nuclear accents and in the sequence and scaling of the tones according to the pitch register of the language. More generally, in addition to those two trends exists a continuum of patterns that are used to mark the contours of the other pragmatic tunes such as requests, plateaux that indicate uncertainty about the utterance, and continuation tunes. Among those, requests for instance have shown a variability in whether or not the intended meaning is a polite request or an imperative, where the emotional state of the speaker is incorporated into the realisation of the contour (Wichmann, 2004).

In Arabic, the two extensive studies on the different melodies for Lebanese Arabic and for Egyptian Arabic in Chahal (2001), Hellmuth and Chahal (2014), and Hellmuth (2006a) are used as general hypotheses in this chapter. Regarding declaratives, both Arabic dialects demonstrate a low falling edge configuration L-L% that includes variable accent shapes in nuclear position. For polar question tunes, again both display a high rising configuration H-H% at the edge of the phrase that combines with a low L* nuclear accent in LA, and a rising L+H* accent in EA. For WH questions, the author states that LA displays an H-H% edge that is inconsistent across speakers. While in EA, WH questions end in a falling edge configuration L-L%. The two dialects use different strategies to mark continuation tunes. In LA, continuation tunes used to indicate non-finality and incompleteness end with a boundary shape of H-L% or L-H% that combines with a H* or L+H* nuclear accent. In EA continuation tunes are marked by a high rising edge configuration of the type H-H% that combines with a downstepped !L+H* accent. Plateau tunes are only generalisable in LA as they show consistent patterns across the analysed data. They are pragmatically used to express reproach in LA and are marked by a downstepped !H-L% edge configuration that combines with different nuclear accent shapes. Finally, regarding imperative and request tunes, neither of the studies on LA nor EA include those pragmatic tunes in their corpora, and thus the thesis aims to expand on the existing inventory of melodic tunes in Arabic by including those two tunes in the analysis. Analysing more sentence types would help observe the range of intonational expression in Arabic, as well as observe how differences in pitch register are used to convey different intonational meanings. Based on the corpus-based study of requests and imperatives in Wichmann (2004), those tunes are marked in British English by a falling or rising edge configurations. Respectively, polite requests including ‘please’ in the sentence, vary according to whether the polite marker is in final or no- final position. When ‘please’ occurs non-finally the contour shows a falling pitch register leading to the boundary L%; however, when it occurs finally, the contour shows a sustained rise over the marker until the boundary H%. The findings of this study will constitute a general hypothesis regarding the trend expected for those tune types.

Regarding the prosodic structure and phrasing levels, as was mentioned in the literature review, this aspect would demonstrate the essential intonational groups that are signalled prosodically in the dialect. As these would show a rhythmic nature whereby alternating prominent and non- prominent (accented vs. unaccented) words and syllables, the intonation researcher would want to observe how prominence is distributed within those phrases, as well as the phonetic correlates used to mark the prominence and boundaries of those domains. In terms of prominence, languages may choose to make a word/ syllable prominent by means of stress, pitch accent, duration, segment realisation, amplitude, or a combination of those. Similarly, languages may choose to mark boundaries by tonal (tone height and direction, complex F0 configuration/ segment realisation, alignment), non-tonal boundary strength effects (duration, lengthening, juncture and pausing), or a combination of those. In addition to the previously mentioned boundary strength effects, the prosodic constituency boundaries demonstrate a number of within- domain tonal implementation phenomena reflective of the domain they represent (discussed in chapter two, section 2.4). Intonational phrases have been reported to be the domain of downstep, upstep, final lowering and phonetic declination while intermediate phrases have been reported to be the domain of pitch reset for example.

Finally, an investigation of the level of accent distribution would reveal what constitutes the designated head of the domain that a language chooses to distribute accents within. That is, languages may distribute their accents on every word making the prosodic word the domain of accent distribution where every lexically stressed syllable is realised with a prominent pitch accent (Egyptian in Hellmuth, 2006a, French in Jun & Fougeron, 2000). In this case both the pitch accent and the stressed syllable are heads of the prosodic word in EA, and head of the Accentual phrase in French. It is worth mentioning that although this tonal marking of prosodic words in EA may seem to be realised in the postlexical domain, there remains uncertainty regarding whether or not those pitch accents are also realised as part of the lexical representation of words, as the prominence marking domain of stress and accent is conflated in the dialect (Hellmuth & Chahal, 2014)[10]. On the other hand, other languages may show a relative prominence ranking of accents; a sparse distribution of accents where the final nuclear accent in the intermediate phrase is realised as the most prominent, while deaccenting or compressing other pitch events in the same domain (Lebanese in Chahal, 2001, English in Ladd, 2008, Beckman & Pierrehumbert, 1986). In this case the nuclear accented word is the head of the intermediate phrase (the last phrase before an IP boundary) and the lexically stressed syllable is the head of that prosodic word. In anticipation of the detailed discussion of focus in the next chapter, it is relevant to mention here that focus is seen as the realisation of the nuclear accent relative to the ranking of accents in languages. The nuclear accent being the most prominent unit in a phrase is equivalent to the focused word being the most important in

[10] Jun & Fougeron (2000) state that in French, APs have an invariant underlying tonal pattern and are the domain of primary and secondary stress.

a sentence in Lebanese or English. However in Egyptian, Hellmuth (2006a, 2014) shows that although every word in EA is realised as prominent by means of a pitch accent, focus is still realised by expanding the pitch range of the nuclear accent on the focused word in comparison to the rest of the words in the sentence.

In light of the discussed aspects, this chapter will propose a model of intonation representative of Jeddah Arabic, as well as touch upon the methods used to figure out and extract the desired intonation information. The sections will proceed as follows: first is a summary of the material, second will be an overview of the annotation process including the phonetic and phonological methods, and finally the proposed system of JA intonation. The forms and patterns presented in this chapter are the across- speaker canonical representations for tonal configurations observed in the dialect.

4.2 Material

Following previous studies on Arabic intonation (Al-Zaidi, 2014, Chahal, 2001, Hellmuth, 2006a) and other languages, material for the study consisted of the controlled and semi-controlled data discussed in detail in chapter 3. For the purposes of this chapter, target material and illustrations were taken from across the corpus, thus incorporating read and re-told story material, scripted speech and some utterances from the focus dialogue. Full list is presented in detail in Appendix [D]. Sample PRAAT figures will accompany the discussions about the proposed JA intonational categories. In the corpus speakers are numbered sequentially from 1-10 for females, and 1-10 for males. Figures will be captioned accordingly by speaker number and gender (F/M). The figures each contain the following information from top to bottom: a pitch trace/contour window, two word tiers in SAMPA and Unicode transcription, two phoneme tiers in SAMPA and Unicode, the syllable tier, orthorgraphy/sentence tier in transliteration, the comments tier containing tune type/information, the Tone tier containing the transcribed intonation categories, the break-index tier containing the juncture/boundary information, and finally a translation tier. Equivalent prosograms are provided in Appendix [E] for illustrative and comparative purposes.

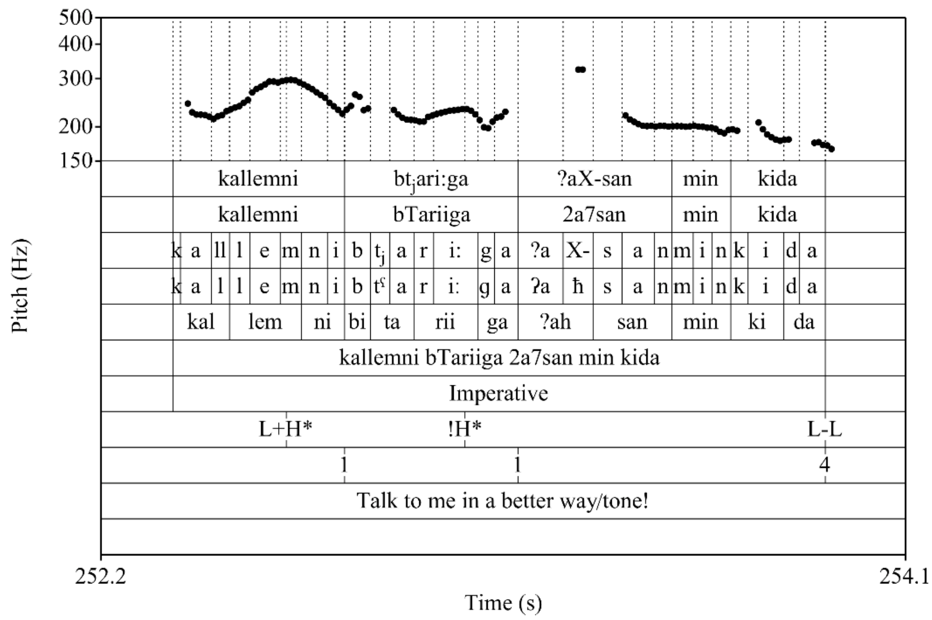


Figure 9: Imperative tune realisation from a female speaker 03_F.

4.3 Method for determining observed tonal compositions

The general criteria for identifying the tonal categories will be discussed in the following in light of previous literature on Arabic and other languages. Specifically, the guidelines in the

ToBI annotation convention (Beckman & Hirschberg, 1994), and the contextual factors guidelines in Arvaniti (2016).

- **The speaker's pitch range:**

Because the corpus contains both male and female speakers whose pitch range differs naturally, each speaker's range was taken into consideration while marking a specific tonal event. Tonal events within a contour make reference to the bottom of the speaker's pitch range, named as the *Baseline* in early Pierrehumbert literature.

- **The Pitch contour:**

Whether the F0 track was overall rising or falling throughout an utterance. Pitch track manner also reveals how/whether a particular tone is targeted and presents evidence for interpolation. This was observed while also taking into consideration any disfluencies caused by microprosdic factors. F0 perturbations causing F0 to fluctuate or disrupt during an utterance include: inherent segment class, such as obstruents causing F0 to dip at closure and rise at release, different phonation types such as glottalisation and creaky voice causing F0 to disrupt or disappear, etc.

- **Pitch configuration at phrase edge:**

Global pitch range was observed in prosodic phrases according to the different syntactic phrases (interaction between prosodic structure and syntactic structure will be discussed in the upcoming section 4.7.1). The manner demarcating the edge of these phrases was marked, e.g. an L- where pitch falls in a phrase, or an H- where it rises for an intermediate phrase edge. Similarly, an L%, or an H% where pitch falls or rises for an Intonational phrase boundary. The configuration is usually observed on the last syllable of the nuclear accented word in ip or IP. This is taken by Chahal (2001) as evidence for a secondary association of phrase tones to nuclear accented words.

- **Pitch configuration around stressed syllable (local):**

Observed tone location: 1) a configuration/movement on a lexically stressed syllable indicates a pitch accent i.e. prominence lending; 2) configuration/movement not on stressed syllables (on edges of phrases) indicates an edge tone (phrase tone or boundary tone) i.e. not prominence lending and cued by sense of juncture and/or a pause.

- **Position of observed tone in phrase:**

The location of an accent in a phrase demonstrates its status according to the prominence hierarchy presented earlier in the focus section. If the accent is final in a phrase accompanied with a visible compression of pitch range, it bears the nuclear status, whereas a non-final accent bears a pre-nuclear status.

- **Phonetic realisation of accent:**

Accents will demonstrate different shapes depending on their location within a phrase (initial, medial, final) and surrounding segmental context, which causes changes to their phonetic identity. These phonetic changes include: the height of an accent 'scaling', 'downstep' and 'upstep', as well as its timing with respect to the stressed syllable 'alignment'.

- **Interpolation:**

Interpolation is the phonetic rule for the transition between two tone targets. In the theory, the intonation contour is made up of a sequence of two F0 targets occurring at specific points: H and L representing the prominent turning points throughout the course of a contour (Pierrehumbert, 1980). The remaining transitional pitch between any two targets is computed via interpolation.

In light of these guidelines and observations, the phonetic component in the theory is manifested via the following implementation rules: Downstep, Upstep, interpolation and declination. Downstep is the process affecting the scaling (level) of an H tone and is represented by adding a (!) diacritic in front of an accent. An accent is scaled in level compared to the previous accent heights, and this process is phonologised in some languages. In English the presence of a bitonal accent triggers the downstep of the following H accent as reported in Pierrehumbert (1980). However, in Arabic as reported by Chahal (2001) a downstepped accent is a voluntary choice that is not triggered by such factors. In other words, a downstepped accent may or may not be preceded by a bitonal accent. Upstep, on the other hand, is the process affecting local pitch range whereby the tone following a H- or !H- phrase tone is raised to the same or higher level as the phrase tone. This takes place in boundary tones of the type: H-H%, H-L% or !H-L%. The high phrase tone raises the sequential tone to a level that is higher (H-H%), or plateau (H-L%, !H-L%). The plateau level we observe in the latter boundary types would explain why these are substantially different from a true high H-H% or low boundary L-L% (Chahal, 2001). Interpolation as was mentioned earlier controls the pitch transition between tones in a contour. Finally, declination is the process affecting accents throughout a declarative neutral tune. The height of successive accents in this tune would demonstrate a gradual descending trend leading to the major boundary of the utterance (Hellmuth, 2006a, Pierrehumbert, 1980).

4.4 Main intonational Tunes

This section presents the main intonational tunes observed in JA. The description of a tune takes into consideration the boundary configuration along with the pitch accents and nuclear accents. The pitch contours of these tunes are generated using Praat. The decided compositions displayed in this chapter are the most common patterns established by obtaining the frequency of the occurrence of the composition of a tune (nuclear accent+ boundary tone) in the corpus. Any further variations are discussed accordingly in the respective section.

Target tune	**Boundary tone**	**Nuclear accent**	**% of occurances**
Declarative	L-L%	H*, !H*, L+H* and L*	52.57%, 27.93%, 12.79% and 6.71%
Yes/no question	H-H%	L+H*	91%
WH question	!H-H%	H*	75.23%
Request	H-H%	L+H*	98.30%
Imperative	L-L%	!H*	61.90%
Plateau 'uncertainty	H-L%, !H-L% and L-H%	H*, !H* and L+H*	46.66%, 46.66% and 6.66%
Continuation	H-H%	L+H*	56%

Table 4: The most common tune compositions for the proposed tunes in the current JA corpus

4.4.1 Declarative Tunes

Declarative sentences in JA display a falling edge configuration that ends with an L-L% boundary combination. This edge boundary can be combined with different pitch accents, including high H*, downstepped !H*, low L*, or rising L+H* accents. This falling manner consequently causes some high accents to gradually decrease in height closer to the boundary in the phenomenon known as phonetic declination. This low boundary combined with a H* nuclear tone in 52.57% of the data, an L* nuclear accent occurred in 6.71% of the data, an L+H* as nuclear accent ocuurred in 12.79% of the data, and a downstepped !H* occurred as the nucler accent in 27.93% of the declarative tune sentences. Noticeably, the common nuclear accent choice in the data seems to be in preference of a high tone. It thus can be initially proposed that that the overall nuclear accent for declarative tunes is a high tone, with around 93% of the occurrances underlyingly ending the declarative in high tone. The following are examples from the corpus:

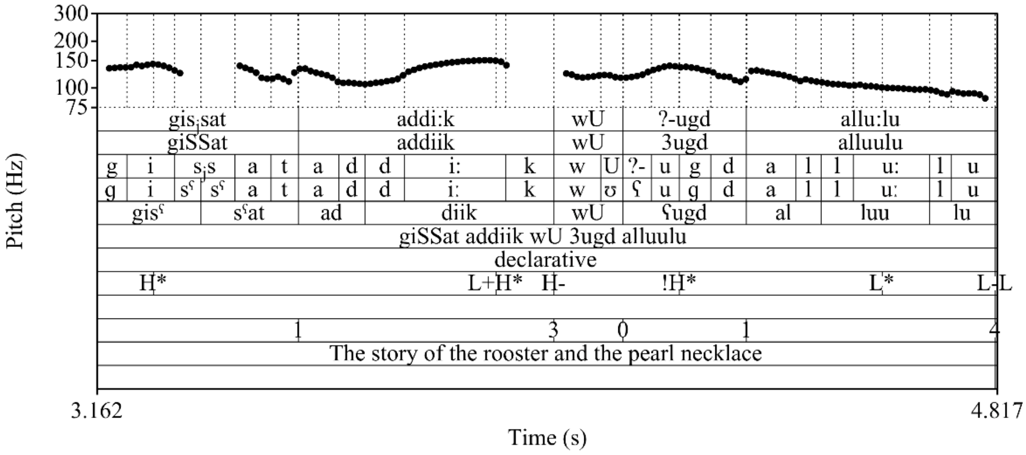


Figure 10: Declarative tune in narration displaying declination. Stressed syllables underlined. Speaker 01_M

gissat 2addiik	wu	3ugd 2alluulu
the story of the rooster	and	the pearl necklace
H* L+H* H-		!H* L* L-L%

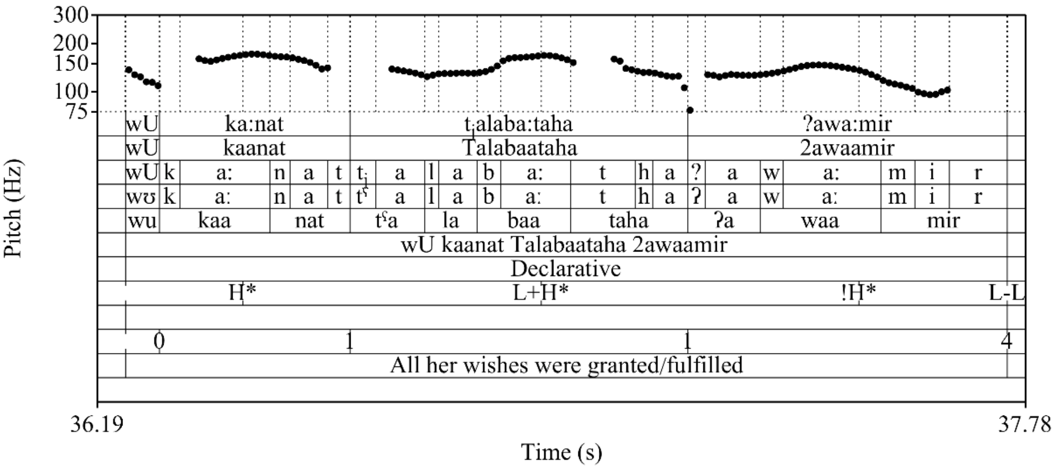


Figure 11: Declarative tune in narration. Speaker 03_M

wu kaanat	Talabaataha	2awaamir
and all	her wishes	were commands
H*	L+H*	!H* L-L%

4.4.2 Question Tunes

The corpus contains both types of rising question tunes: WH questions and Yes/No questions. WH question tunes were elicited by using a WH question word and ending the utterance in a question mark, while for yes/no questions, a question mark was suffice in prompting speakers to produce a question tune. During the recording, the participants were presented with slides containing maximally 5 utterances, and WH and yes/no questions were randomised in order to prevent them occurring on the same slide. These measures were taken as a precaution to avoid “list intonation” effect. The dialect displays a difference in the marking of each question type. Both types end in a high rising boundary tone H-H%, however the nuclear accents, as well as a difference in pitch range distinguish the two types, along with a categorical difference in meaning. The examples in the corpus see final words varying in stressed syllable position (stressed syllable is medial or final), as well as varying in the length of sentences. Yes/no questions were comprised of 3 word and 1-word sentences, which helps validate the true nuclear accent for this sentence type when the sentence is shorter (cf. /Darabatu/ in the upcoming examples). The nuclear accent in yes/no questions is a rising L+H* followed by a distinctively expanded range towards the boundary. This combination occurred in 91% of the yes/no question data, while the other 9% of the examples combined an H* or an !H* as nuclear accents with this high boundary.

For WH questions, the question word plays an important role regardless of its position within the sentence (initial or medial in a sentence). A simple monotonal H* constitutes the nuclear accent of WH questions that is followed by a much compressed pitch range towards the boundary, which surfaces as a level trendline. This combination occurred in 75.23% of the compositions of this WH question tune, while the other 24.77% percent surface an L+H* nuclear accent with this downstepped edge tone. When the WH question word is in initial position as in figures [17-18] below, it is allocated the nuclear accent and followed by a monotonous stretch of compressed pitch level until the boundary. This stretch is evidence for a downstepped phrase tone !H- that is stretched over the material until the boundary. When the question word is in medial position [Figure 19], it is also allocated a nuclear accent and followed by a downstepped edge. The final rise in the WH question tunes is categorically different from the final high rise edge configuration found in the yes/no tune. Moreover, the WH edge never surfaces as a true fall compared to an L-L% declarative edge where F0 falls to a true low even following a H* nuclear accent. Therefore, the two edge configurations were excluded from the WH tune analysis and a downstepped edge is proposed. Note the following examples:

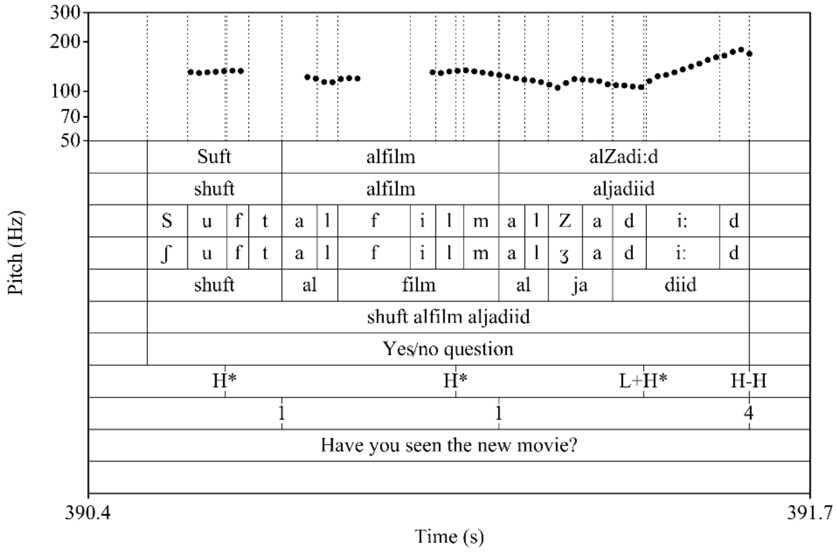


Figure 12: yes/no question. Nuclear accent on final syllable. Speaker 05_M

shuft 2alfilm 2aljadiid

did you see the film new

H* H* L+H* H-H%

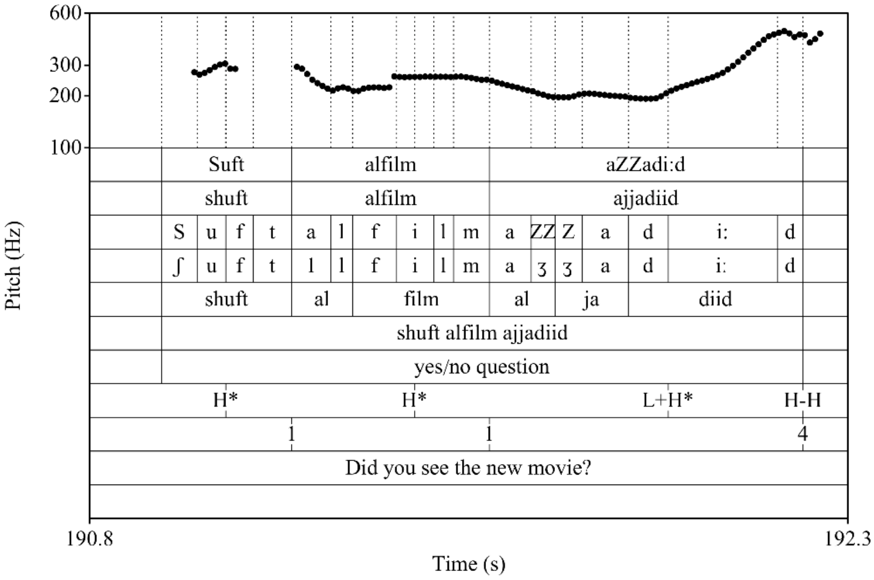


Figure 13: yes/no question. Nuclear accent on final syllable. Speaker 07_F

shuft 2alfilm 2aljadiid

Did you see film the new

H* H* L+H* H-H%

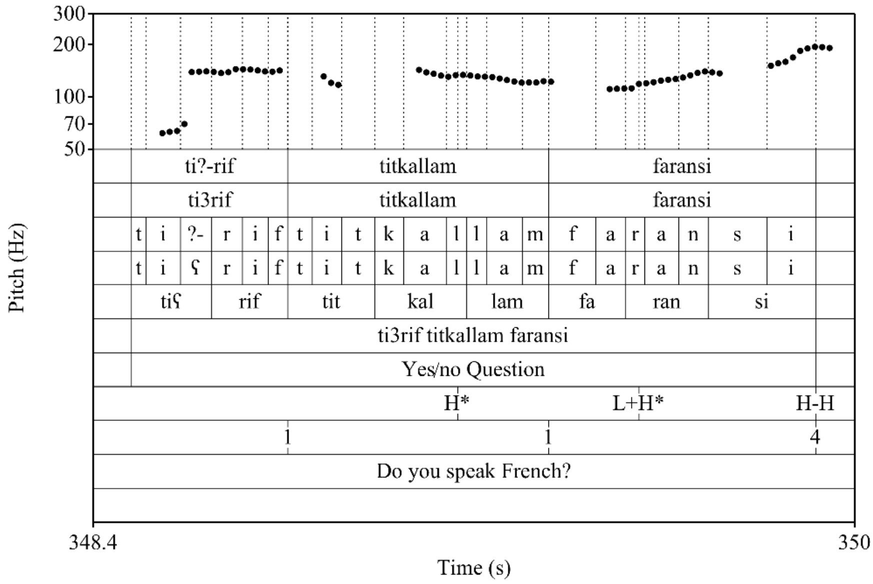


Figure 14: yes/no question. Nuclear accent on medial syllable. Speaker 07_M

ti3rif titkallam faransi

Do you speak French

H* L+H* H-H%

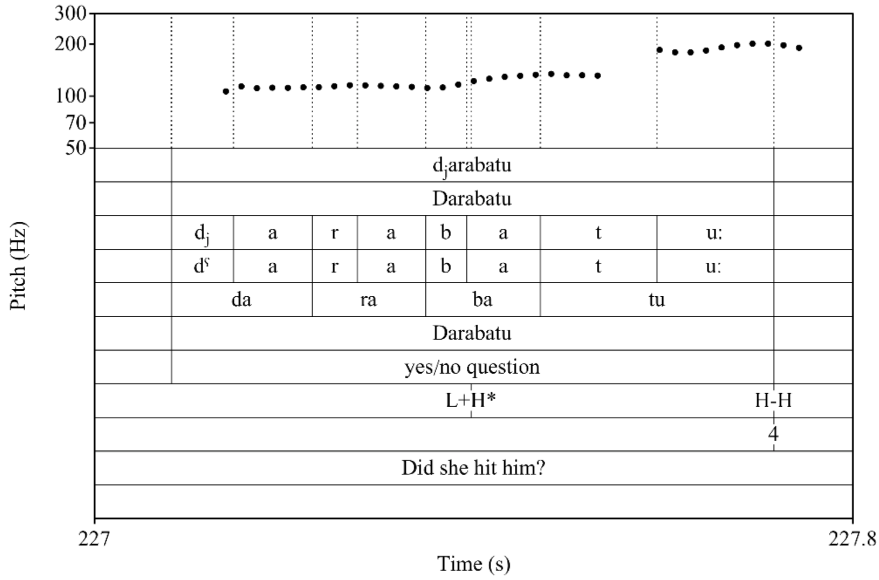


Figure 15: yes/no question. Nuclear accent is on medial syllable in one-word sentence. Speaker 10_M

Darabatu

Did she hit him

L+H* H-H%

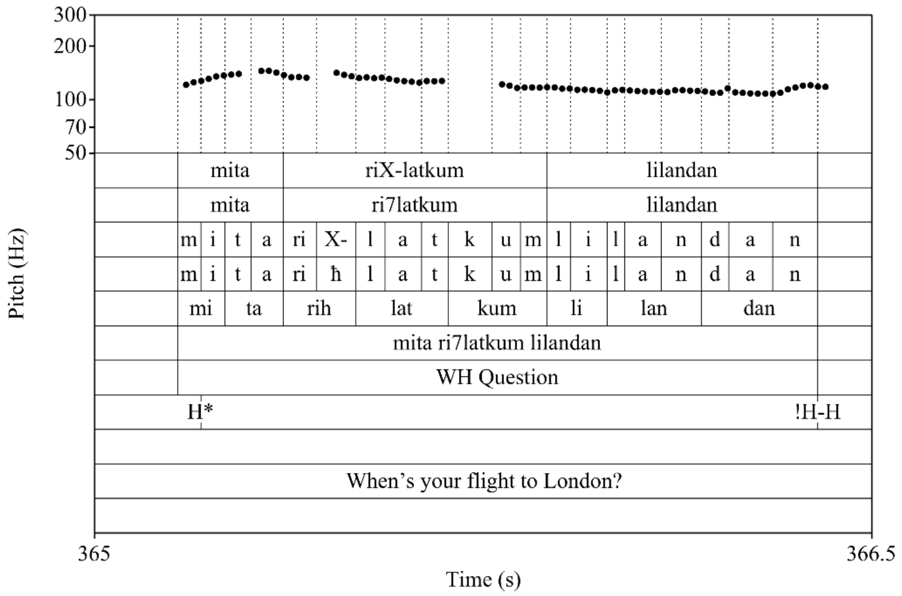


Figure 16: WH question. Question word in initial position. Speaker 02_M

mita riHlatkum lilandan

when's your flight to London

H* !H-H%

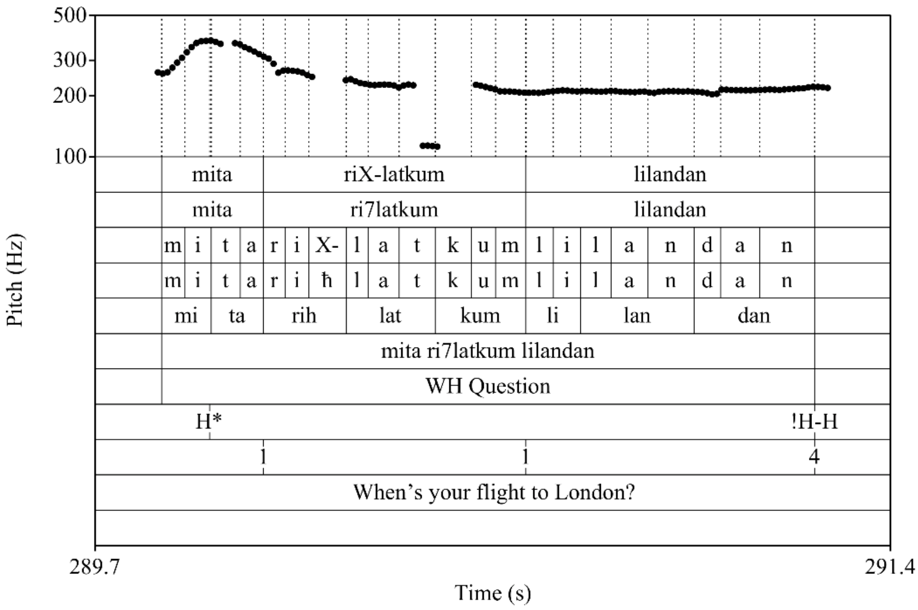


Figure 17: WH question. Question word in initial position. Speaker 06_F

mita riHlatkum lilandan

when's your flight to London

H* !H-H%

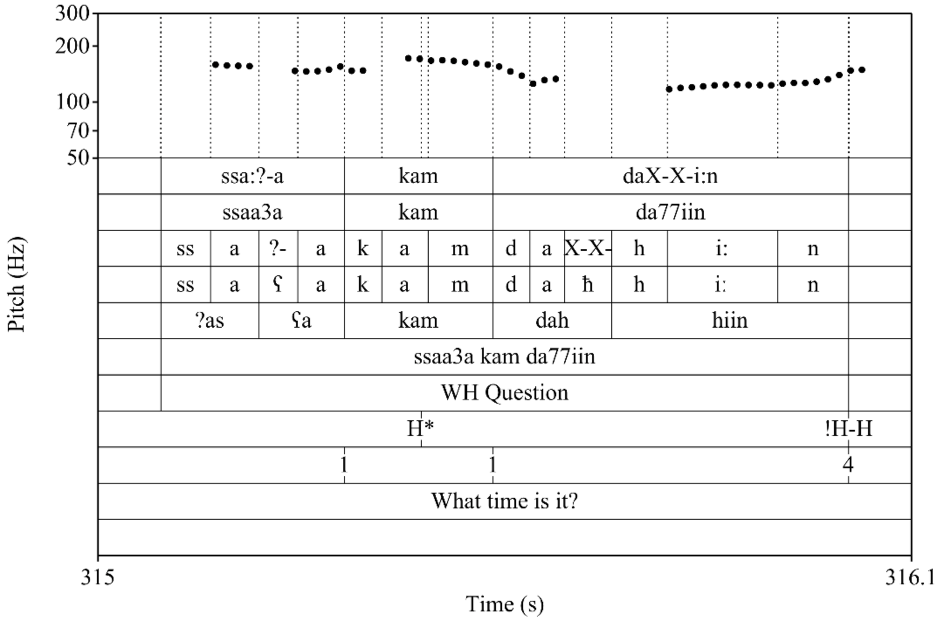


Figure 18: WH question. Question word in medial position. Speaker 03_M

2assaa3a	kam	daHHiin
The time	what's	now
	H*	!H-H%

4.4.3 Request Tunes and imperatives

In addition to the core utterance tunes demonstrated above, the corpus also included other tunes, such as polite requests and imperatives. The request sentence in the corpus contains the polite marker word /mum.kin/ that means both 'please' and 'can you' indicating to the speakers to use a polite tone. This polite tune was produced as a "polite" question by speakers, which demonstrates a similar pattern to a yes/no tune ending in a rising boundary. The speakers may be able to distinguish between these two tunes only depending on the existence of the polite marker- otherwise they are similar. The request tune in the corpus only ocurrs with 'please', as it is very uncommon to express requests without 'please' in the dialect. A discriminatory investigation that synthesises the two: yes/no vs. request constructions by excluding the lexical information of sentences may yield further differences; however at this stage, the difference between the two tunes can be said to be syntactic and semantic rather than intonational. Speakers generally marked a request tune with a high H-H% boundary alongside an L+H* nuclear accent, which occurred in 98.30% of this tune's data. Less than 2% of the data surfaced with an H* nuclear accent.

In a similar vein, imperatives across the corpus are intonationally marked similar to declarative tunes, whereby the semantic function of the tune is encoded in the choice of wording. Speakers marked imperatives with a low L-L% boundary combined with a downstepped !H* nuclear accent in 61.90% of the data. Around 40% of the occurances varied in the use of the nuclear accent between L+H* and H* alongside the L-L% boundary to mark this tune.

More generally, neither pair, the yes/no questions and requests, nor the declarative tunes and imperatives show major differences in the global trendlines of the F0 contours as a whole. This could be due to the length of sentences. A further analysis incorporating longer stretches of material may be better able to demonstrate their differences. Note the following examples for request and imperative contours:

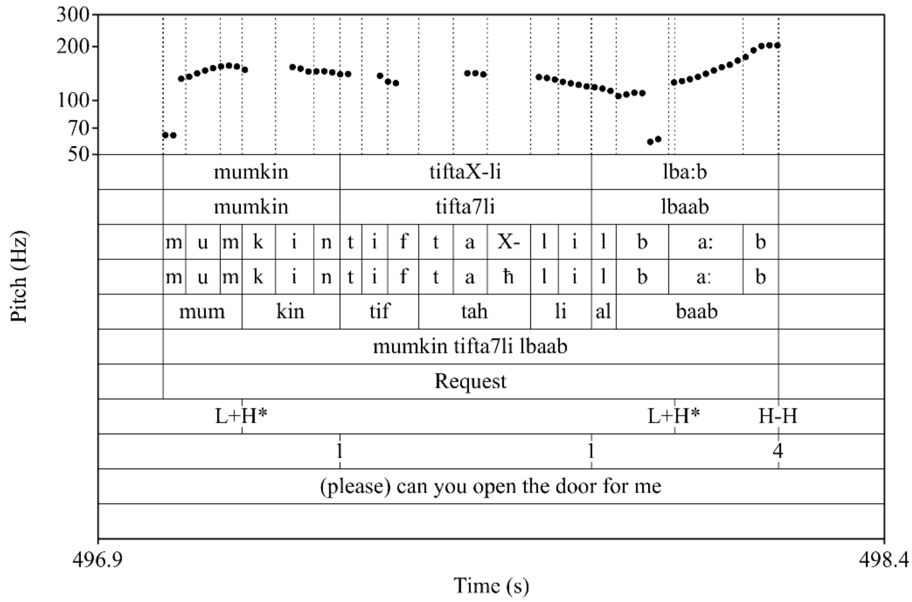


Figure 19: Polite Request tune. Speaker 10_M

mumkin	tiftaHli	albaab
'please'	open for me	the door
L+H*		L+H* H-H%

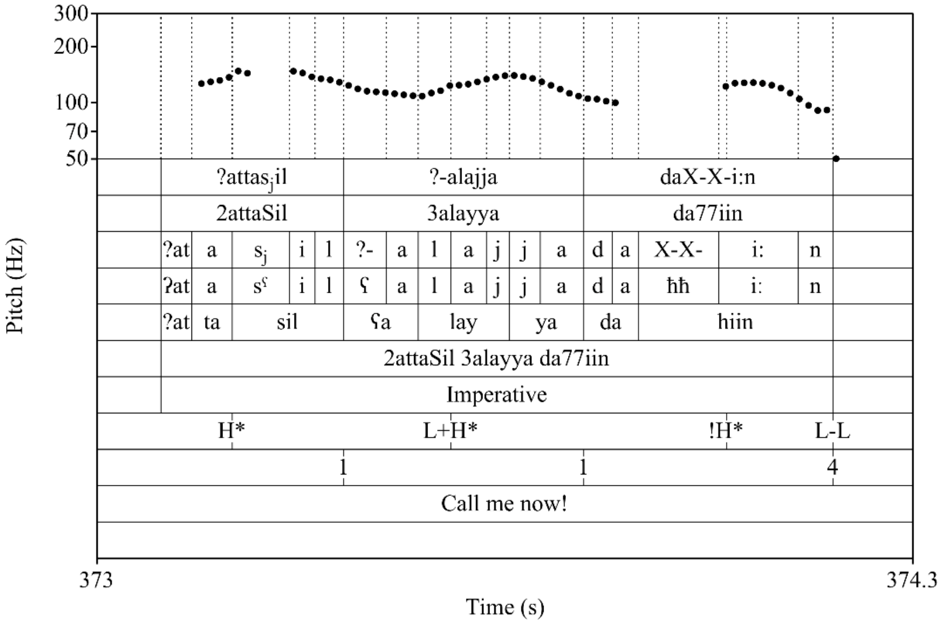


Figure 20: Imperative Tune. Speaker 04_M

2attaSil 3alaya daHiin

Call me now

H* L+H* !H* L-L%

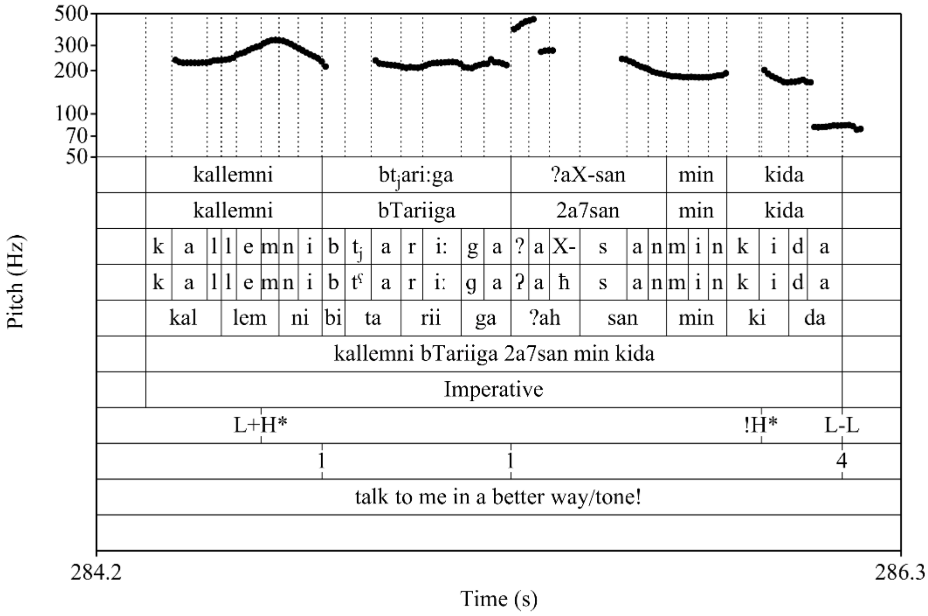


Figure 21: Imperative Tune. Speaker 03_F

Kallemni biTariiga ?aHsan min kida

Talk to me in a way better than this

L+H* !H* L-L%

4.4.4 Continuation and (uncertainty) tunes

The corpus contains a considerable number of continuation and plateau tunes to mark incomplete utterances, i.e. non-finality and uncertainty. An edge configuration of a falling H-L% and H* nuclear accent occurred in 46.66% of the data, an !H-L% and a downstepped !H* nuclear accent in the other 46.66%, while a rising L-H% and an L+H* nuclear accent 6.66% with pitch notebly remaining at the middle of a speaker's range can mark the end of these tunes alongside the different nuclear accents. Continuation tunes on the other hand are marked by a high rising edge H-H% combined with a high H* in 18.66% of the data, downstepped high !H* in 25.33% of the data, or a rising L+H* pitch accent in 56% of the continuation tune compositions. Note the following examples:

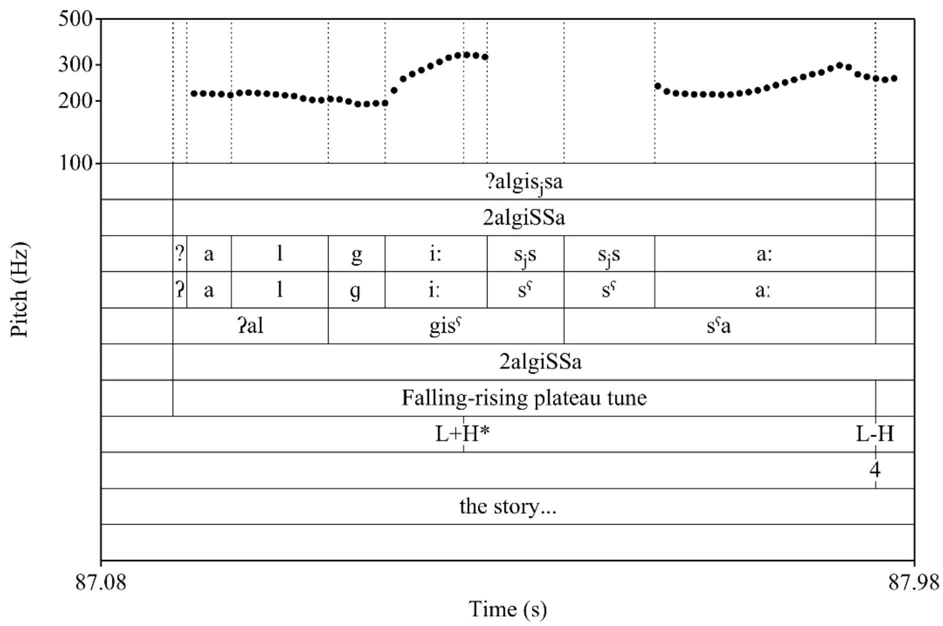


Figure 22: Rising plateau tune in narration. The L phrase tone stretches following from the pitch accent until the edge of the the phrase where it rises again. The prosogram in appendix [E] provides clearer targets. Speaker 03_F

2algiSSa…. 2innu kaanat fii 2amiira

The story [cont….] that there was a princess

L+H* L-H%

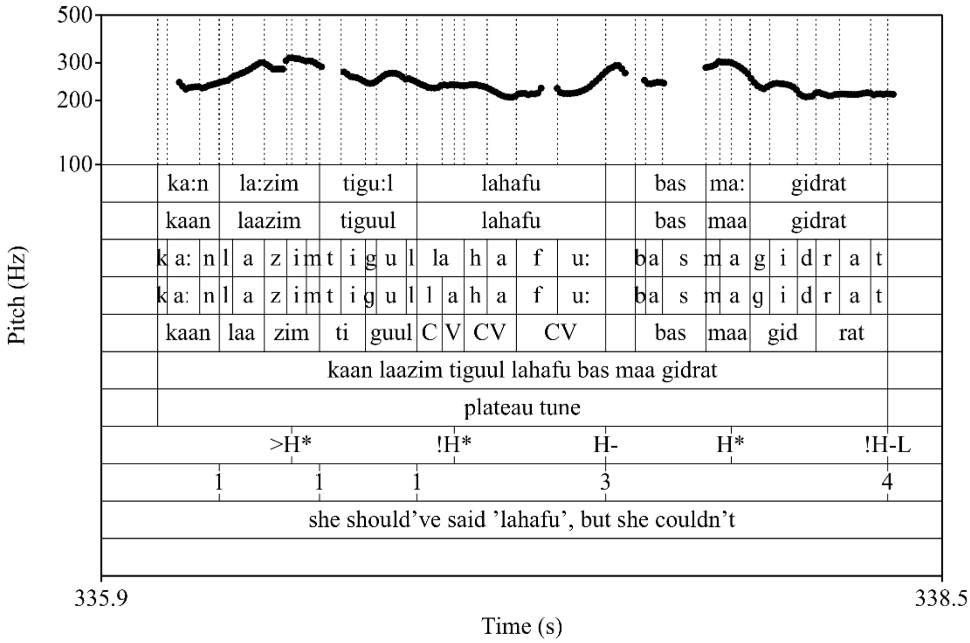


Figure 23: Falling plateau tune. Speaker 07_F

kaan laazim tiguul lahafu, bas maa gidrat

she should've said 'lahafu', but she couldn't

H* H* !H* H- H* !H-L%

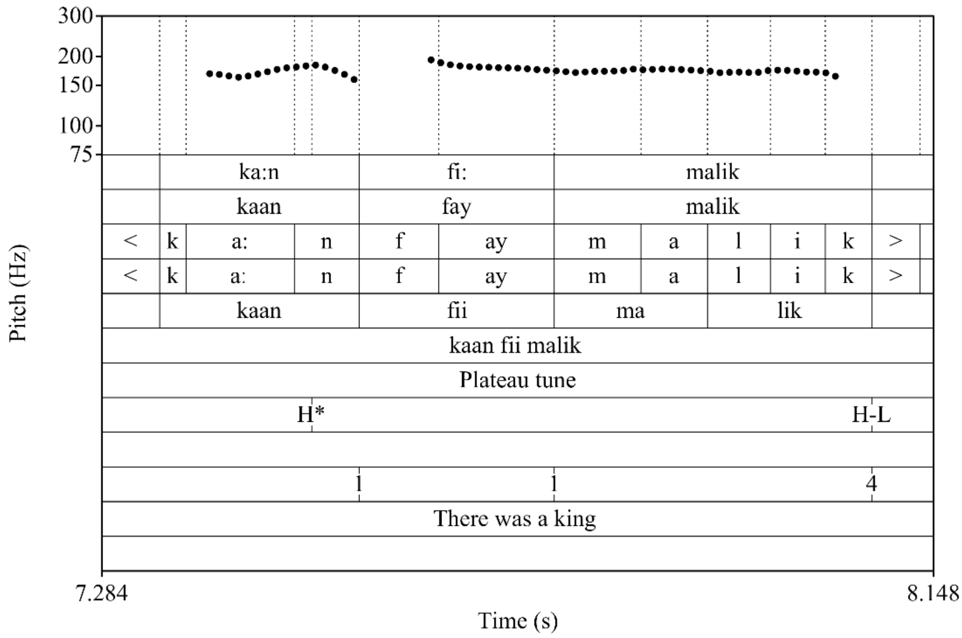


Figure 24: Falling plateau tune in narration. The prosogram in appendix [E] provides clearer targets. Speaker 09_M

kaan fii malik

there was a king

H* H-L%

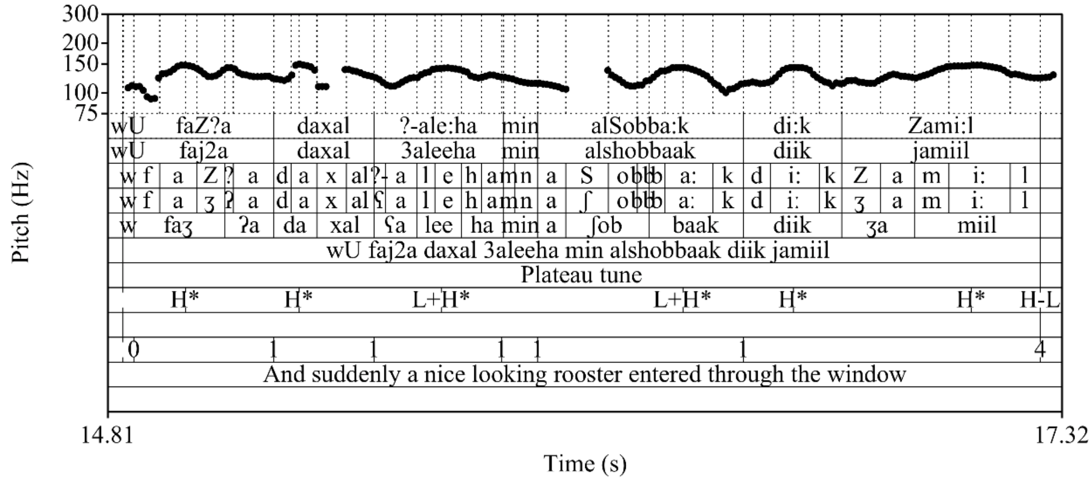


Figure 25: Falling plateau tune in narration. Speaker 03_M

faj2a daxal 3aleeha alshobaak diik jamiil

suddenly entered from the window a rooster pretty

H* H* L+H* L+H* H* H* H-L%

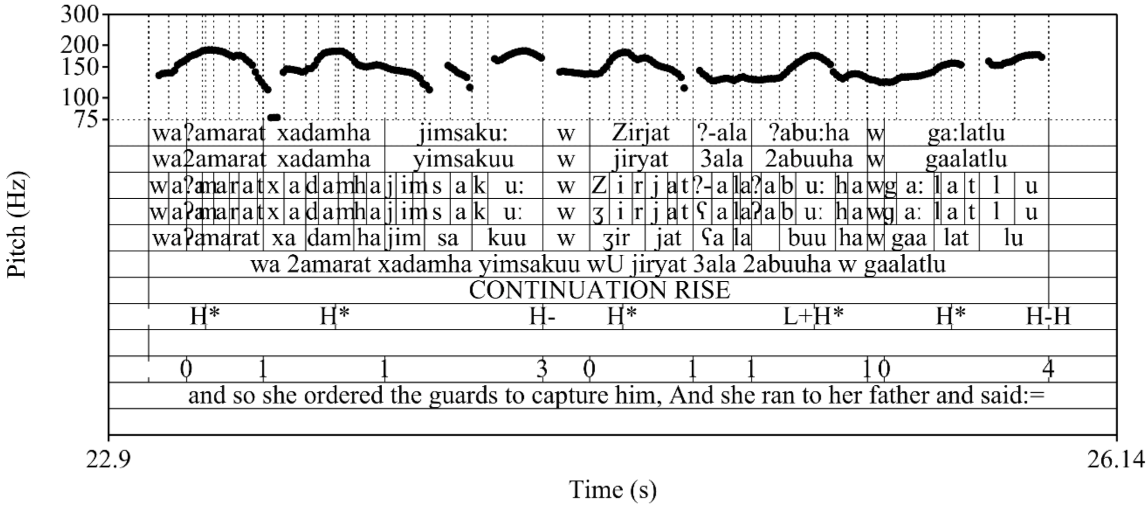


Figure 26: Rising continuation tune in narration. Speaker 05_M

yimsakuuh wu jiryat 3ala 2abuuha wu gaalatlu

hold him and she ran up to her dad and said:

L+H* H- H* L+H* >!H* H-H%

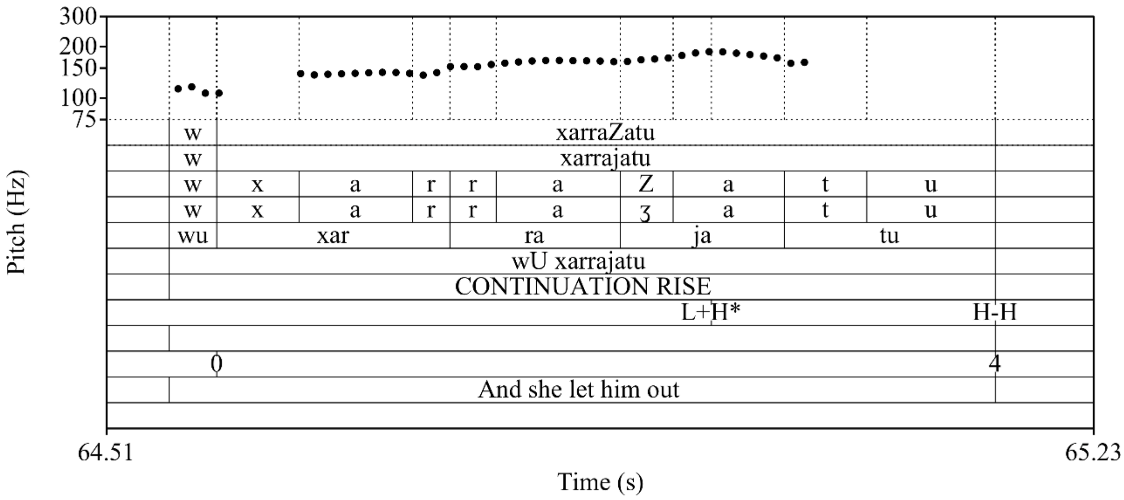


Figure 27: Rising continuation tune in narration. The prosogram in appendix [E] provides clearer targets. Speaker 09_F

wu xarraja̲tu:

and she got him out:

L+H* H-H%

4.5 Pitch Accent inventory

For pitch accents, JA shows H*, L*, !H*, L+H* tones, as the case with the other Arabic dialects reported in this thesis. For phrasal tones, it demonstrates H-, !H- and L- as demarcating the edge of ip. The existence of these edge tones is further evidence for the existence of the ip as a prosodic constituent at the level of phrasing, as those edge tones constitute the tonal correlates for ip edges reported for Arabic along with the tonal alignment patters. It is worth noting that an H- tone occurs in Lebanese, but is absent in Taif Hijazi (Al-Zaidi, 2014). For boundary tones, both L% and H% boundary tones demarcating ends of IPs were observed in JA data, which was also taken as evidence for the existence of the IP as a prosodic constituent. Note the following summary table of the tones and their observed F0 patterns:

Tone	Specification
H*	F0 peak in the mid-upper speaker range.
!H*	Downstepped high accent.
L*	F0 valley in lower region of speaker range.
L+H*	F0 low-high movement where the peak aligns within syllable rhyme and valley aligns at syllable onset.
L-	ip low phrase tone.
H-	ip high phrase tone.
!H-	Downstepped high phrase tone.
L%	IP boundary: F0 contour falls markedly.
H%	IP boundary: F0 contour rises markedly.

Table 5: F0 patterns in Jeddah Arabic.

4.5.1 Pitch accents

4.5.1.1 H* and !H*

H* is a monotonal, unmarked tone in the dialect. It is realised as a peak in the highest of speaker's pitch range. The alignment of this tone with respect to the accented syllable varies according to its location in an intonational phrase. Before an IP boundary (nuclear position) this tone is pushed as leftward as possible in the accented syllable, i.e. the more final the syllable, the earlier the alignment, while non-finally it is realised within the span of the vowel (more on this is discussed in detail in the alignment chapter). The height of this tone also varies according to the position of the tone in a sentence. An H* occurring phrase initially, for example, is higher than the last H* accent in a declarative (declining) tune. This accent has a downstepped counterpart !H* that is scaled to a lower level peak than a plain H*. A downstepped accent is realised in the middle part of the speaker's range.

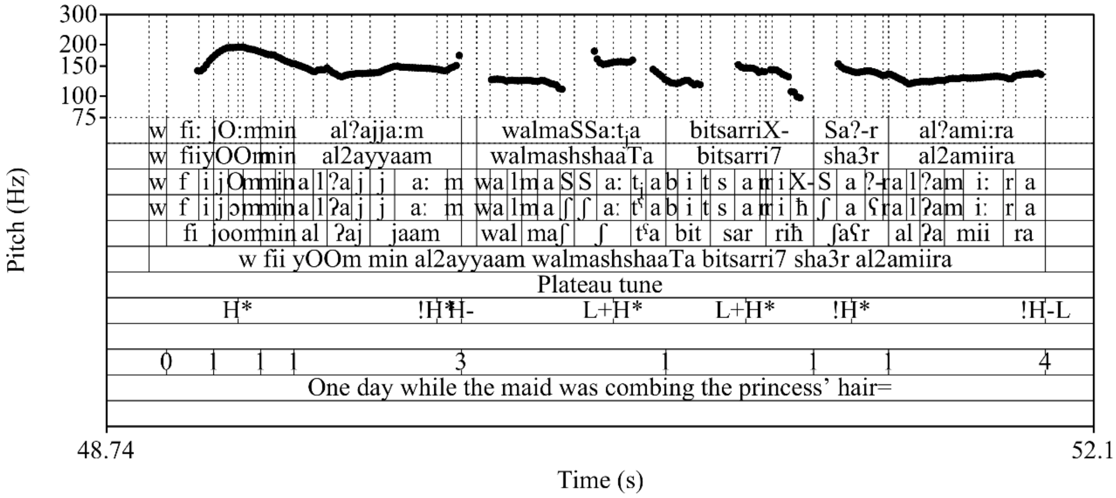


Figure 28: Utterance showing a plain H* on /joom/ 'day' where contour rises steadily to reach the peak, followed by a downstepped !H* on /2al2ajaam/ 'the days' and /sha3r/ 'hair'. Speaker 01_F

4.5.1.2 L+H*

The second most common and the only bitonal accent type in the dialect. This accent is comprised of a leading L tone that starts at or just before the onset of the accented syllables, followed by a high peak that is realised within the rhyme of that syllable. These two tones exhibit a close timing coordination in that they occur consecutively on the same accented syllable. The L starts at the lower pitch range and the H is realised as a sharp turning point at the higher ranges. Evidence that L is targeted is when the pitch falls considerably towards a low pitch range at or just before the onset of accented syllable. A few occurrences of a downstepped phonetic counterpart of this accent L+!H* were also observed in the dialect, but they were not enough to make it part of the accent inventory. Note the examples for those accents in the following figures:

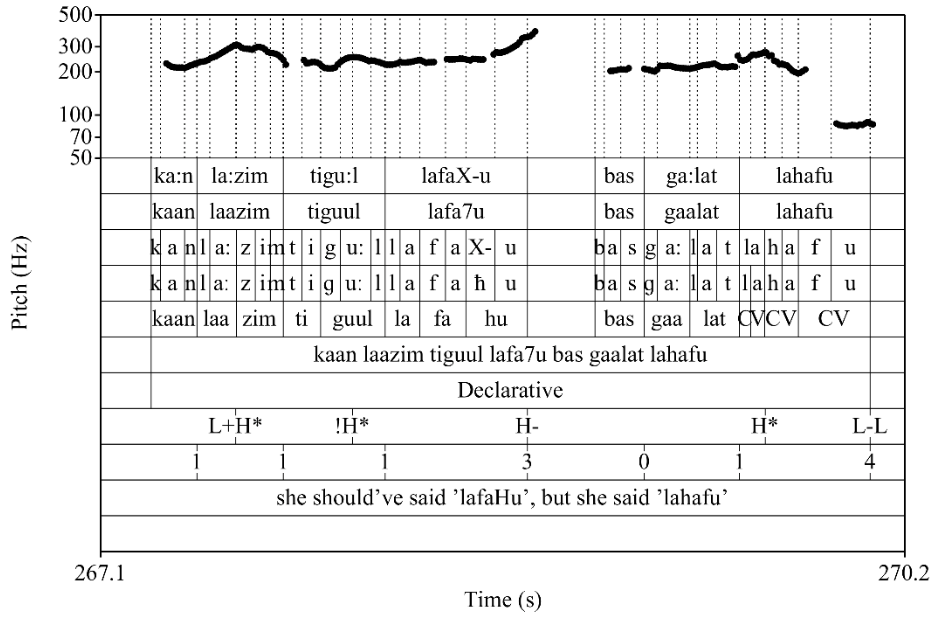


Figure 29: L+H* on /laazim/ 'necessary'. Speaker 08_F

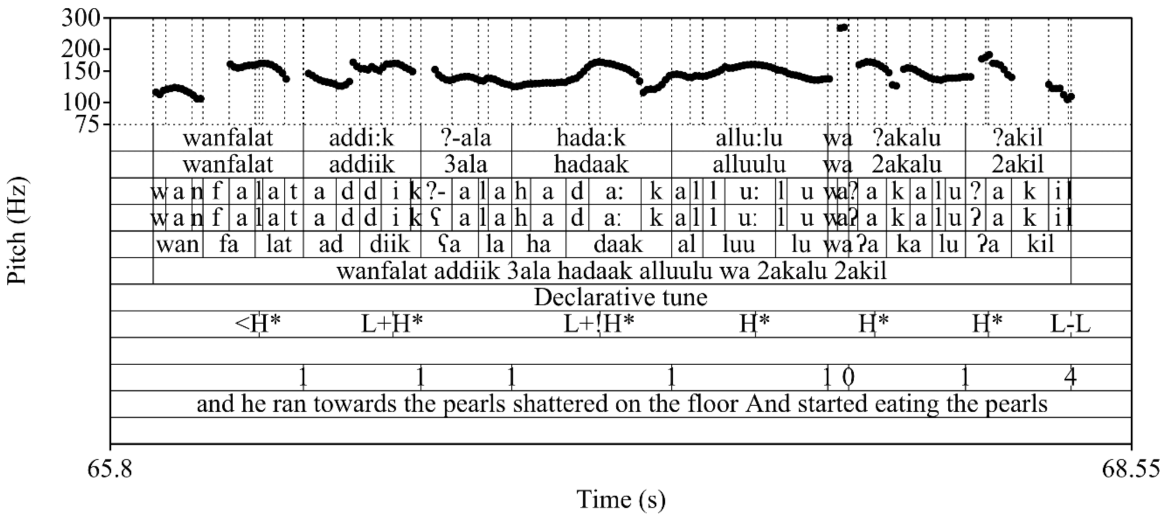


Figure 30: Utterance showing a downstepped L+!H* on /hadaak/ 'that' in narration. Speaker 01_M.[11]

*4.5.1.3 L**

Monotonal accent realised at the lower range of the speaker's pitch range. It occurs most commonly as a monotonic low stretch spanning the whole accented syllable, and it occurs at

[11] One could also argue that the accent on /hadaak/ is a plain L+H* followed by a L- phrase tone, however on the basis of vowel and syllable duration (the ip boundary non-tonal correlates), as well as auditory juncture and pitch range of the following word, it is argued that a downstepped accent choice may seem more plausible.

the lowest level finally in an utterance but scaled higher non-finally. This accent is absent phrase medially.

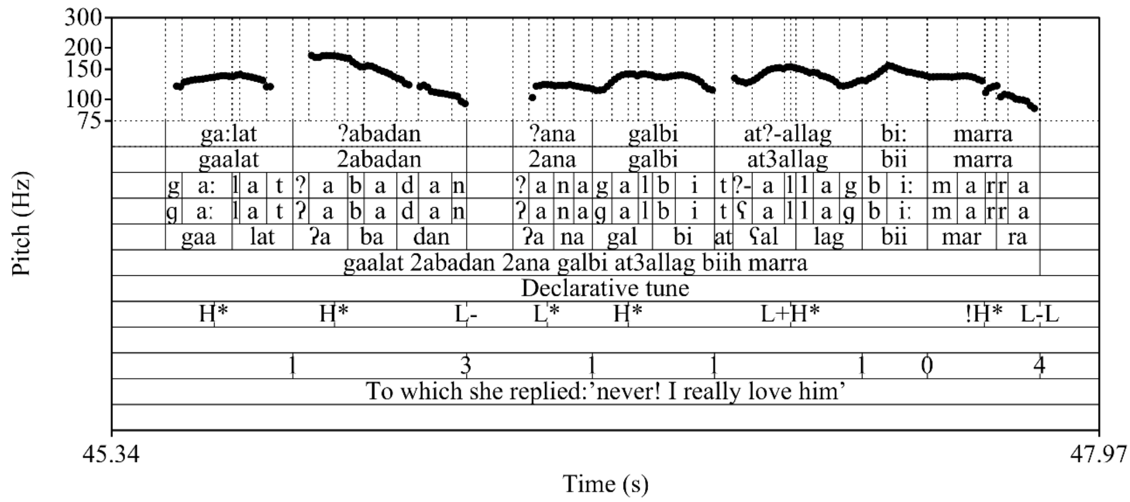


Figure 31: Declarative tune in narration with L* on /2ana/ 'I'. The prosogram in appendix [E] provides clearer targets. Speaker 05_F

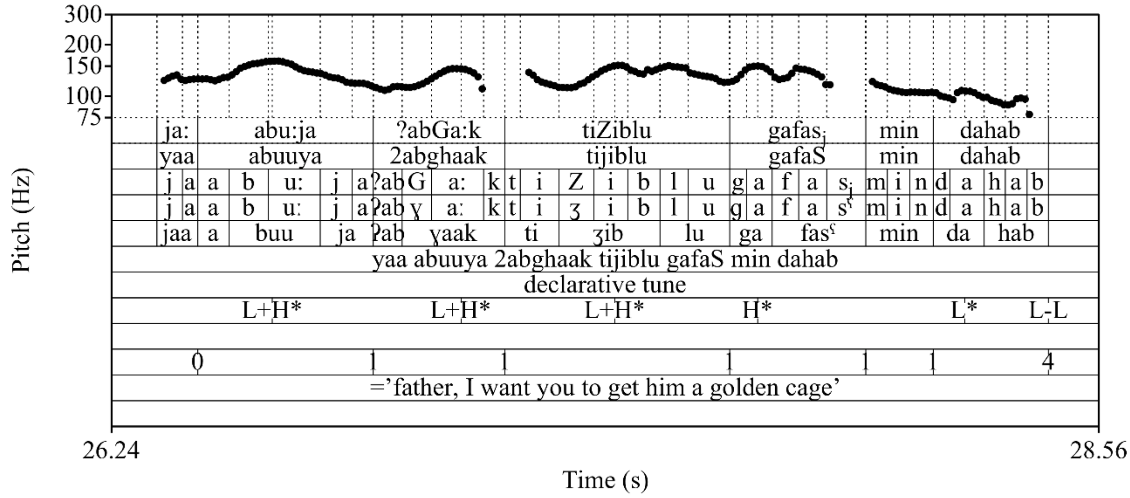


Figure 32: Declarative tune in narration with L* on /dahab/ 'gold'. Speaker 03_M

4.6 Edge tones (phrase accents and boundary tones)

4.6.1 Phrase accents

The corpus also included phrase accents delimiting the right edge of intermediate phrases. The occurrence of a phrase accent was determined tonally via the pitch configuration at certain points in the contour. This configuration is different from a pitch accent in that it occurs on word edges and not on lexically stressed syllables. Also, this configuration is auditorily recognisable via a sense of juncture and an infrequent silent pause. Final lengthening is also observed before those phrase accents, however to a different degree than what would be observable before IP boundary tones. Alongside these characteristics, phrase tones occurring at edges of ips may contribute in controlling and resetting the pitch of the following material. In long-tailed sentences, the phrase tone may be separated from the boundary tone by a number of accented or unaccented material, especially if the nuclear accent occurs early in the sentence. For example, in figures [34] and [35] below, it can be seen how the material following the phrase tone pitch configuration starts at a new pitch register. In figures [36] and [37], the pitch level of the material following the phrase tone is controlled via gradient pitch range manipulation.

4.6.1.1 H-

A high phrase accent realised in the mid-upper pitch range, where high pitch is sustained following a pitch accent.

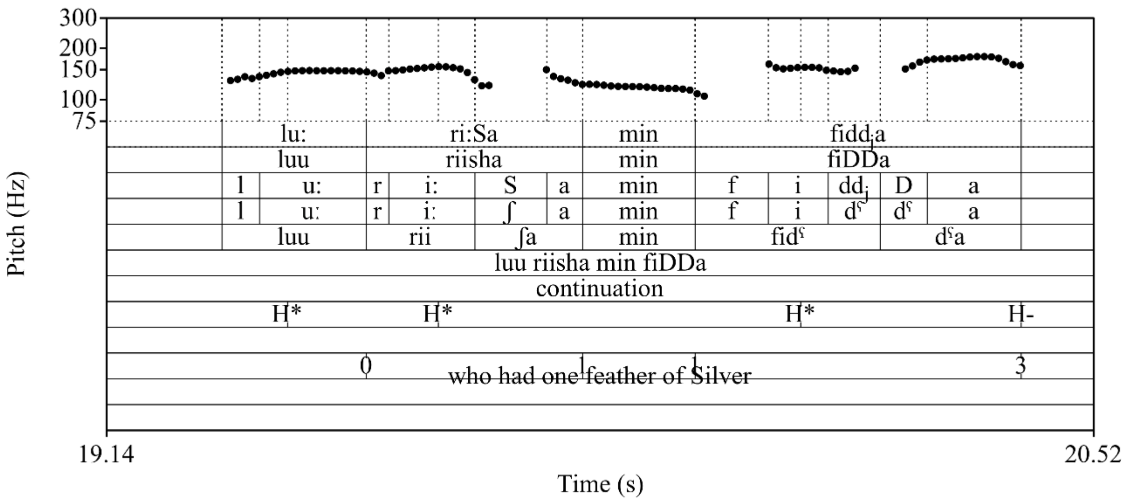


Figure 33: High phrase tone on right edge of the word /<u>fiD</u>Da/ 'silver' in narration followed by an audible and acoustic pause. Speaker 09_M

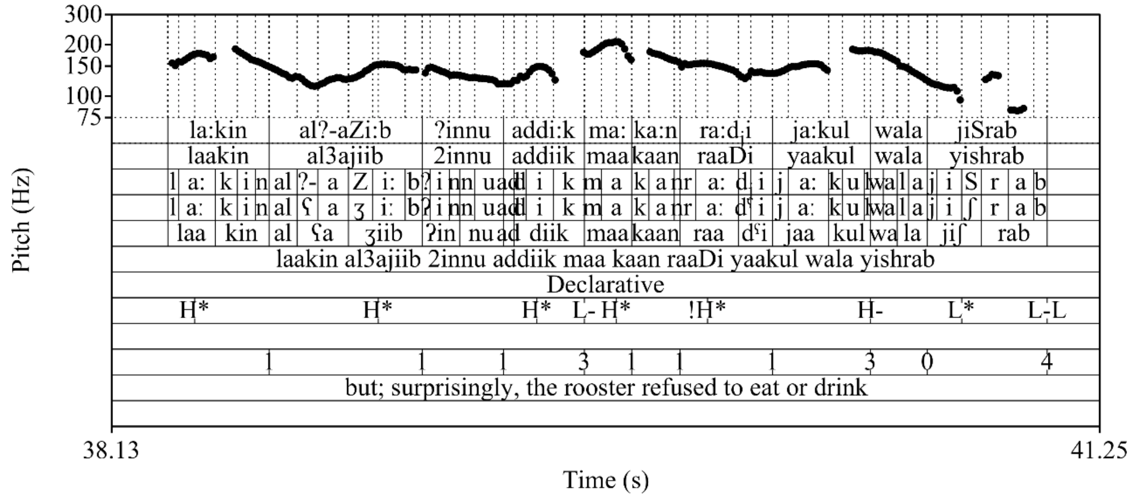


Figure 34: High phrase tone on right edge of the word /jaakul/ 'he eats' in narration. No pause is present in this case, but an auditory sense of juncture. Speaker 01_M

4.6.1.2 !H-

This phrase accent is different from the previous in terms of pitch range. It is realised as a high plateau configuration in the middle of a speaker's pitch range.

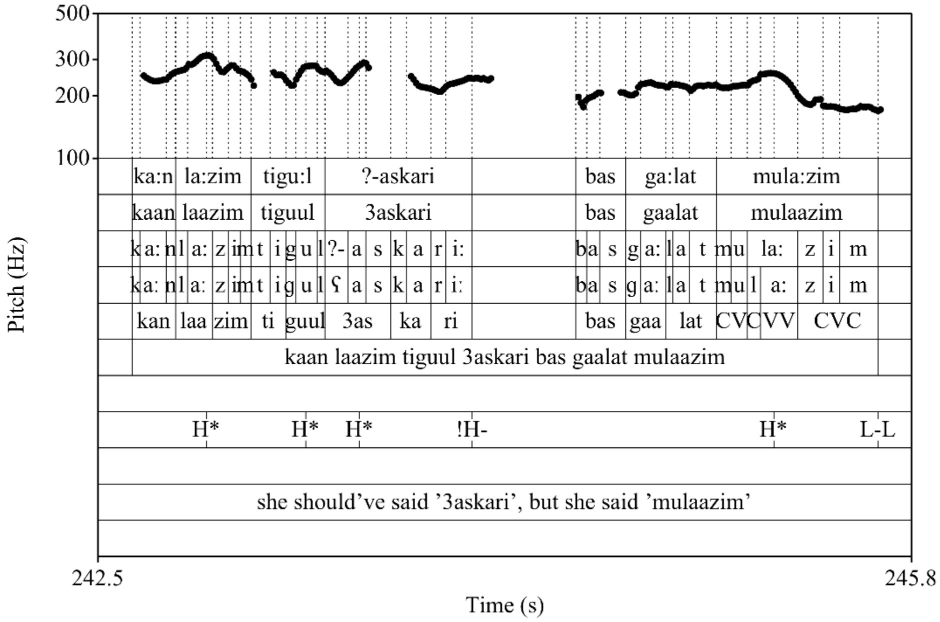


Figure 35: Downstepped phrase tone !H- on right edge of the word /3askari/ 'military officer'. A high accent on the lexically stressed syllable in /ˈ3askari/ located within the mid-upper range. Speaker 07_F

4.6.1.3 L-

This is a low phrase accent occurring at the right edge of a word. This phrase accent is often realised as a sharp drop in pitch in anticipation of an intermediate phrase boundary. Moreover, depending on the previous target configuration, it could also surface as a flat low stretch in other cases.

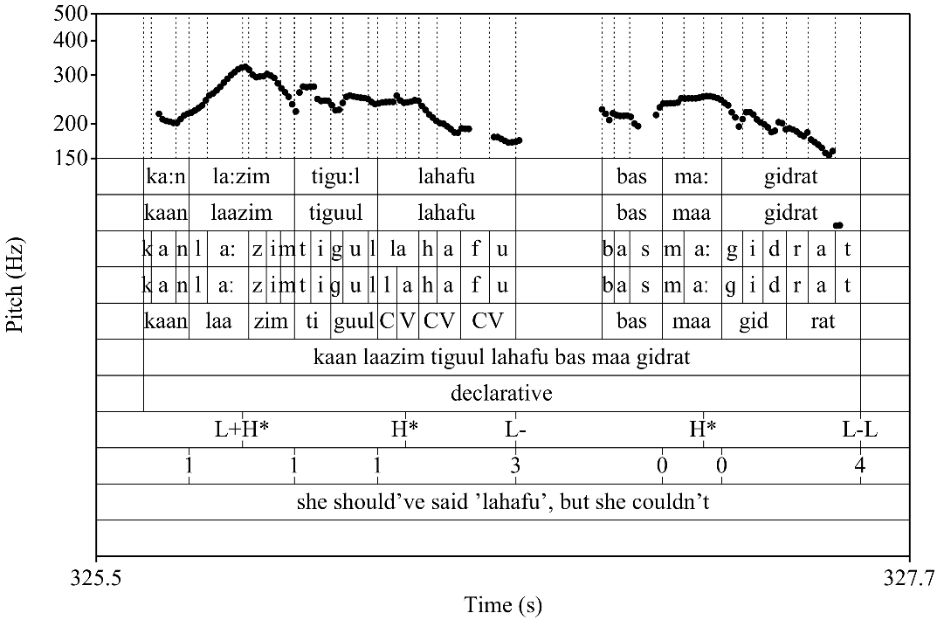


Figure 36: Low phrase tone L- on the edge of /<u>la</u>hafu/ which shows a downstepped high !H* accent. Speaker 01_F

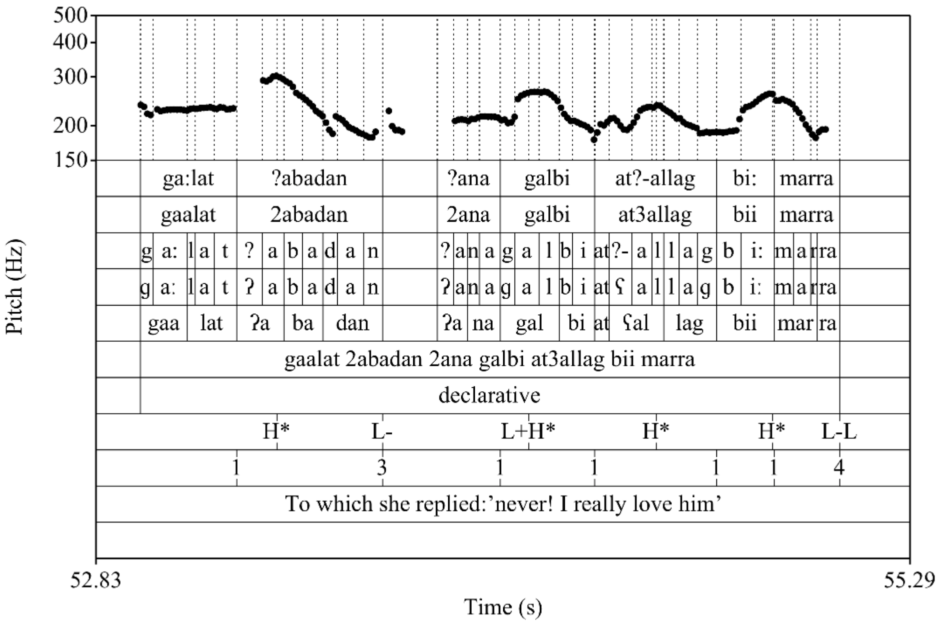


Figure 37: Low phrase tone on the edge of /2abadan/ which shows a high H* accent. Speaker 01_F

4.6.2 Boundary tones

Within AM theory and ToBI conventions, a boundary configuration is a combination of a phrase tone at the edge of the word (independent of lexical stress) followed by a boundary tone movement marking the end of a contour. The boundary tone movement marking a full intonational phrase can end in a high H% or a low L%. This phrase tone- boundary tone combination yields several shapes, among them in JA are: L-L%, H-H%, !H-H%, L-H%, H-L%, !H-L%. As was presented in section 3 discussing the main intonational tunes above, these boundary tones paired with specific pitch accents are used to make up a particular tune with a contrastive meaning. Note the following schematic representation of these configurations:

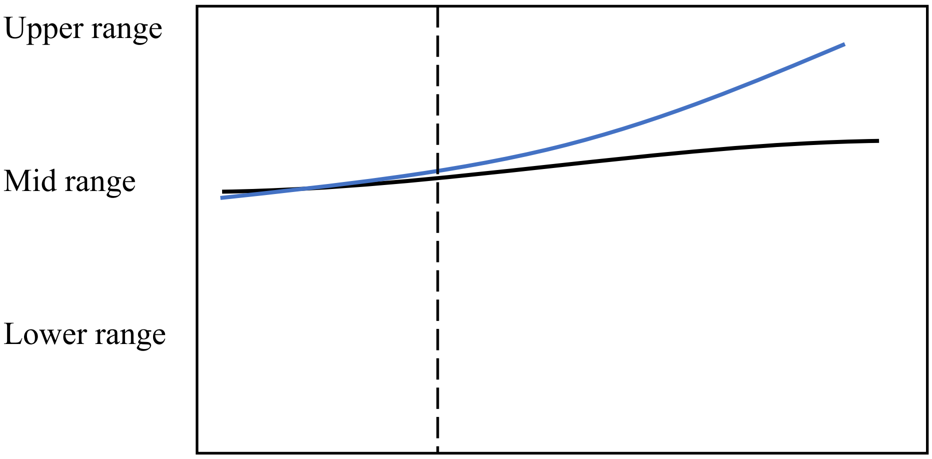


Figure 38: Boundary tone combinations in JA. Left of the vertical dashed line represents phrase tone configuration & level, the right of it represents the boundary tone configuration. Top curve is the 'High-rising' yes/no tune transcribed as: H-H%, Bottom curve is the 'downstepped-high' WH question tune transcribed as: !H-H%

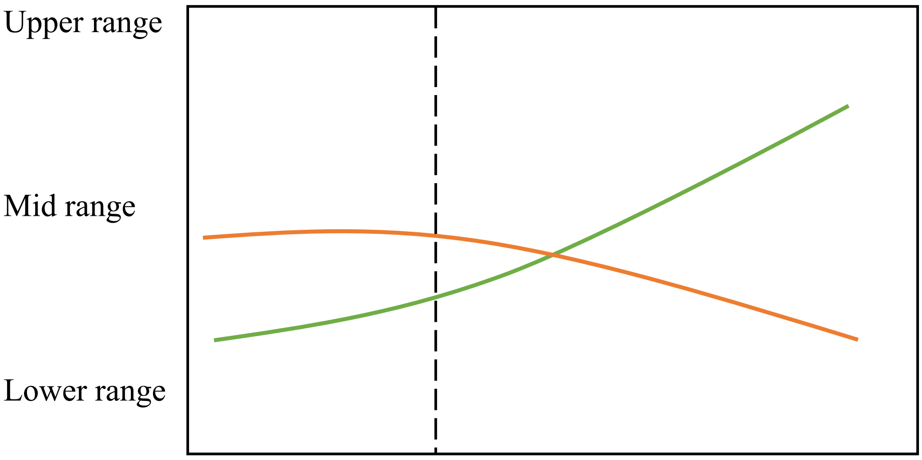


Figure 39: Top curve is the 'Falling' declarative tune transcribed as: L-L%, Bottom curve is the 'low-rise' continuation tune transcribed as: L-H%

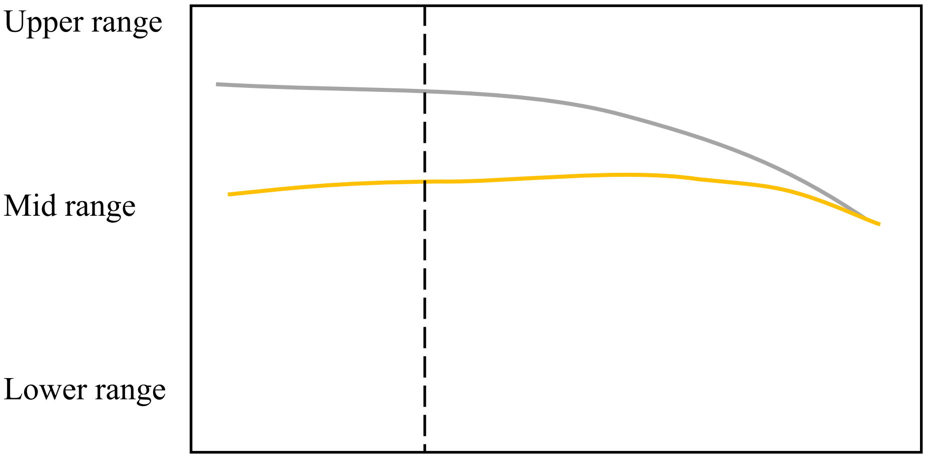


Figure 40: Top curve is the 'rise-low plain' plateau tune transcribed as: !H-L%, Bottom curve is the 'rise-low downstepped' plateau tune transcribed as: H-L%

4.7 Prosodic structure in Jeddah Arabic

This section discusses the constituency (phrasing) and prominence hierarchy in the dialect. As can be inferred from the examples in section 4.4 above, there is a clear indication of an intonational level of phrasing that contributes to the 'rhythmic grouping' of utterances in Jeddah Arabic. The level of phrasing is tonally demarcated via the tonal configurations signalling the edges of those phrases, as well as the acoustic and auditory juncture cues. Here details of the function of the phrases, the tonal and non- tonal phrasal correlates are discussed, while the phonetic correlates of prominence as related to accent distribution and tonal alignment are discussed in the upcoming chapters.

4.7.1 The constituency Hierarchy

This Arabic dialect, as the case with the other Arabic dialects, displays three types of prominence: lexical stress, pitch accents and nuclear accents. Every prosodic word has an obligatory stressed syllable marking a phonetic and phonological difference between stressed and unstressed syllables. Stress is assigned phonologically in the lexical level according to the moraic rules presented earlier in section 0. Similarily, all intonational phrases as observed in section 4.4 above show obligatory pitch accents, which are realised on the chosen lexically stressed syllable(s). Findings from the upcoming alignment chapter confirm the association of pitch events with lexically stressed syllables, thus demonstrating a phonotactic constraint that demands pitch events to be associated with this lexical unit. This intonational level too marks a phonetic and phonological difference between accented and unaccented syllables according to the accent distribution rules within and across domains. The intonational levels as discussed show subordination among the prosodic constituents, as well as syntagmatic relations within those constituents. In this sense, the intonational level of representation in Arabic may be said to conform to the *Strict Layering Hypothesis* (Selkirk, 1984, Nespor & Vogel, 1986, Chahal, 2001) demonstrating a hierarchical organisation whereby the highest level is composed of at least one of the lower levels.

4.7.1.1 The intermediate Phrase

The ip is a level of phrasing below an IP, and is made of at least one syntactic phrase (NP, VP, PP) that carries part of the information of the whole utterance (Chahal & Hellmuth, 2014). The intermediate phrase is reported to be signalled by the phonetic cues of early tonal alignment; in Arabic, it was found that these cues differ in strength as compared to what would happen at an IP boundary (Chahal, 2001, Chahal & Hellmuth, 2014). This is the case in Jeddah Arabic, as tonal alignment patterns discussed in chapter [6] report a difference between ip and IP boundaries. Additionally, as was observed in the examples in sections 4.4 to 4.6 above, an infrequent silent pause may be detected following this constituent in Jeddah Arabic, which serves as a break or disjuncture that speakers would insert to express non-finality. This constituent ends in a phrase tone T- that is either: high H-, high scaled !H- or a low L-. The phrase tones are observed mid-sentence[12] when the contour following the accented syllable displays an abrupt pitch configuration at word edge that cannot be attributed to a trailing tone or a result of interpolation. The intermediate phrase is the domain of pitch reset and pitch control where the new phrase following an ip edge demonstrates a new pitch register. Regarding accent distribution in this phrase, accents are placed from left to right according to the focus realisation of the utterance and reflect relative prominence relationships. Thus words can be nuclear accented, pre-accented, or unaccented. Among these pitch accents, the ultimate nuclear accent bears the highest prominence, and is considered the head of that intermediate phrase.

[12] There are cases where the phrase tone is separated from the boundary tone by a number of unaccented material in a pattern known as 'tone spreading', hence a phrase tone is not immediately followed by a boundary tone as how it usually occurs at the end of IPs. Evidence for this is discussed in the chapter on focus, where a phrase accent occurs mid-sentence and its phonetic value spreads over the unaccented material until the boundary target.

4.7.1.2 The Intonational Phrase

The highest intonationally- demarcated phrase that ends in a boundary tone T% on its right edge[13]. This prosodic phrase coincides in length and function to a full syntactic sentence "syntactic root clause" and carries the complete information and meaning of the utterance (Lebanese and Egyptian in Chahal & Hellmuth, 2014, English in Pierrehumbert, 1980, and Cole, 2015). In Jeddah Arabic it is signalled by prominent boundary cues such as early tonal alignment and significant silent pauses at the juncture. As can be seen in the examples throughout section 4.4, this constituent is the domain for phonetic declination of peaks leading to the boundary, upstep, downstep, as well as reset of pitch for a new phrase following the boundary. These phenomena are reflective of the main characteristic difference between an IP and an ip, i.e. complex pitch configurations at an IP edge (e.g., H-H%) versus a much simpler monotonal configuration at an ip edge (e.g., -H). Regarding accent distribution, subject to focus interpretation of the sentence, every IP is comprised of at least one intermediate phrase with optional pre-nuclear accents, one nuclear accent, and no post-nuclear accents. In the case that an IP is made up of more than one ip, the last pitch accent before the boundary is considered the nuclear head of that phrase, after which no pitch accents are permitted in the same phrase. This nuclear accent combines with a boundary configuration to form one of the pragmatic tunes discussed previously. In anticipation of the upcoming chapter on focus realisation, it might be worth mentioning here that facts from phrasing in narrow focus utterances confirm that in JA the nuclear accent is the most prominent accent in an intermediate phrase. After this accent no post-nuclear accents are seen in the same IP. Moreover, facts from broad focus show that in the same IP, there may be pre-nuclear accents, but the last word before the boundary is the most prominent nuclear accent. An analysis on the effect of focus prominence on the quantitative detail is presented in the respective chapter accordingly.

The previous paragraphs discussed the significant tonal and juncture characteristics of intonational and intermediate phrases in Jeddah Arabic, including the tonal implementation phenomena observed within those domains. Recall that it was discussed earlier that prosodic boundaries (IP/ip) show both tonal and non-tonal correlates that are taken as evidence for their occurrence. In addition to pausing and juncture, the non-tonal cues include pre-boundary

[13] IPs can also be marked by a boundary on their left edge to indicate how a contour starts, this is language-specific as originally proposed in pierrehumbert (1980). For Arabic, chahal proposes an initial %L but only in cases where the contour unusually starts at a low range. She maintains that the default starting point in Arabic is mid range. JA shows no such low initial boundary, and contours start at "neutral value, as Pierrehumbert names it.

lengthening whose degree is taken to be indicative of boundary strength. Final lengthening has been shown in numerous studies to be an indicator of an IP boundary affecting syllables immediately preceding the boundary. Syllables at this boundary have been reported to be longer in duration compared to their counterparts in the vicinity of ip boundaries in Lebanese Arabic (Chahal, 2001), English (Ladd, 2008, Pierrehumbert, 1980, Beckman & Pierrehumbert, 1986) and European Portuguese (Frota, 2000), among other languages. For Egyptian Arabic, Hellmuth suggests that the tonal cues for prosodic boundaries were more consistent than the non-tonal cues, and pre-boundary lengthening although observed was an inconsistent cue in the data (Chahal and Hellmuth, 2014). To examine pre-boundary lengthening in the current corpus, data from the alignment experiment in chapter 6 are used. The alignment experiment places words of varying stressed syllable types and stress positions in the phrase-mid (PW) contexts, ip and IP boundaries in order to analyse their tonal alignment patterns. It is perhaps worth mentioning here that the respective chapter reports prevalent tonal alignment differences between IPs and ips in Jeddah Arabic, which contribute to the marking of their boundaries. Bearing this behaviour in mind, the pre-boundary lengthening effects are discussed in the following.

As was presented in the methodology chapter [3], the material for the analysis of the prosodic structure pre-boundary lengthening effect is the tonal alignment stimuli. The full list of sentences can be found in Appendix [D]. The stressed/accented syllable location in the tonal alignment experiment varies in distance to the prosodic boundary in question (ip or IP). The stressed/accented syllable can be located two syllables from the boundary /`la.ha.fu/, one syllable from the boundary /mu.`laa.zim/and /riH.`lat.hum/, or zero syllables from the boundary /lil.Haf.`laat/. The PRAAT script used to extract the tonal alignemnt meausrements in chapter [6] also includes durational measurements of target syllable duration, vowel duration and word duration. Subsequently, statistic significance of means and post-hoc comparison tests were carried out to interpret the results. One-way ANOVAs are conducted to analyse the within-word mean differences (in syllable duration, vowel duration and word duration) according to prosodic boundary (ip, IP and PW). Tukey HSD post hoc analysis for multiple comparisons with (vowel duration, syllable duration and word duration as dependent variables, and condition (ip, IP, PW) as independent variables.

In order to investigate the lengthening correlate in Jeddah Arabic in the absolute proximity case first, the word /lil.Haf.`laat/ was thus included in the corpus, as the stressed syllable is final in this word and would be final by default before a boundary. The average duration of the syllable and vowel in the vicinity of each boundary including the durations in prosodic words is presented in the following:

Boundary	Vowel duration	Syllable duration	Word duration
ip	161.78 ms (SD 40.39)	306.62 ms (SD 63.43)	695.65 ms (SD 99.34)
IP	127.54 ms (SD 24.98)	258.52 ms (SD 48.89)	621.70 ms (SD 75.66)
PW	114.71 ms (SD 33.98)	218.84 ms (SD 56.94)	568.43 ms (SD 97.63)

Table 6: Average vowel, syllable and word durations of the word /lil.Haf.laat/ with final stress.

As can be inferred from the previous table, the word shows lengthening before major prosodic boundaries, however in a converse pattern to what has been reported for other languages. The stressed syllable (head foot), long vowel (bimoraic) and whole prosodic word were all lengthened to a greater degree before ip boundaries than before IP boundaries. Analysis of variance and pairwise comparison results for all durational measurements are significant for this word (p= 0.00: vowel duration F= 40.49, df= 2; word duration F= 39.41, df=2; syllable duration F= 52.30, df=2).

It is proposed that the pre-boundary lengthening effect is also observed on units further away from the boundary. For example, in the secondary target word /mu.`laa.zim/ the stressed syllable is separated from the boundary by one intervening syllable. Note the lengthening patterns in this case:

Boundary	Vowel duration	Syllable duration	Word duration
ip	155.83 ms (SD 23.82)	195.65 ms (SD 27.30)	551.46 ms (SD 81.15)
IP	143.42 ms (SD 23.42)	182.00 ms (SD 28.48)	476.14 ms (SD 80.63)
PW	121.18 ms (SD 22.75)	158.70 ms (SD 29.58)	437.00 ms (SD 67.34)

Table 7: Average vowel, syllable and word durations of the word /mu.laa.zim/ with medial stress.

Again, the same conclusions can be drawn; while major boundaries do show lengthening, proximity to ip boundaries cause units to be longer in duration compared to IP boundaries. Anova and pairwise comparison results are all significant for this word (p= 0.00: vowel duration F= 56.92, df= 2; word duration F= 54.04, df=2; syllable duration F= 41.01, df=2). The same pattern is observed in the other two target words in the corpus containing syllables further removed from boundary /`la.ha.fu/ and /riH.`lat.hum/. Analysis of variance and pairwise comparison results are all significant (p=0.01 for /lahafu/: vowel duration F=19.14, df= 2; word duration F= 56.11, df=2; syllable duration F= 40.73, df=2), and p= 0.03 for /riHlathum/: vowel duration F= 20.38, df= 2; word duration F= 50.04, df=2; syllable duration F= 29.26, df=2). The words include stressed syllables separated from boundary by one and two syllables:

Boundary	Vowel duration	Syllable duration	Word duration
ip	67.67 ms (SD 9.60)	125.53 ms (SD 16.18)	444.77 ms (SD 63.75)
IP	60.35 ms (SD 11.33)	112.81 ms (SD 21.67)	412.15 ms (SD 61.44)
PW	58.66 ms (SD 8.90)	102.39 ms (SD 15.55)	357.00 ms (SD 54.34)

Table 8: Average vowel, syllable and word durations of the word /la.ha.fu/ with initial stress.

Boundary	Vowel duration	Syllable duration	Word duration
ip	74.29 ms (SD 9.79)	192.10 ms (SD 25.50)	616.62 ms (SD 87.48)
IP	68.48 ms (SD 9.21)	189.00 ms (SD 27.67)	556.78 ms (SD 81.49)
PW	65.36 ms (SD 8.94)	166.10 ms (SD 26.52)	503.00 ms (SD 69.87)

Table 9: Average vowel, syllable and word durations of the word /riH.lat.hum/ with medial stress.

The previous pre-boundary lengthening analysis suggests that durational manipulation constitutes a non-tonal correlate for intermediate phrases in Jeddah Arabic and the observed differences in means were all significant at the 0.05 level. Relying on the consistency of the results across syllable/vowel types and prosodic boundaries, it can be said that intermediate phrases in Jeddah Arabic show a durational pattern opposite to what has been proposed in other Arabic dialects, English and European Portuguese.

One possible explanation for this lengthening pattern in JA is the fact that ips in the dialect are not often followed by pauses, making the disjuncture at this boundary much less defined. Studies in this regard have shown that lengthening of pre-boundary units is correlated with the occurrence of pauses, as well as the length of the pause (Wightman et al., 1992, on American English, Frota, 2000 on European Portuguese). Thus, English and European Portuguese have reported that a unit at a boundary accompanied with a pause was longer than one that is not accompanied by a pause, and longer pauses caused greater lengthening than shorter pauses. In addition to this, it is also suggested that major prosodic constituents tend to rely more on tonal cues to define their boundaries, while durational cues are used more often in smaller boundaries (ibid). The current lengthening analysis in JA does not control for ip boundaries depending on whether or not they are accompanied by pauses nor takes into account the length of the pause. In this case, it may be initially suggested that the observed lengthening at this boundary is to compensate for the absence of the pause in the ips with no pauses, or the ips with short pauses. In turn, this lengthening pattern may be used to compensate for the prosodic strength of this boundary in comparison to the IP boundary, i.e. signal how different an ip edge is. Jeddah Arabic lengthening patterns here resemble those reported for Spanish in Rao (2010). The three Spanish dialects he analysed show that ip boundaries are accompanied by short (or no pauses) and significant pre-boundary lengthening of syllables and whole words. As an explanation for this pattern, he suggests that ip edges show greater lengthening as a way to signal the non-finality of an idea, while greater pauses and less lengthening at IP boundaries signal the termination of an idea. It is easy to see how this is replicated in JA, as ip phrases occur half way through a sentence and constitute part of the information of the whole sentence. Bearing this discussion of boundary lengthening effects in mind, it must also be noted that the researcher is aware that another confounding factor may be present in the current JA analysis. The current analysis does not control for the inherent speech rate differences in the corpus that may influence the segment durations. This could be done in a follow up analysis by using a duration normalisation measure, as well as incorporating the pause analysis mentioned above to discriminate ips from IPs.

Based on the investigations in this section, it is concluded that prosodic phrasing in Jeddah Arabic shows consistent tonal correlates (edge configuration, alignment patterns and within- domain pitch patterns), as well as non-tonal correlates in the form of pausing, juncture and lengthening effects. More specifically, the discussed correlates demonstrate how Jeddah Arabic patterns support evidence for cross-dialectal variation in boundary effects. The conclusion that there are indeed dialectal differences in the degree and direction of lengthening before prosodic boundaries.

4.8 Jeddah Arabic comparison with LA and EA

The model of intonation proposed here for Jeddah Arabic includes a tonal inventory that is comprised of four pitch accents: H*, !H*, L+H*, L*. A pitch accent of the phonetic shape L+!H* was observed in some instances but there was not enough evidence to include it as part of the phonological inventory. The JA model shows phrase accents of the shape H-, L-, and !H-, as well as boundary tones of the shape H% and L%. The dialect demonstrates the following tonal implementation rules: Downstep, upstep, phonetic declination, and interpolation.

There are two postlexical phrasing levels above the prosodic word in JA: the intermediate phrase and the intonational phrase. The two domains account for the rhythmic alternation of words in JA sentences. The IP is the highest level of tonally marked phrases and it is realised via distinct, complex pitch configurations on the right edge. The ip is a level of phrasing below an IP and above a prosodic word, also showing specific pitch configurations at the right edge. The dialect uses both tonal and non-tonal correlates to mark these two levels of phrasing, and there has not been evidence supporting a further level of phrasing. Recall the literature in chapter 1 introduce a further level of phrasing between an ip and a prosodic word in the prosodic hierarchy, namely an *Accentual Phrase* (AP). Evidence for the existence of this further level of phrasing however has not been supported by the current JA analysis, and thus only two levels of tonally marked phrases above the word are proposed for the dialect.

The dialect bears similarities and differences to the other intonational models of Lebanese Arabic and Egyptian Arabic. The similarities JA shares with EA are the existence of an L+H* pitch accent shape, phrase and boundary tones H-, L-, H% and L%. Moreover, the two dialects are similar in terms of marking declaratives with an L-L% boundary, while both yes/no questions and continuation tunes are marked with an H-H% boundary. However, JA shows more stylised boundary configurations (presented in section 4.4.2) that makes use of the phonological downstep property to mark WH tunes and plateau tunes.

JA also shares similarities and differences with Lebanese Arabic. Like Lebanese, JA shows a number of pitch accent types, as opposed to only one pitch accent type in EA L+H*. The two dialects are also similar in their use of the downstep property to mark different accents, phrase tones and boundary tones: !H*, !H-, !H-L%, !H-H%. The main differences between the two dialects is the existence of a further phonologised L+!H* pitch accent in LA, and the way they mark yes/no questions. In LA yes/no tunes are of the shape: L* or H* + H-H%, while in JA they are unmarkedly L+H* + H-H%. The following chapters will further demonstrate how JA is different from LA and EA in terms of tonal alignment and Focus patterns. Adding to those differences is the direction and degree of pre-boundary lengthening of units at ip and IP edges. While Lebanese Arabic reports greater degree of lengthening before IP boundaries than before ip edges, Jeddah Arabic saliently lengthens units in a converse manner. It is reported here that ip edges in JA cause units to be longer in duration than before IP boundaries.

4.9 Conclusion

This chapter introduced the main tunes and their composition, the accent shapes used by the dialect, their function as prominence lending or non-prominence lending, and the prosodic structure in the Jeddah Arabic dialect. It has been argued that there are intonational phrases and intermediate phrases, and the ways they coincide with syntactic grouping were presented. The tonal and non-tonal correlates of these phrasing levels have also been presented along with the implications for the phrasing patterns in the dialect. It can be initially proposed that prominence distribution of accents within a phrase would be subject to focus interpretation of the utterance that shows a tendency for prominence to be right-headed. Nuclear accents would be the most prominent accents on the right of an intermediate phrase, pre-nuclear accents would be permitted, but post-nuclear accents would be prohibited. Based on this, it can be suggested that in JA there is a relative hierarchy of accent prominence. It is now time to investigate the experimental evidence for those observations, how these accents are distributed in the respective phrases as an effect of manipulations in focus type and sentential position, the phonetic and quantitative detail signalling prominence; and how tones align with the intonational structure proposed in this chapter. Focus is investigated to see what happens to phrasing levels when we change the nuclear accent's expected position, and tonal alignment is investigated to see what happens to the alignment of the tone when there is phonetic manipulation of time pressure.

Chapter 5. Focus and Prominence

5.1 Introduction

Focus is a term related to the information status of an utterance. It can be regarded as a tool to convey information about or highlight specific parts in the utterance. This takes place in order to serve a communicative function or a pragmatic interpretation of some sort. The manifestation of focus has been reported in many studies on many languages to be through the use of prosody- that is, intonationally- among other lexical and pragmatic tools (Xu & Xu, 2005, Hellmuth, 2011). In particular, accentuation (or accentual prominence), which is phonetically signalled by F0 configurations and de-accentuation are proposed to be the prosodic reflexes to distinguish between focused and non-focused information (Xu & Xu, 2005, Arvaniti, 2017, Hellmuth, 2011). Focus, then, is prosodically achieved by allocating the target enhanced accentual prominence.

A terminological distinction concerning the definition and nature of focus is reported in the literature. Hence, two types of focus definitions regarding its nature emerged: information focus and contrastive/ identificational focus. Information focus distinguishes between new information to the utterance and given information, whereby new information entails unpredictability and originality, contrastive information entails a contrasting relationship among elements mentioned in the discourse (Al-Zaidi, 2014). In his thesis employing syntactic and prosodic methods to analyse information structure in the Hijazi variety spoken in Taif, Al-Zaidi (2014) reports that in Hijazi, specific question types can evoke two types of focus. A WH-question such as ‘what’ evokes general, ‘broad’ information about a sentence, while a yes-no question triggers contrastive, ‘emphasised’ focus.

A further distinction regards focus to be defined by its scope as well as by its nature. Advocates of this approach are Hellmuth (2011) and Ladd (2008) among others. The scope defining the domain of emphasis that can either place the whole utterance under focus “Sentence focus”, or a specific component under focus (Hellmuth, 2011). Ladd (2008, p. 215) regards the former as “Broad Focus” on a large unit such as the sentence, while the latter as “Narrow Focus” on single unit/ word. The experiment in this chapter is thus concerned with both as related to the domain and scope of focus. Although the reported studies use narrow focus and contrastive focus interchangeably, I will use the term narrow focus in the analysis.

5.2 Experimental Literature

Phonetic studies on narrow focus across languages report qualitative and quantitative correlates signalling different focus conditions. In Arabic, a comparative study by Yeou et al. (2007a, b) on contrastive focus patterns in mid- sentence position in Moroccan Arabic, Kuwaiti Arabic and Yemeni Arabic analysed the acoustic correlates of F0 alignment, F0 excursion size and vowel duration, in addition to tone contour shape and phrasing. For the purposes of the experiment in this chapter, only the results of qualitative (contour shape and phrasing), duration and excursion size will be presented, while the results for F0 alignment patterns are discussed in the next chapter. Yeou et al.'s study reports a great effect of focus on target vowel duration and excursion size in all three dialects. A difference of 25, 29, and 49 milliseconds between Broad and Narrow focus vowels was reported for each dialect. Regarding excursion size in target syllables, the study reports a larger excursion size (or a larger F0 change between L and H) in syllables under narrow focus, with a greater excursion size difference between focus conditions reported for Moroccan, and the least difference reported for Yemeni. The small difference in excursion size in Yemeni might indicate that this acoustic measure is not used to signal different focus conditions in this Arabic dialect, while it seems that it is a significant correlate in the other two dialects. The excursion size differences are largest in Moroccan with a reported 5.33 Semitone differences between Broad and Narrow syllables, followed by a 3.25 Semitone difference in Kuwaiti and lastly a 0.63 semitone difference in Yemeni. The qualitative results report a systematic use of pausing, phrasing, and de-accentuation to mark contrastive focus. The results of the de-accentuation patterns of pre- and post- focal material will be reported in the next section. The study provides evidence that in both conditions a contour shape of L+H* can be used on the target word, suggesting that contour shape/ accent type is not used as a qualitative measure. However, in the narrow focus condition the peak is raised considerably in order for the tone to be distinguished from the surrounding string. In addition to this and only in Kuwaiti a high rising tone H* can also be used alongside L+H* to mark focalisation. Phrasing is also only used as a qualitative measure in the Kuwaiti dialect represented by a high rise F0 movement after the target tone followed by a pause of about 115 millseconds following target word. The maintained high tone after the target syllable is indicative of a high intermediate phrase tone H- as shown in their graph (Yeou et al.: 2007b, p. 325).

A series of studies by Chahal (1999, 2014, 2003, 2001) on Lebanese Arabic reports that narrow focus in three different sentential positions in this dialect is also signalled by a combination of quantitative and qualitative measures. First, she reports that accent type is not a qualitative measure to mark different focus conditions since both H* or L+H* were used by speakers. Rather, the qualitative measure of de-accentuation was systematically used to mark narrow focus. The results of this will be presented and discussed in the next section. Second, the target word in broad focus utterances was not allocated a nuclear accent (although it was pitch accented) but the last word in the Intonational Phrase (IP) was promoted as nuclear accented (even when it was not the target word) and was distinguished by higher intensity, duration and peripheral vowel formant structure. Moreover, broad focus utterances were in the majority of cases realised as one IP suggesting that phrasing is not used to mark this condition. Third, the researcher reports that narrow focus is marked by a higher peak (increased F0) and intensity, a larger pitch range/ excursion size, peripheral vowel qualities, and longer vowel duration. Alongside this, phrasing was also systematically used to mark this focus condition in this Arabic dialect, whereby the three participants in her study insert an IP (H%, L%) or intermediate phrase ip (L-, H-) boundary after the target. She notes that when phrasing is used to break a Narrow focus utterance, the target word is realised as a separate phrase with a nuclear accent and a boundary tone. The following string is thus realised as a separate phrase with its own accent distribution, but whose pitch range is nevertheless dramatically compressed compared to the previous phrase. Moreover, she reports that the peaks in the consecutive phrase are also lower compared to the peak of the target in the previous prosodic phrase. She therefore concludes that the realisation of a focus condition in this dialect takes into account the target's status in the accentuation distribution mechanism. That is, a nuclear accented target will demonstrate more distinctive phonetic characteristics (F0, duration, RMS and formant frequencies) than a pitch accented target, which is also more distinctive that an unaccented target. This reflects the relationship in the prominence hierarchy presented in section 2.2.4, **Table 1**.

In Egyptian Arabic, Hellmuth (2006a, b, 2014) conducts a series of studies in an attempt to analyse the prosodic reflexes of focus in this dialect. Pitch accent type is not used as qualitative measure in the dialect, since all words bear an L+H* accent. However, phrasing is reported in Hellmuth (2014) to occasionally mark contrastive focus by the insertion of a L- low phrase tone after target, even though this pattern was not taken by the author to be a systematic reflex of focus. The main quantitative measure used to mark focus in Egyptian is excursion size. The pitch range of the word under contrastive focus was reported to be larger in Semitones than an unfocused word, despite the fact that the latter is still pitch accented. The author therefore maintains that pitch range manipulation (either expansion or compression) to be the main reflex of focus in this dialect. This method of marking of focus using one main acoustic correlate is not limited to the Egyptian dialect. In Jordanian Arabic, De Jong and Zawaydeh (1999, 2002) show that their participants varied considerably in how they marked focus, and confirm that neither duration nor vowel quality were employed systematically to mark narrow focus (The study does not report any other quantitative or qualitative results). In other studies on Egyptian Arabic, El-Zarka (2011, 2013) argues that while EA shows an invariant pitch accent shape, her results show alignment differences between the two focal accents according to focus interpretation. Her analysis accounts for the pitch configuration on the whole word under focus. Thus, she proposes a special LHL accent sequence, where the first rise LH coincides with the beginning of the stressed syllable with the peak realised therin, while the second L which she calls "the closing L tone" is the result of the fall from the previous peak. The results from the three speakers in her (2011) study show that under narrow focus, the peak consistently shows earlier alignment with the stressed syllable, and the second L tone- equally as important in enhancing prominence- consistently shows early alignment with the edge of the focal word compared to broad focus accents. She continues to argue that the alignment differences between the two conditions take place in absence of any differences in pitch height of the peak or valley, as well as being highly speaker-dependent. A point worth bearing in mind here that there is a great deal of cross-dialectal and inter-speaker variation in Arabic regarding focus marking.

Al-Zaidi (2014) on Taif Arabic reports results comparing contrastive focus and broad focus in initial and medial positions. He reports that accent type is not a qualitative measure to distinguish focus condition in this dialect, since both L+H* or H* can be used in either condition. It is worth noting that some of the examples provided in Al-Zaidi's thesis, do in fact demonstrate that phrasing could also be used as a qualitative measure to mark focus, however the author does not state, nor analyse this. This is evident from the sustained high F0 stretching to the end of the target word and following the high tone on the stressed syllable, which is pitch accented in his examples. Whether or not all his speakers used this strategy is not mentioned. As for quantitative measures, in utterance initial position he found that the F0 peak of contrastive focused words is considerably higher than the same word in the broad focus utterance. The duration of the word and syllable in contrastive focus utterances was also found to be longer than the broad focus counterparts, with a reported difference of around 31 milliseconds. Moreover, word excursion size and intensity were found to be larger in contrastive focus syllables compared to broad focus (a significant difference of around 4 Semitones and 3 dB, respectively), thus suggesting an expansion effect under focus. The results are similar for the utterance medial position. An effect of contrastive focus on peak height, excursion size, duration and intensity was also evident in this position. Thus, a higher peak, a longer duration and an expanded range were reported for contrastive focus. A difference of 3 Semitones, 2 dB, and 16 Milliseconds was observed between contrastive and broad focus targets in medial position.

One point of interest and relevance to the current thesis is Al-Zaidi's observation regarding the domain and location of the F0 peak in relation to the lexically stressed syllable. The author shows that throughout his study and across sentence types, focus conditions and syllable types, F0 maxima were always aligned within the confines of the vowel and do not span across the TBU or prosodic word. This suggests that peak alignment is not a syllable structure (open vs. closed, heavy vs. light) differentiating prosodic method in the dialect, nor is it a quantitative measure differentiating focus conditions. This supports a *Segmental Anchoring Hypothesis* analysis of the dialect maintained in AM theory, which states that the lexically stressed syllable is the domain of accent placement. It also implies that the dialect possesses an L+ H* bitonal accent, and not an L*+ H bitonal accent as well. More on this topic will be discussed in the chapter on tonal alignment. The results of post -focal compression in this dialect will be discussed in the next section on post focus.

In a follow up analysis on Taif Arabic by Al-Zaidi, Xu and Xu (2019), the authors provide extensive acoustic results of focus prosody. The analysis compares contrastive and broad focus types in initial and final sentential positions and includes the analysis of the pre and post focus region. For the acoustic measurements, they report results for maximum F0, mean F0, excursion size, intensity, and duration by 8 male and 8 female participants. Their study reports that the focused constituent (on focus) shows higher maximum and mean F0 and intensity, longer duration, and larger excursion size in both initial and final positions. Their results on pre- and post focal region is discussed in section 5.2.1.

Studies on cross language behaviour regarding focus reveal interesting cross-linguistic implications. A study by van Donzel and Beinum (2000) on information structure in Dutch declaratives revealed that the language uses F0 variations to mark focus in discourse. Their study distinguishes between focused information that is more likely to be allocated a prominent pitch accent and given information that is inferable from context to not be marked by any special F0 movements. They also corroborate the findings stating that modifiers and adverbial clauses of time and place may also be allocated pitch accents-therefore not deaccentuated- even when unfocused because they contribute to the linguistic information of an utterance. They report that phrasing and pausing (indicative of phrase ip and boundary IP tones), accent placement all had an effect on perceived prominence by listeners. Thus, higher peaks and phrasal tones on targets were also perceived as a cue to more prominence, and hence more nuclear accent judgments.

In a study on Spanish contrastive focus patterns, Face (2002) states that in addition to sentence word order as a reported method to convey focus in previous studies on the language, there are also qualitative and quantitative intonational methods. The author uses controlled answer-question paradigm to elicit broad and narrow focus in initial and medial position. The study reports that pitch accent type is a qualitative measure used to convey contrastive focus in Madrid Spanish. An L+ H* accent is exclusively used to mark narrow focus, while an L*+ H is used to mark both broad and narrow targets. When an L*+ H accent type is used in narrow focus, the peak is considerably higher than in its counterpart condition. This is also indicative of a peak height difference distinguishing both conditions, which is also reported as a quantitative measure in this study. A reported 28 Hz height difference occurs between conditions. The language also strategically uses phrasing as a method to convey focus by inserting an L- or H- phrase tone after the target word. The author reports that following this phrase tone; an optional pause of 50-100 Milliseconds may or may not occur. This study also reports a significant difference in rise duration as affected by focus condition. Rise duration (Ms) and rise speed (Semitones per second) are quantitative measures used to assess the temporal distance and rate in rising tones L+ H* and L*+ H across languages (Xu & Xu, 2005, Xu & Sun, 2002, Flemming & Cho, 2017). A longer rise time and slower speed between the (L) and the (H) is reported for L*+ H bitonal tones. In L+H*, a longer rise time is accompanied by a faster speed as both tones need to be realised in the same stressed syllable, and therefore would need a more rapid and spacious transition. The measures of rise duration and speed in Xu and Sun (2002) are used to analyse the speed and duration of the pitch change in bitonal accents that correlates them with changes in excursion size. They view the bitonal accent movement as involving 3 stages: the acceleration ‘rise’, the glide ‘reaching the peak’, and deacceleration ‘the fall following the peak’. Observing this pattern led them to conclude that the speed of pitch change in bitonals varies with the excursion size of the accent; that a larger duration and excursion size causes a faster speed from the L to reach the peak. This pattern neatly colligates the realisation of bitonal accents that is reported under focus prominence; the increase in the duration of the syllable is correlated with the larger excursion sizes, and faster speeds in L+H* accents under contrastive focus. In the following is a schematic representation of Xu and Sun’s (2002) proposed mechanism. The middle ramp in the figure represents the three stages of the bitonal accent production. Response time corresponds to the time it takes to reach 75% of the pitch change in an LHL sequence.

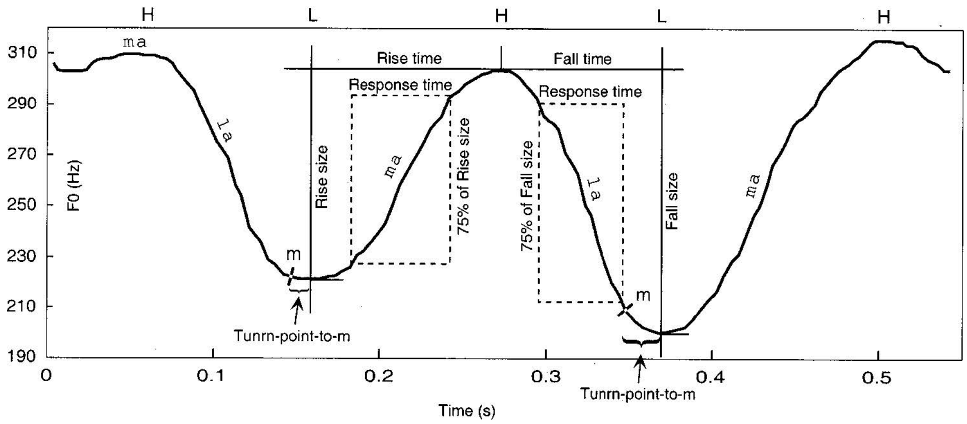


Figure 41: Schematic illustration of rise and fall times and speed calculation. Xu & Sun (2002).

Indeed, it has been reported that speakers systematically increase the excursion size, rise speed and the duration under contrastive focus, as one way of enhancing the tone. Face's study reported a 136 Milliseconds rise duration in initial position (138 in medial position) for this rising tone in Spanish, compared to 125.3 (Ms) for Chinese, and 142.2 (Ms) for English as reported in Xu and Sun (2002). Finally, Face's study reports that duration is also a quantitative measure used to convey focus in this language. A reported 49 Milliseconds difference occurred between the two conditions in initial position, and a 59 Milliseconds difference in medial position. It can be said then that peak height, duration, rise duration, phrasing and accent type are used to convey focus in Madrid Spanish.

Many studies on English (Ladd, 2008, Gussenhoven, 1983, Xu and Xu, 2005, among others) report that focus is conveyed quantitatively in the language, and qualitatively through deaccentuation. Under focus, English words have higher F0 peaks, larger excursion sizes, longer durations, and faster rise speed compared to a neutral focus counterpart (Xu & Xu, 2005). The position of the target under narrow focus also showed quantitative positional differences regarding max F0 and excursion size. The peak was highest (10.9 St) in initial position, followed by medial (9.9 St) and final (8.1 St). The excursion size showed a similar pattern with the largest excursion reported for initial position (3.2 St), then medial (3.0) and final (2.4 St). Rise speed was faster for narrow focus targets (23.4 ST/s) as compared to broad focus (9.5 ST/s). Pre- and Post- focus results will be discussed in the deaccentuation/ compression section 5.2.1.

For German, a perceptual study by Baumann and Winter (2018) on the linguistic variables that convey narrow focus, reports that in this language listeners perceive a word as focused when it is marked by a rising accent type (L+H), and when it is located in nuclear position (at the end of an IP). For the acoustic variables, they report duration, maximum F0, and RMS intensity to also be reliable factors prompting positive prominence judgments, though not as important as the prosodic factors of accent type and position. They create a complex model that evaluates many prosodic and acoustic variables using the statistic mining algorithm 'Random forests' to observe the weight and contribution of each in predicting prominence judgments. They conclude that their results further confirm previous production studies on the language whereby focus is conveyed by a number of acoustic as well as prosodic factors. The findings from German that prosodic factors are equally as important in marking focus is discussed in the following.

Similarly, in Neapolitan Italian, studies show that speakers and listeners rely on accent type to differentiate focus conditions. A series of experiments by D'Imperio (1997a, 1997b) on broad and narrow focus in three sentential positions, show a number of patterns. In broad focus where the focus is on a larger constituent such as the utterance, a downstepped nuclear pitch accent of the shape !H* (a low peak succeeding a high peak) is used. In narrow focus, the pitch accent type H*+ L is used. Moreover, D'Imperio's studies report that F0 excursion size, peak alignment and duration are also used to convey focus in this variety. The author therefore maintains that focus interpretation is a very complex strategy in Italian, one that takes a number of factors into consideration.

In Mandarin Chinese, a language with contour tones, Xu (1999) shows that focus largely affects the F0 variations and durations of words. These F0 variations are mainly in the form of peak height and excursion size changes. A reported combined difference of around 81 Hz in height, and around 127 Hz in excursion size occurs between narrow and broad focus across positions. Regarding duration, the study reports a combined 25-55 Milliseconds difference between focus conditions in different positions. Under focus, positional differences were also reported in Mandarin that are similar to English discussed above. Xu shows that these factors are mainly the most significant reflexes of focus in this language, along with F0 compression in the post-focal area, which will be discussed shortly. In another study on Chinese, Xu and Xu (2005) report results of an analysis on rise duration, rise speed, excursion size and duration as an effect of focus prominence. They found there to be an effect of narrow focus on all those measures. That is, under focus, excursion size is larger, syllable duration is longer, and the speed of transition in bitonal accents is faster. They also contend that rise speed is highly correlated with excursion size, which led them to conclude that speakers deliberately use the rise spead measure to enhance the prominence of a unit.

5.2.1 Pre- and Post- Focus

What happens to the rest of the accents preceding and following the nuclear accent in narrow focus utterances is a point of much interest and research, as it also signals cross-linguistic differences. Phonetically, it is reported in many languages that after the nuclear accent is placed on a focused constituent- wherever that might be in the sentence- the rest of the string- pre and post focus- is acoustically realised in a notably compressed pitch range. This in AM is regarded as de-accentuation- pitch range compression or absence/deletion of F0 tonal events. Also called *Post Focus Compression* (PFC) whereby pitch range and amplitude are considerably reduced after the focal constituent (Xu, 2011). This phenomenon is reported in Lebanese Arabic (Chahal & Hellmuth, 2014, Chahal, 2001), Taif Arabic (Al-Zaidi, 2014), Mandarin Chinese (Xu, 2011, 1999), and most Germanic languages including English (Ladd, 2008). As well as in Moroccan, Kuwaiti, and Yemeni Arabic (Yeou et al., 2007b, b).

In Lebanese Arabic, Chahal (2014, 2001) finds that in narrow focus utterances, words following the nuclear accented target are realised in a very compressed pitch range with absent peak variations, indicative of PFC. Material preceding the target have their F0 height intact (though lower than target), which she believes is indicative of prenuclear accents. She further notes that even when the narrow focus utterance is produced as 2 prosodic phrases, the phrase preceding/ following the target phrase has its peaks lower, and its range significantly compressed.

In a similar vein, Taif Arabic as reported by (Al-Zaidi, 2014) contrastive focused targets in initial position see the post- focal string realised in a lower and compressed range than the broad focus condition. The author also reports that the mean intensity of the post- focal string is significantly less than its counterpart in broad focus. In medial position, the post- focal string was reported to also show significant intensity differences, with lower intensity in the post-focal region in contrastive focus. However, the study did not show significant pitch range differences in the post- focal string between the two focus conditions. The pre- focal string seems to be largely intact across focus conditions. Additionally, the results on Taif Arabic in Al-Zaidi, Xu and Xu (2019) report that the post focus region systematically shows less expanded excursion sizes, and lower F0 and intensity values, which they conclude as evidence for PFC in the dialect. On the hand, results on the pre- focus region show less systematic results.

In Kuwaiti, Yemeni, and Moroccan Arabic, Yeou et al. (2007a, b) show that there is variation in the realisation of pre- focal material across dialects. In this region, Yemeni and Kuwaiti Arabic pattern together in accenting pre-focal material (though realised with lower peaks than focused targets), whereas Moroccan markedly deaccents pre-focal material. As for post- focal material, the three dialects significantly deaccent/ compress the post-focal string.

English and Mandarin Chinese- although belonging to different language families- show a somewhat uniform pattern regarding pre- and post- focal material. In English (Xu & Xu, 2005) quantitatively demonstrate that the peaks of all post- focal words are significantly lowered after narrow focus compared to broad focus. The study also shows that the F0 peaks of the pre- focus string are unchanged. The authors therefore conclude that there are 3 pitch range adjustments as an effect of focus: “expansion under focus, compression after focus, and little or no change before focus” (p. 186). The same effect was reported for Mandarin Chinese in Xu (1999).

Finally, other phonological strategies are reported in other languages, which demonstrate an absence of de-accentuation altogether, whereby every word is accented regardless of focus condition, for example, Egyptian Arabic (Hellmuth, 2006a). Also reported for Italian (Avesani & Vayra, 2005). Suggesting that post focal compression does not take place in these languages. However, the author did find in subsequent acoustic experiments on Egyptian Arabic (Hellmuth, 2011, 2006b, 2014) that subject to the giveness status of the material, unfocused items in post- focal context were nevertheless realised with a notably reduced pitch range compared to a focused constituent. This was taken as evidence for PFC in this Arabic dialect.

5.3 Experiment design and recordings

The design of the focus material constitutes of lexically and grammatically identical declarative sentences as answers to questions in the form of a short dialogue. The use of a dialogue paradigm to elicit focus is encouraged in many studies (Al-Zaidi, 2014, Al-Zaidi, Xu & Xu, 2019, Jun & Fletcher, 2014, Chahal, 2011, Xu, 2011, Face, 2002). The target declarative sentence is controlled for length: a medium length sentence with 11 syllables, as advised in Wang and Xu (2011). The stimuli in the current experiment is influenced by the methodology in Chahal (2011), where the dependence is on question type to evoke the desired focus condition. If the question contains information contrasting 2 or more items, narrrow focus on one constituent is evoked. On the other hand, if the question is of a more general type, such as “what happened today”, broad focus on the whole sentence is evoked. Every target sentence corresponded to those specific questions that elicited the desired focus type and position. The target answer word is a disyllabic initially stressed sonorant noun /ˈlaa.na/ in a trochaic syllable structure. The choice of a proper noun to be the target in this experiment is to facilitate its placement in any location within the sentence without evoking syntactic ambiguity. The answer constituent is accordingly placed in three different sentential positions (initial-medial-final) in the two focus conditions, in order to assess whether the observed local intonation patterns are affected by word position (Face, 2002, Wang & Xu, 2011). As was discussed earlier in the chapter, this is a contrastive, constituent focus experiment in comparison with broad focus; a target is narrow focused when it is highlighted against one or two alternatives in the prompting question. The rest of the string is given information, as it is understood and repeated in the dialogue. In the following are details about the material as adapted from Chahal (2001)- note the focused item in brackets:

Narrow focus:

1. miin xaraj ma3 ritaal gabil 2ams? saja? 'Who went out with Ritaal two days ago, was it saja?
(initial) [laana] xarajat ma3 ritaal gabil 2ams.
2. ritaal xarajat ma3 xalid gabil 2ams?
(medial) ritaal xarajat ma3 [**laana]** gabil 2ams.
3. ritaal xarajat ma3 saja gabil 2ams?
(final) gabil 2ams ritaal xarajat ma3 [**laana].**

These targets were compared to their counterparts in the following:

Broad focus:

?esh Saar gabil 2ams? 'what happened two days ago (Tell me)?'
(initial) [laana] xarajat ma3 ritaal gabil 2ams. 'Laana went out with Ritaal two days ago'.
(medial) ritaal xarajat ma3 [**laana]** gabil 2ams.
(final) gabil 2ams ritaal xarajat ma3 [**laana].**

Figure 42: Narrow focus (Top) and Broad focus material (bottom).

As can be noted, the material is designed to yield comparable and controlled targets- controlled in sentence length and segmental composition- in different sentential positions and under different focus conditions. Each identical target pair (broad/narrow) is examined to observe the phonetic manifestation of the categorical differences that occur between the two conditions. A total of 360 utterances were collected for this experiment:

2 focus positions * 3 sentential positions (initial, medial and final) * 3 repetitions * 20 speakers.

An overall number of (51) utterances were excluded. These utterances were eliminated for reasons such as:

(a) Wrong speech rate, e.g. utterances produced quickly, or uncarefully.

(b) Producing the wrong type of focus on target, as compared to the rest of the productions.

(c) Technical reasons, such as noisy backgrounds.

A total of 309 utterances were analysed as a result.

5.3.1 Recordings

Studies on information structure regarding focus in intonational frameworks are constructed mainly on controlled lab data. This format as was seen allows the researcher to control for specific variables in order to obtain ultimate results from participants. For example, this type of lab speech can be controlled for sentence and syllable count- 11 syllables in my data (Xu & Xu, 2005). It can also be controlled for speech rate and segmental composition in some studies. In addition to tune and/or melody choice in others.

Interacting with the researcher, the subjects were instructed to read target material as normal as possible at normal speech rate. For the narrow focus material, they were specifically instructed to emphasise the target (Xu & Xu, 2005, Wang & Xu, 2011, De Jong & Zawaydeh:, 2002). The questions and answers in the dialogue were randomised and scripted on a PowerPoint slide, the questions were read by the researcher and the answers were read by the participant. The dialogue also contained 20 filler sentences in between.

5.3.2 Quantitative and qualitative measurements

With reference to the F0 contour and spectrogram in Praat version: 6.0.10 (Boersma & Weenink, 2016), the data was labelled using a modified ToBI transcription for Arabic. To aid subsequent qualitative and quantitative analyses, target sentences were transcribed to mark the occurrence of pitch accents, phrase and boundary tones. That is, to observe whether a target is pitch accented, and how this accent is realised compared to the surrounding string. This also includes the shape of the accent and observing the phonological distribution of accents (relative prominence) within the target sentence.

A Praat script developed with the help of supervisor Jalal Al-Tamimi was used to automatically obtain pitch analyses and other measurements. For the pitch analyses, a two-pass method was used to allow proper computations of the pitch that are adapted to speaker gender and range. The two-pass method (Hirst, 2011) starts by computing the general pitch contour with default settings, i.e., 50 Hz floor and 600 ceiling. Then the first and third quartiles are obtained and these are multiplied by a coefficient. These new values are then added as the new floor (minima) and ceiling (maxima). These are adapted to each speaker's range and allow for an automated method of correction. For more details and a sample of the script, see: https://github.com/JalalAl-Tamimi/Praat-f0-Accurate-Estimation. Subsequently, automated F0 contours are extracted, corrected and smoothed. Smoothing and correction of F0 contours helps eliminate pitch irregularities and jumps caused by microprosodic factors, and only maintain the 'linguistically' significant pitch events. These corrections were carried out throughout the whole corpus.

In the script, the target word in each focus condition was the domain for phonetic measurements. The F0 maxima are identified where the highest F0 point occurs within a domain (vowel: /aa/ in /laana/, and /aa/ in /ritaal/, syllable: /laa/ in /laana/ and /taal/ in /ritaal/, word /laana/ and /ritaal). The minima are identified with reference to the maxima using an intermediate measure named "SyllableMaxTime". The script would locate the time point of the Max in the syllable then compute the lowest point within the domain preceding the located Max F0 in Hertz and Semitones.

The target stressed syllable was measured for F0 maxima and minima in Hertz and Semitones, duration, intensity, excursion size, rise time and rise speed. The whole word was measured for excursion size in Semitones (F0Max – F0min) and F0 maxima and minima in Hertz and Semitones. The stressed vowel was measured for duration, F0 maxima and minima in Hertz and Semitones. The semitone conversion is: 1 ST = 100 HZ. Moreover, the whole sentence excluding target word was considered for phonetic measurements. The pre and post-focal target /riˈtaal/ was measured for F0 maxima and minima, intensity as well as excursion size and duration. These measurements were obtained with the following aims:

Duration (vowel, syllable): to assess the quantitative effect of focus type narrow/ broad on length.

F0 Hz/ST (vowel, syllable, word, sentence): to assess peak height and range as a result of focus type.

Pitch range Hz/ST (syllable, word, sentence): to assess effect of focus on excursion size of target word and pre/post string.

Syllable Rise speed (ST/S) and rise duration (MS): to assess the effect of focus on the rise speed in rising tones L+ H*.

Intensity dB (vowel, pre/post focal string): to assess the effect of focus on amplitude.

5.3.3 Experiment questions

The focus experiment comparing the two focus conditions was carried out to investigate the following research questions:

1. Is focus encoded intonationally in this Arabic dialect?
2. How are narrow and broad focus utterances distinguished phonetically?
   - What quantitative methods (duration, F0, intensity, rise time, rise speed) are used to mark focus in JA that makes it distinguishable from broad focus?
   - Are these quantitative methods affected by the target word's position/ location?
   - What happens quantitatively to the pre/post focus /ritaal/ under narrow and broad focus?
3. Qualitatively, how does focus affect accent distribution in target utterances (phrasing, accent type, accent distribution)? And what happens qualitatively to the pre/post focal material?
4. Having observed and analysed the results for each focus type, which of the quantitative or qualitative variables can be said to indicate/ predict a particular focus interpretation?

5.4 Results

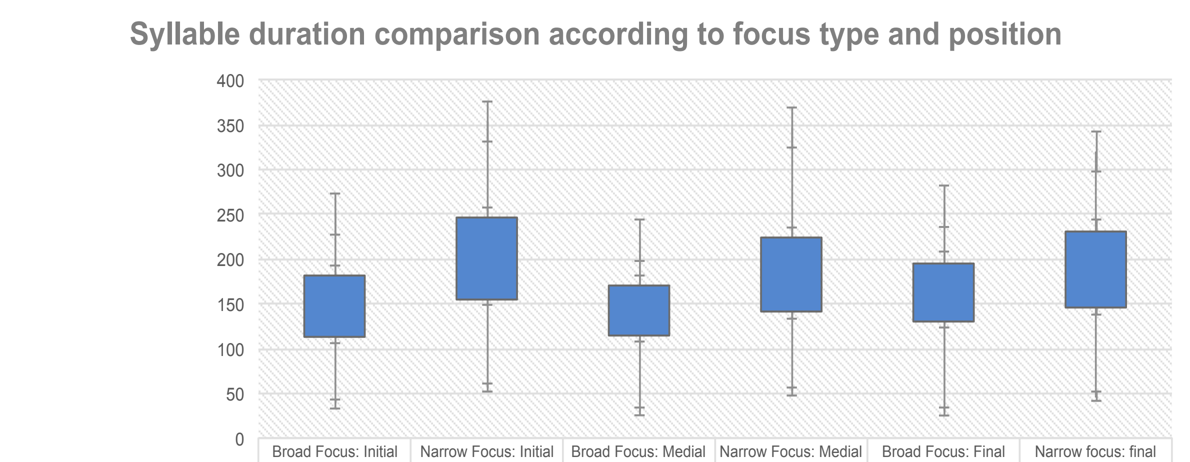
5.4.1 Comparison between narrow and broad focus

First in the analysis, the comparison between narrow and broad focus patterns is presented, together with detailed analyses of the effect of focus on each quantitative/ phonetic measure. The analysis in this section aims to figure out which of the quantitative or qualitative variables can be said to be indicative of, or the most significant predictor of a particular focus interpretation. The following tables and figures demonstrate the comparison between the two focus types according to different quantitative measures and sentential position.

Summary statistics

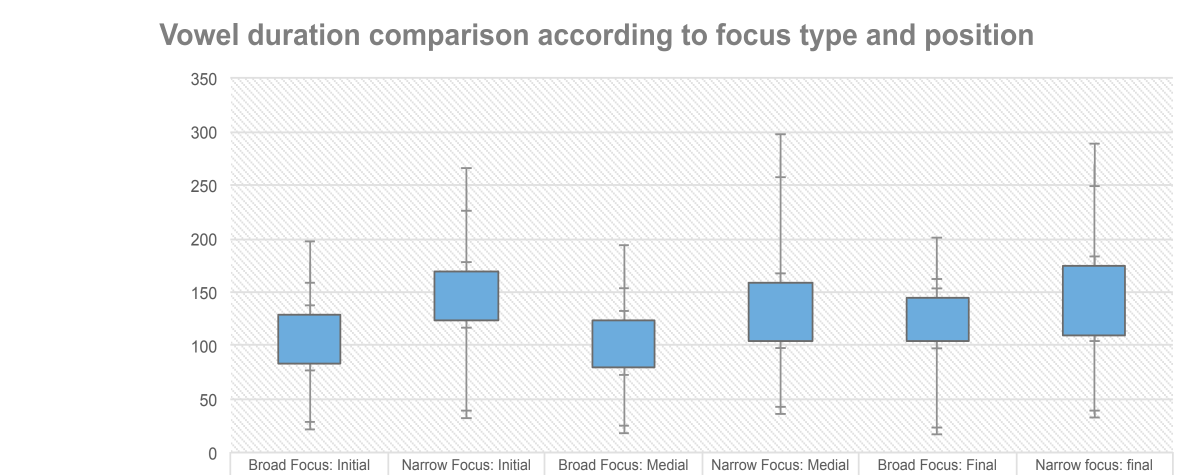
	BF		NF	
	mean	std. deviation	mean	std. deviation
v1	132	24.1	168	37.2
v2	182	34.3	233	51.1
v3	194	57	213	55
v4	10.7	5	12.5	4.52
v5	194	56.7	213	54.8
v6	10.7	5.25	12.5	4.53
v7	34.6	23.3	47.8	23.8
v8	3.38	2.03	4.62	2.50
v9	69.1	4.85	70.9	4.49
v10	94.3	101	115	55.4
v11	32.9	15.6	42.7	18.1
v12	197	58.72	214.46	57.65
v13	11	5	13	4.65
Variables in this table: v1: phonemedurationMS, v2: syllabledurationMS, v3: SyllablemaxF0 Hz, v4: SyllablemaxF0ST, v5: phonememaxF0, v6: phonememaxF0ST, v7: SyllableRangeF0Hz, v8: SyllableRangeF0ST, v9: IntensitymaxdB, v10: SyllableRiseTimeMS, v11: RiseSpeedSTsec, v12: WordmaxF0 Hz, v13: WordmaxF0ST				

Table 10: summary statistics for the comparison between broad and narrow focus in all positions.



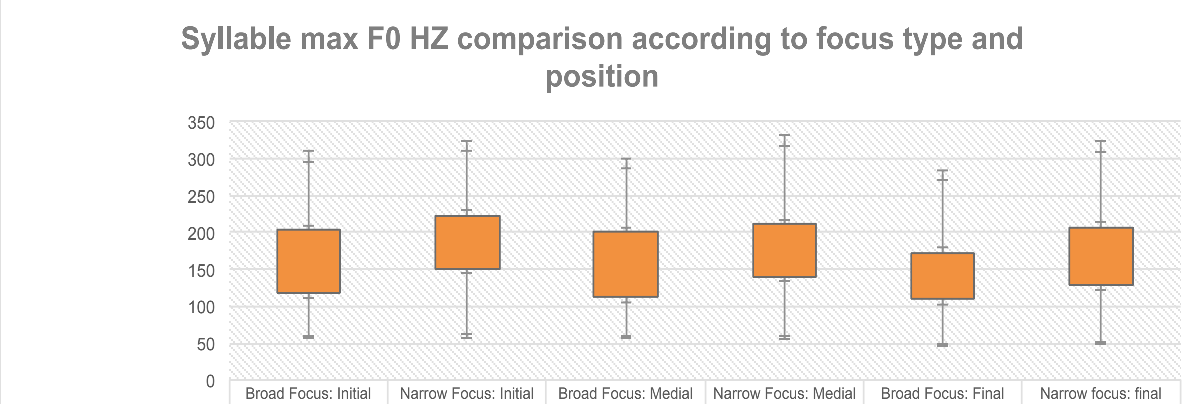
	Broad Focus: Initial	Narrow Focus: Initial	Broad Focus: Medial	Narrow Focus: Medial	Broad Focus: Final	Narrow focus: final
Average of SYLLABLEduration	181.22	245.72	170.25	223.98	196.035	231.67
StdDev SYLLABLEduration	37.67	56.13	30.13	51.88	28.84	46.26
Max. of SYLLABLEduration	250	353.41	220.61	346.76	258.57	320.49
Min. of SYLLABLEduration	113.049	155.85	114.14	140.35	130.014	144.92

Figure 43: NF/BF syllable duration comparison by position.



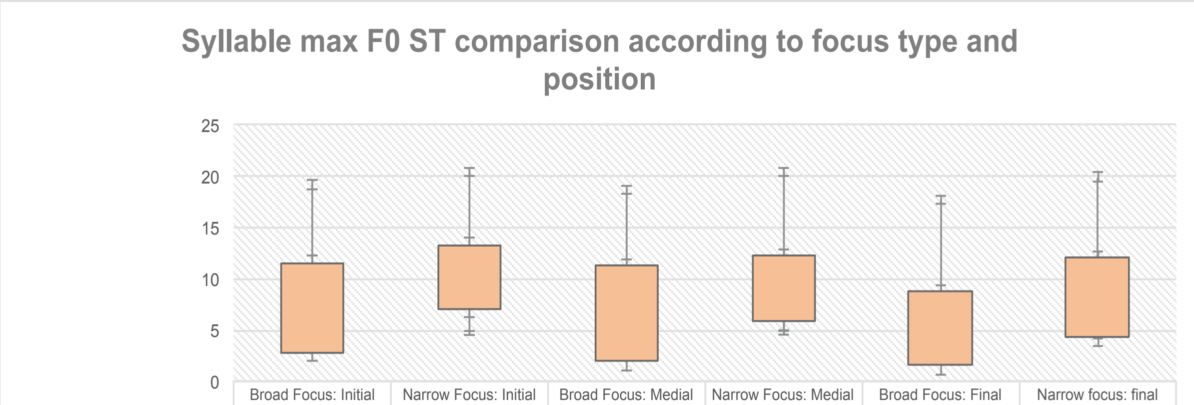
	Broad Focus: Initial	Narrow Focus: Initial	Broad Focus: Medial	Narrow Focus: Medial	Broad Focus: Final	Narrow focus: final
AveragePHONEMEduration	129.65	169.88	123.21	158.88	145.34	174.84
StdDev PHONEMEduration	25.012	35	21.06	39.47	20.41	36.15
Max. of PHONEMEduration	177.8	245.85	173.41	277.24	181.64	268.9
Min. of PHONEMEduration	83.59	123.56	79.025	104.54	104.26	110.3

Figure 44: NF/BF Vowel duration comparison by position.



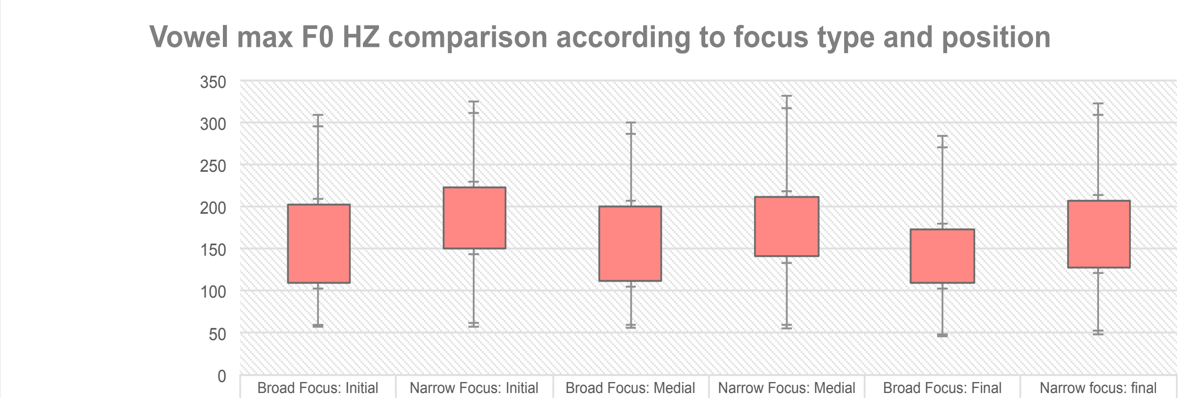
	Broad Focus: Initial	Narrow Focus: Initial	Broad Focus: Medial	Narrow Focus: Medial	Broad Focus: Final	Narrow focus: final
AverageSyllableMaxF0Hz	203.69	222.87	201.16	211.015	173.1	207.15
StdDev SyllableMaxF0Hz	58.33	59.35	58.4	57.59	48.1	49.83
Max. of SyllableMaxF0Hz	302.69	317.84	294.12	324.63	278.039	316.15
Min. of SyllableMaxF0Hz	117.94	150.93	112.033	140.59	109.18	128.62

Figure 45: NF/BF syllable max F0 HZ comparison by position.



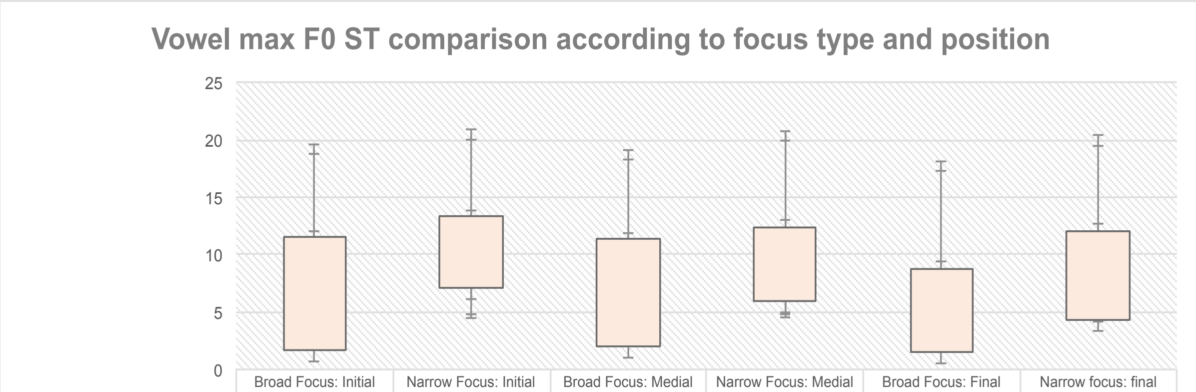
	Broad Focus: Initial	Narrow Focus: Initial	Broad Focus: Medial	Narrow Focus: Medial	Broad Focus: Final	Narrow focus: final
Average SyllableMaxF0ST	11.57	13.32	11.32	12.33	8.84	12.096
StdDev SyllableMaxF0ST	5.17	4.66	5.37	4.7	4.84	4.33
Max. of SyllableMaxF0ST	19.17	20.46	18.67	20.38	17.7	19.92
Min. of SyllableMaxF0ST	2.85	7.12	1.96	5.89	1.52	4.35

Figure 46: NF/BF syllable max F0 ST comparison by position.



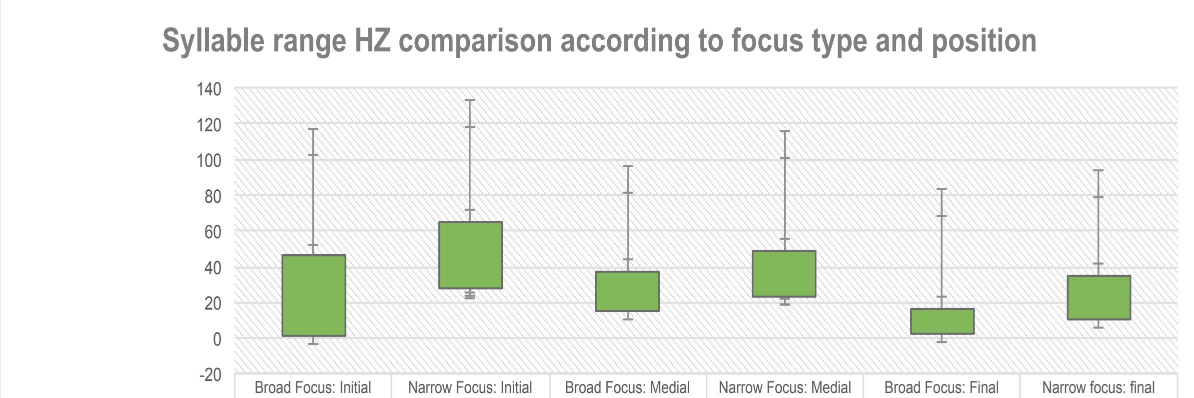
	Broad Focus: Initial	Narrow Focus: Initial	Broad Focus: Medial	Narrow Focus: Medial	Broad Focus: Final	Narrow focus: final
Average phonememaxF0 HZ	202.87	222.67	201.025	211.015	172.16	206.58
StdDev phonememaxF0 HZ	58.49	59.04	58.3	57.59	47.11	49.8
Max. of phonememaxF0 HZ	302.69	317.84	294.12	324.63	278.039	316.15
Min. of phonememaxF0 HZ	109.77	150.93	112.033	140.59	109.18	128.62

Figure 47: NF/BF vowel max F0 HZ comparison by position.



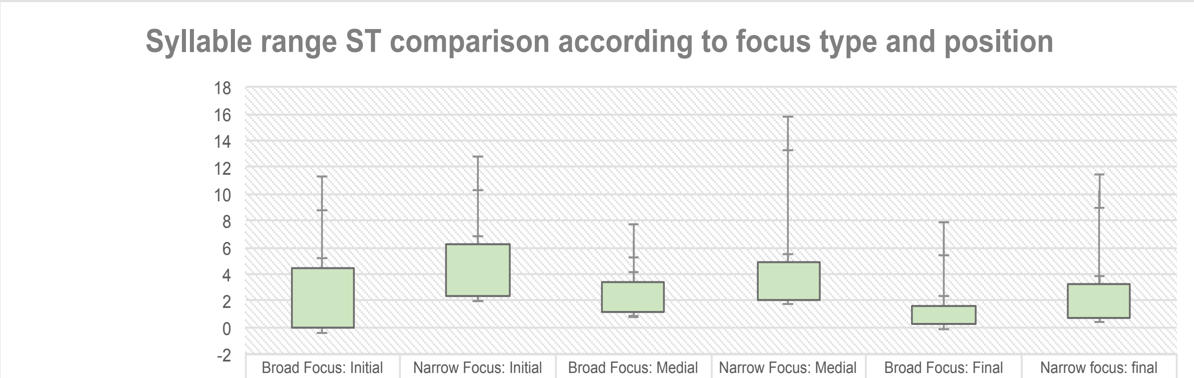
	Broad Focus: Initial	Narrow Focus: Initial	Broad Focus: Medial	Narrow Focus: Medial	Broad Focus: Final	Narrow focus: final
Average phonememaxF0ST	11.49	13.31	11.31	12.33	8.76	12.048
StdDev phonememaxF0 ST	5.22	4.64	5.37	4.7	4.77	4.33
Max. of phoneme maxF0 ST	19.17	20.46	18.67	20.38	17.7	19.92
Min. of phonememaxF0 ST	1.61	7.12	1.96	5.89	1.52	4.35

Figure 48: NF/BF vowel max F0 ST comparison by position.



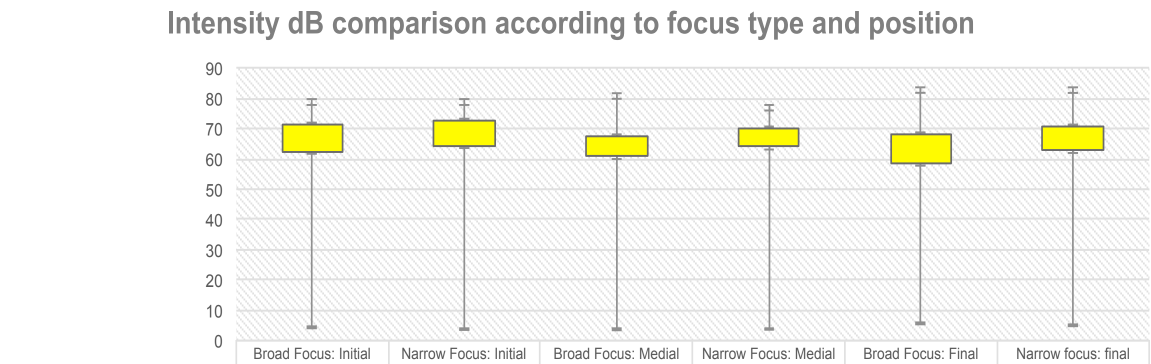
	Broad Focus: Initial	Narrow Focus: Initial	Broad Focus: Medial	Narrow Focus: Medial	Broad Focus: Final	Narrow focus: final
Average SyllableRangeF0 Hz	45.87	65.036	36.9	48.72	16.25	34.85
StdDev SyllableRangeF0 Hz	24.74	23.98	18.093	20.6	13.53	18.011
Max. of SyllableRangeF0 Hz	110.033	125.33	88.96	108.19	75.77	85.88
Min. of SyllableRangeF0 Hz	1.11	28.13	14.82	23.12	2.104	10.18

Figure 49: NF/BF syllable range F0 HZ comparison by position.



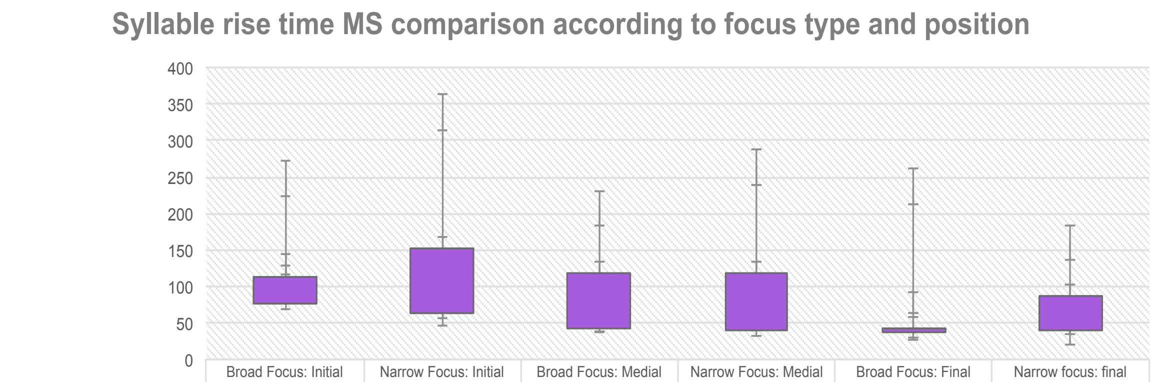
	Broad Focus: Initial	Narrow Focus: Initial	Broad Focus: Medial	Narrow Focus: Medial	Broad Focus: Final	Narrow focus: final
Average SyllableRangeF0 ST	4.51	6.26	3.46	4.91	1.69	3.2
StdDev SyllableRangeF0 ST	2.29	2.26	1.088	2.64	1.15	1.61
Max. of SyllableRangeF0 ST	10.007	11.53	6.5	14.52	6.66	10.21
Min. of SyllableRangeF0 ST	0	2.39	1.19	2.12	0.28	0.79

Figure 50: NF/BF syllable range F0 ST comparison by position.



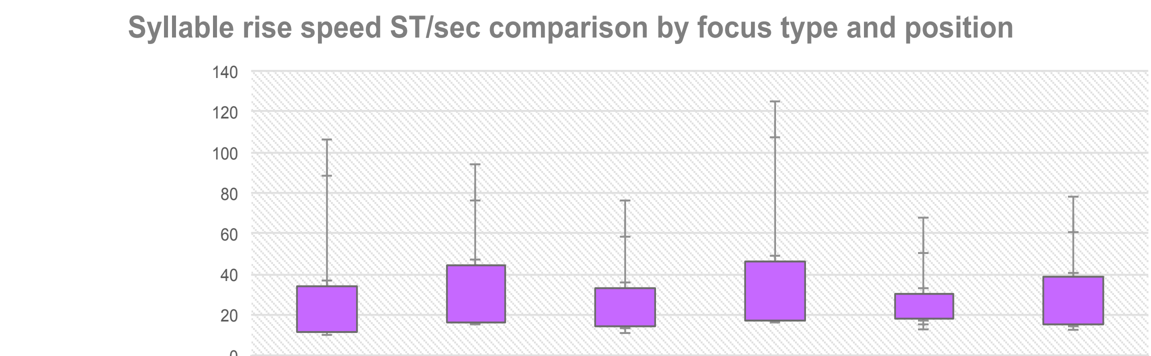
	Broad Focus: Initial	Narrow Focus: Initial	Broad Focus: Medial	Narrow Focus: Medial	Broad Focus: Final	Narrow focus: final
Average phonemeintensitymax dB	71.23	72.59	67.47	69.79	67.84	70.6
StdDev phonemeintensitymax dB	4.27	3.85	3.65	3.92	5.71	5.084
Max. of phonemeintensitymax dB	78.59	78.75	80.85	76.99	82.83	82.8
Min. of phonemeintensitymax dB	62.58	64.21	60.85	63.99	58.54	62.85

Figure 51: NF/BF intensity comparison by position.



	Broad Focus: Initial	Narrow Focus: Initial	Broad Focus: Medial	Narrow Focus: Medial	Broad Focus: Final	Narrow focus: final
Average SyllableRiseTimeMS	112.64	152.65	117.97	118.015	41.99	86.51
StdDev SyllableRiseTimeMS	130.3	60.051	52.095	53.83	77.33	34
Max. of SyllableRiseTimeMS	248.18	338.43	207.72	263.48	237.025	160.67
Min. of SyllableRiseTimeMS	75.41	63.06	43.52	38.71	36.38	40.2

Figure 52: NF/BF rise time comparison by position.



	Broad Focus: Initial	Narrow Focus: Initial	Broad Focus: Medial	Narrow Focus: Medial	Broad Focus: Final	Narrow focus: final
Average Syllable RiseSpeedST/sec	34.52	44.25	33.17	46.4	30.35	38.48
StdDev syllable RiseSpeedST/sec	18.43	17.11	12.49	22.26	14.29	14.074
Max. of syllable RiseSpeedST/sec	97.21	85.023	67.22	116.24	59.22	69.52
Min. of syllableRiseSpeedST/sec	11.12	15.93	14.1	17.25	18.041	15.28

Figure 53: NF/BF syllable rise speed ST/sec comparison by position.

A series of one-way Anovas were carried out to first observe the significance of the mean differences between the two focus types/conditions as indicated in the previous table. There isn't a significant difference between the average of variable 10: RiseTime between the two focus types (v10: F=2.142, p=0.145). For all of the other variables, there is a significant difference between the averages of the respective variables under the two focus types. Note the following table:

ANOVA results					
			Dependent variable: 'phonemedurationMS (v1)'		
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
focusType	1	215867	215867	69.16	9.63e-15***
Residuals	219	683559	3121		
			Dependent variable: 'syllabledurationTier5MS (v2)'		
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
focusType	1	219400	219400	70.65	5.43e-15***
Residuals	219	680064	3105		
			Dependent variable: 'phonememaxF0 (v5)'		
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
focusType	1	28262	28262	7.104	0.00826**
Residuals	219	871208	3978		
			Dependent variable: 'phonememaxF0ST (v6)'		
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
focusType	1	28587	28587	7.189	0.0079**
Residuals	219	870883	3977		
			Dependent variable: 'SyllableRangeF0Hz (v7)'		
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
focusType	1	71715	71715	18.97	2.04e-05**
Residuals	219	827755	3780		
			Dependent variable: 'SyllableRangeF0ST (v8)'		
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
focusType	1	56736	56736	14.74	0.000161***
Residuals	219	842734	3848		
			Dependent variable: 'IntensitymaxdB (v9)'		
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
focusType	1	36785	36785	9.338	0.00252**
Residuals	219	862685	3939		
			Dependent variable: 'SyllableRiseTimeMS (v10)'		
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
focusType	1	8714	8714	2.142	0.145
Residuals	219	890756	4067		
			Dependent variable: 'RiseSpeedSTsec (v11)'		
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
focusType	1	58321	58321	15.18	0.00013***
Residuals	219	841149	3841		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 11: anova results for the comparison between broad and narrow focus

According to these results it can be said that narrow focus targets are longer in vowel and syllable durations, show higher peaks and excursion sizes, higher intensity and faster transitions of units per second than their broad focus counterparts. However no significant differences are reported for v10: syllablerisetimeMS. This is confirmed in the post-hoc pairwise comparisons:

Results of Tukey HSD Test					
		diff	lwr	upr	p adj
v1		63.97019	48.81	79.13039	0
v2		64.49151	49.37011	79.6129	0
v3		28.14642	6.031388	40.26146	0.0082628
v4	nf-bf	23.27912	6.167279	40.39096	0.007895
v5		36.87125	20.18849	53.554	2.04e-05
v6		32.79551	15.96249	49.62852	0.0001612
v7		26.40702	9.375907	43.43813	0.0025231
v8		33.25047	16.43329	50.06765	0.0001296

Variables in this table: v1: PHONEMEdurationMS, v2: SYLLABLEdurationTier5MS, v3: phonememaxF0, v4: phonememaxF0ST, v5: SyllableRangeF0Hz, v6: SyllableRangeF0ST, v7: phonemeintensitymaxdB, v8: RiseSpeedSTsec

Table 12: post-hoc results for the comparison between broad and narrow focus

In order to evaluate which of the quantitative variables is the most significant in the marking of a particular focus type/condition across gender and speakers, mixed effects logistic regression analysis of the relationship between them was carried out. The employment of logistic regression was to observe the relationship between focus type/condition: NF, BF and the acoustic predictor variables. In particular, the analysis was carried out to figure the probability of a certain focus interpretation as a function of those predictors. In the model: the Outcome is the binomial focus condition/type: Narrow focus/broad focus. Variables (v1-9: phoneme duration, syllable duration, phonememaxF0Hz, phonememaxF0ST, syllablerangeF0Hz, syllablerangeF0ST, phonemeintensitymax, syllablerisetime, rise speed) + gender, repetition + position (initial/medial/final) are fixed effects, while speaker + word are random effects. Visual inspection of residual plots did not reveal any obvious deviations from homoscedasticity or normality.

Firstly, a model was run which includes all of the independent variables (v1-9). But, values regarding the variables named (v3), (v4), (v5), and (v6) were quite high. The correlation between the variable named (v3: phonememaxF0Hz) and the variable named (v4: phonememaxF0ST) was quite high and significant (corr=0.992, $p<2.2e-16$). Furthermore, the correlation between the variable named (v5: syllablerangeF0Hz) and the variable named (v6: syllablerangeF0ST) was also quite high and significant (corr=0.856, $p<2.2e-16$). Therefore, the variables in Hertz named (v3) and (v5) are removed from the model. Then the new model is run. According to this model, the two most significant variables are: (v9: risespeedSTsec) (1.31e-05), and (v2: syllabledurationMS) (p=0.017). As a result one can say that each one-unit change in (v2) will increase the log odds of in favor of narrow focus by (0.032). Also, each one-unit change in (v9) will increase the log odds of in favor of narrow focus by (0.061).

Deviance Residuals:					
	Min	1Q	Median	3Q	Max
	-2.1549	-0.7191	-0.2815	0.5664	2.7275
Coefficients					
	Estimate	Std. Error	z value	Pr (>|z|)	VIF
(Intercept)	-1.608e+01	3.353e+00	-4.795	1.63e-06 ***	
PHONEMEdurationMS (v1)	9.885e-03	1.737e-02	0.569	0.5692	6.068022
SYLLABLEdurationMS (v2)	3.206e-02	1.345e-02	2.383	0.0172 *	6.823573
phonememaxF0ST (v4)	2.235e-02	3.804e-02	0.588	0.5568	1.082605
SyllableRangeF0ST (v6)	9.949e-03	1.210e-01	0.082	0.9345	1.895930
phonemeintensitymaxdB (v7)	7.196e-02	4.163e-02	1.729	0.0839 .	1.244582
SyllableRiseTimeMS (v8)	-6.277e-04	3.610e-03	-0.174	0.8619	1.712858
RiseSpeedSTsec (v9)	6.184e-02	1.419e-02	4.359	1.31e-05 ***	1.514188
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					
Null deviance=296.30, Residual deviance=189.83, AIC=205.83					

Table 13: regression results for the comparison between broad and narrow focus

5.4.2 Detailed results for each focus condition

5.4.2.1 Broad Focus intonational patterns

Observing broad focus targets reveals that the target can be allocated an H* or a L+H* accent, similar to its narrow focus target. Broad focus utterances were generally produced as a single Intonational Phrase, indicating that phrasing is not used as a prosodic resource in the broad focus condition. There was variation regarding which words received prominence in terms of pitch accents and nuclear accents. According to position, the target /ˈlaa.na/ was sometimes nuclear accented and usually pitch accented, while a nuclear accent was more often placed on the other proper noun in the utterance /riˈtaal/ or on the adverbial phrase /gabil ?ams/.

In medial positions, when the target /ˈlaa.na/ is nuclear accented, pre-nuclear accents are also observed. When the /`laa.na/ is pitch accented in initial position, a nuclear accent is placed on another word in the string.

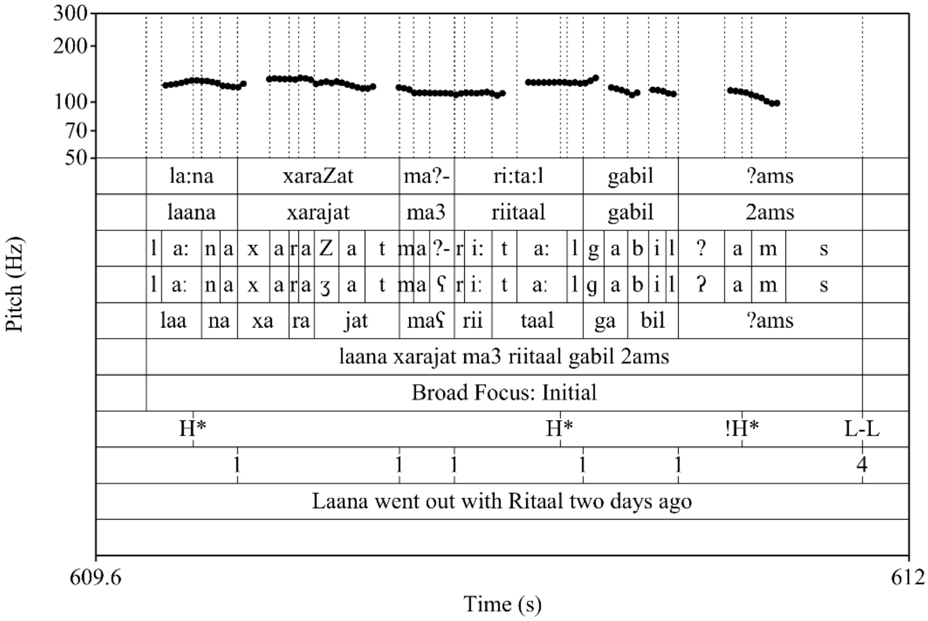


Figure 54: Broad focus utterance with target in initial position. Speaker 04_M

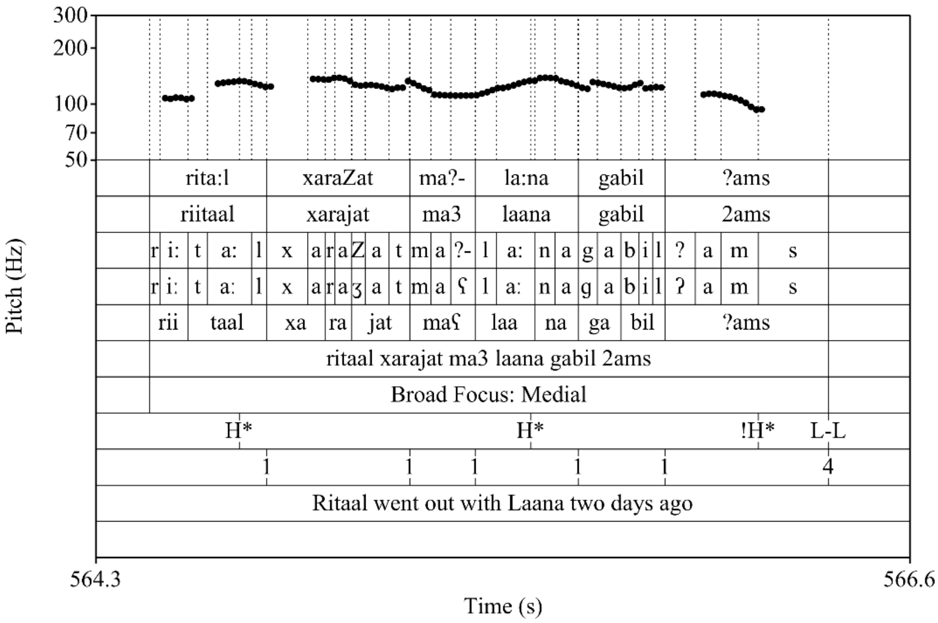


Figure 55: Broad focus utterance with target in medial position. Speaker 02_M

In final position, many patterns emerge. The most common pattern observed is allocating the target a nuclear accent, while allowing for pre-nuclear accents as well.

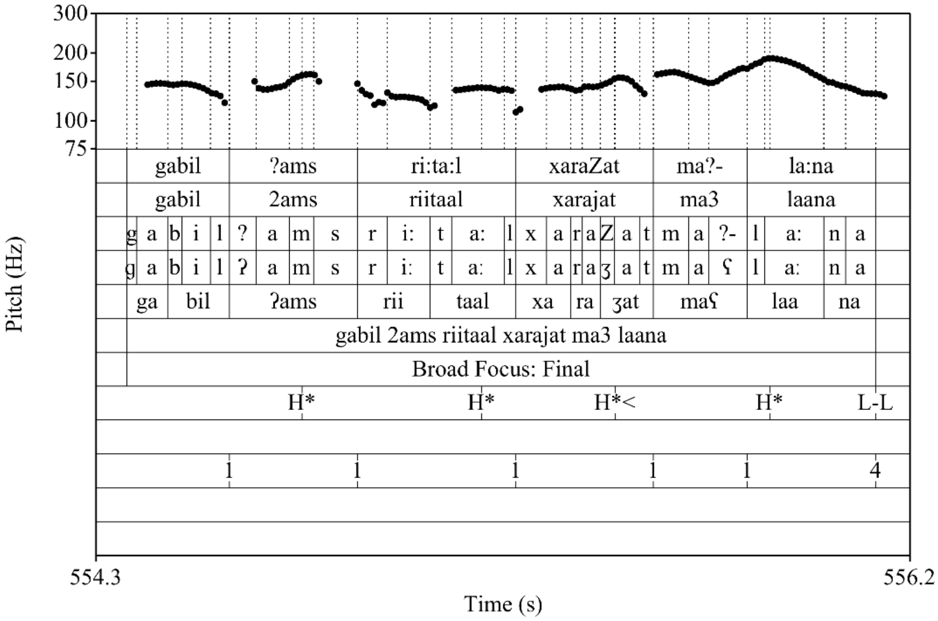


Figure 56: Broad focus utterance with target nuclear accented in final position. Speaker 01_M

5.4.2.2 Quantitative detail

Broad focus utterances, as mentioned previously, are produced in one whole Intonational Phrase. The relationships between the durations, intensity, peaks and excursions are reported accordingly. Throughout the analysis, reference to the acoustic variables will be abbreviated by a ‘v’ and a sequence number accordingly. The analyses will include results for both the target word /laana/ and secondary target word /ritaal/. All throughout the statistical analyses in this thesis, the assumption of normal distribution is considered. The assumption of normal distribution was tested by using the Shapiro-Wilk's test. The assumption that homogeneity of variance within the population is equal was also considered in order to conduct relative tests. Tests of homogeneity were conducted by using the Levene test. When satisfied, sets of Analysis of variance tests were carried out accordingly. When violated, firstly the outlier values were removed. Then the assumptions were checked again. If the assumptions were still violated, rank transformations of these variables were taken into consideration. In the analyses gender and position are included as independent variables. Alongside those detailed tests, logistic regression tests were carried out taking into consideration variations along the many dimensions included in the experiment in an attempt to construct a more generlised vision of the data.

1. Duration

This table shows the vowel and syllable durations when target is placed in different sentential positions. As can be noticed the vowel and syllable are longest in final position, followed by initial and medial.

Summary statistics

Laana	Initial		Medial		Final	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
v1	130	25.0	123	21.1	145	20.4
Gender	**F 136.3**	**M 122.7**	**F 124.7**	**M 121.4**	**F 155.4**	**M 138.7**
v2	181	37.7	170	30.1	196	28.8
Gender	**F 193**	**M 169**	**F 173.4**	**M 166.6**	**F 213.8**	**M 184.4**
v1: Phonemedurationms, v2: Syllabledurationms						

Table 14: duration summary statistics for /laana/ under broad focus.

Two-way Anovas by position and gender were carried out to investigate whether the difference in means is statistically relevant.

Anova results

Dependent variable: 'PhonemedurationMS (v1)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	2679	2679	5.577	0.0197 *
position	2	11781	5890	12.262	1.34e-05 ***
gender:position	2	1013	506	1.054	0.3515
Residuals	128	61489	480		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
Dependent variable: 'SyllabledurationMS (v2)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	10117	10117	10.210	0.001759 **
position	2	16368	8184	8.259	0.000423 ***
gender:position	2	2850	1425	1.438	0.241205
Residuals	128	126835	991		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 15: duration Anova results for /laana/ under broad focus

As can be inferred from the tables, these mean differences in durations by sentential positions and gender are indeed significant individually, however relationships between position and the variables does not depend on gender. The same is carried out for the secondary target word /ritaal/ under broad focus:

Ritaal	Initial		Medial		Final	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
v1	125	26.2	131	36.8	130	29.5
Gender	**F 129.3**	**M 121.3**	**F 135.4**	**M 125**	**F 129**	**M 131**

v2	251	48.6	241	56.3	244	51.1
Gender	**F 269.5**	**M 232.2**	**F 251.4**	**M 229.4**	**F 253.8**	**M 237.5**
v1: Phonemedurationms, v2: Syllabledurationtierms						

Table 16: duration summary statistics for /ritaal/ under broad

Anova results

Dependent variable: 'PhonemedurationMS (v1)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	471	470.6	0.304	0.582
position	2	1154	577.0	0.373	0.690
gender:position	2	607	303.3	0.196	0.822
Residuals	128	198255	1548.9		
Dependent variable: 'SyllabledurationMS (v2)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	13232	13232	9.235	0.00288 **
position	2	3137	1568	1.095	0.33780
gender:position	2	708	354	0.247	0.78146
Residuals	128	183412	1433		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 17: duration anova results for /ritaal/ under broad focus

For this word, there is not an interaction nor effect of gender or position, except for (v2).

As for the comparison between the two words in this case, according to pairwise comparisons, there is a significant difference between words (ritaal-laana: diff=-18.8209, $p<0.05$) for (v1). For (v2), there is also a significant difference according to word (v2: $F=2.623$, $p<0.05$).

2. **Pitch variables**

F0 Peak height

The following are the peak height results in Hertz and Semitones.

Summary statistics

Laana	Initial		Medial		Final	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
v3	203	58.5	201	58.3	172	47.1
Gender	**F 256**	**M 147.7**	**F 250.6**	**M 144**	**F 218.1**	**M 142.1**
v4	11.5	5.22	11.3	5.37	8.77	4.77
Gender	**F 16.2**	**M 6.6**	**F 15.8**	**M 6**	**F 13.3**	**M 5.7**
v3: phonememaxF0HZ, v4: maxF0ST						

Table 18: peak height summary statistics for /laana/ under broad focus

Anova results

Dependent variable: 'phonememaxF0 (v3)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	147049	147049	422.440	< 2e-16 ***
position	2	6860	3430	9.853	0.000105 ***
gender:position	2	2033	1017	2.920	0.057518 .
Residuals	128	44556	348		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
Dependent variable: 'maxF0ST (v4)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	147049	147049	422.440	< 2e-16 ***
position	2	6860	3430	9.853	0.000105 ***
gender:position	2	2033	1017	2.920	0.057518 .
Residuals	128	44556	348		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 19: peak height anova results for /laana/ under broad focus

As can be noticed, the pitch variables show significant differences according to gender and position, as well as an interaction between the two in signaling the variables.

Ritaal	Initial		Medial		Final	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
v3	216	60.1	219	67.2	195	60.0
Gender	**F 268.8**	**M 160.6**	**F 278.4**	**M 151.2**	**F 260.8**	**M 152.8**
v4	12.6	5.12	12.7	5.75	10.8	5.48
Gender	**F 17**	**M 8**	**F 17.7**	**M 7**	**F 16.5**	**M 7**
v3: phonememaxF0HZ, v4: maxF0ST						

Table 20: peak height summary statistics for /ritaal/ under broad focus

For /ritaal/, there are also differences according to position. He results for their significance and interaction with gender are in the following:

Anova results

Dependent variable: 'phonememaxF0 (v3)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	443191	443191	741.091	<2e-16 ***
position	2	1506	753	1.259	0.287
gender:position	2	2631	1316	2.200	0.115
Residuals	128	76547	598		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
Dependent variable: 'maxF0ST (v4)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	149445	149445	395.454	<2e-16 ***
position	2	1108	554	1.466	0.235
gender:position	2	1573	786	2.081	0.129
Residuals	128	48372	378		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 21: peak height anova results for /ritaal/ under broad focus

As can be seen, for this word there is a gender difference, but not an effect of position. Regarding the comparison between the two words in this case, according to pairwise comparisons, there is a significant difference between words (ritaal-laana: diff=-18.8209, p<0.05) for (v3). For (v4), there is also a significant difference according to word (v2: F=2.623, p<0.05).

Excursion size

In the following are the syllable excursion sizes for the targets under broad focus.

Summary statistics

Laana	Initial			Medial			Final		
		mean	std. deviation		mean	std. deviation		mean	std. deviation
v5		45.9	24.7		36.9	18.1		16.3	13.5
Gender	**F**	**57**	**M 34.2**	**F**	**47**	**M 25.2**	**F**	**21.7**	**M 12.6**
v6		4.51	2.30		3.47	1.09		1.70	1.16
Gender	**F**	**4.50**	**M 4.52**	**F**	**3.5**	**M 3.3**	**F**	**1.85**	**M 1.60**
v5: SyllableRangeF0Hz, v6: SyllableRangeF0ST									

Table 22: excursion size summary statistics for /laana/ under broad focus

Anova results

				Dependent variable: 'SyllableRangeF0Hz (v5)'	
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	34817	34817	43.70	9.36e-10 ***
position	2	62123	31062	38.99	5.99e-14 ***
gender:position	2	1578	789	0.99	0.374
Residuals	128	101980	797		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
				Dependent variable: 'SyllableRangeF0ST (v6)'	
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	2680	2680	2.789	0.0974 .
position	2	74785	37393	38.912	6.28e-14 ***
gender:position	2	29	14	0.015	0.9851
Residuals	128	123003	961		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 23: excursion size anova results for /laana/ under broad focus

Notice that there is a significant difference in means for this variable according to gender and position. Now compare those for the secondary target word in the following.

Ritaal	Initial		Medial		Final	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
v5	62.5	23.2	63.7	29.4	51.9	24.6
Gender	**F 78**	**M 46.6**	**F 83.7**	**M 40.7**	**F 67.7**	**M 41.6**
v6	5.91	1.69	5.83	1.86	5.26	1.83
Gender	**F 6**	**M 5.84**	**F 6.25**	**M 5.33**	**F 5.29**	**M 5.22**
v5: SyllableRangeF0Hz, v6: SyllableRangeF0ST						

Table 24: excursion size summary statistics for /ritaal/ under broad focus

Anova results

Dependent variable: 'SyllableRangeF0Hz (v5)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	90405	90405	111.281	<2e-16 ***
position	2	2871	1435	1.767	0.175
gender:position	2	3235	1618	1.991	0.141
Residuals	128	103987	812		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
Dependent variable: 'SyllableRangeF0ST (v6)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	6.0	5.989	1.879	0.173
position	2	8.9	4.460	1.399	0.251
gender:position	2	4.9	2.467	0.774	0.463
Residuals	128	408.1	3.188		

Table 25: excursion size anova results for /ritaal/ under broad focus

For this word, there is a significance according to gender, however no effect of position is reported. According to pairwise comparisons, the differences between the two words is significant for both (v5) and (v6) (ritaal-laana: diff=72.00746, p=0.00000<0.001).

3. Intensity

The following are the intensity results for the targets under broad focus:

Summary statistics

Laana	Initial		Medial		Final	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
v7	71.2	4.27	67.5	3.66	67.8	5.71
Gender	**F 69.5**	**M 73**	**F 66.1**	**M 69**	**F 64.3**	**M 70**
v7: IntensitymaxdB						

Table 26: intensity summary statistics for /laana/ under broad focus

Anova results

Dependent variable: 'IntensitymaxdB (v7)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	26936	26936	25.057	1.80e-06 ***
position	2	34490	17245	16.042	6.07e-07 ***
gender:position	2	1474	737	0.686	0.506
Residuals	128	137597	1075		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

There is a significant difference in intensity means according to position and gender for target word /laana/ under broad focus. These are compared to the secondary target word in the following:

Ritaal	Initial		Medial		Final	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
v7	70.4	5.12	70.9	4.47	71.7	5.54
Gender	**F 67.4**	**M 73.3**	**F 68.6**	**M 73.3**	**F 68.2**	**M 74**
v7: IntensitymaxdB						

Table 27: intensity summary statistics for /ritaal/ under broad focus

Anova results

Dependent variable: 'IntensitymaxdB (v7)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	1011.7	1011.7	55.266	1.33e-11 ***
position	2	14.3	7.1	0.389	0.678
gender:position	2	8.6	4.3	0.235	0.791
Residuals	128	2343.1	18.3		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 28: intensity anova results for /ritaal/ under broad focus

For /ritaal/, significant differences are reported according to gender but not according to position. According to pairwise comparisons, the difference between the two words is significant for this intensity variable (ritaal-laana: diff=1.840675, p=0.00).

4. Pitch accent movement variables

The following presents results for the rise time and speed in L+ H* tones under broad focus. These too will be compared to their counterparts in narrow focus in the upcoming section.

Summary statistics

Laana	Initial		Medial		Final	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
v8	112.6	37.3	118	54.1	42	20.3
Gender	**F 138.8**	**M 85.4**	**F 117.3**	**M 118.7**	**F 30.64**	**M 49.39**

v9	34.52	11.1	33.17	6.74	30.35	10.8
Gender	**F 33.2**	**M 35.8**	**F 32.7**	**M 33.6**	**F 27.6**	**M 32.1**
v8: SyllableRiseTimeMS, v9: RiseSpeedSTsec						

Table 29: pitch movement summary statistics for /laana/ under broad focus

Anova results

					Dependent variable: 'SyllableRiseTimeMS (v8)'
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	5158	5158	5.263	0.0234 *
position	2	68054	34027	34.717	9e-13 ***
gender:position	2	1831	915	0.934	0.3957
Residuals	128	125455	980		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
					Dependent variable: 'RiseSpeedSTsec (v9)'
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	1189	1189.1	0.771	0.382
position	2	1179	589.7	0.382	0.683
gender:position	2	697	348.3	0.226	0.798
Residuals	128	197432	1542.4		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 30: pitch movement anova results for /laana/ under broad focus

The differences in means according to gender and position are significant for rise time (v8), but not for rise speed for the target word.

Ritaal	Initial		Medial		Final	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
v8	113	72.9	111	82.7	131	85.6
Gender	**F 129.4**	**M 96**	**F 112.7**	**M 108.7**	**F 123.3**	**M 135.3**
v9	69.0	40.0	92	115	67.2	114
Gender	**F 66**	**M 72.1**	**F 116.2**	**M 65.1**	**F 93.8**	**M 49.8**
v8: SyllableRiseTimeMS, v9: RiseSpeedSTsec						

Table 31: pitch movement summary statistics for /ritaal/ under broad focus

Anova results

					Dependent variable: 'SyllableRiseTimeMS (v8)'
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	36	35.6	0.023	0.879
position	2	3415	1707.4	1.113	0.332
gender:position	2	687	343.7	0.224	0.800
Residuals	128	196359	1534.1		
					Dependent variable: 'RiseSpeedSTsec (v9)'
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	102	102	0.069	0.7933
position	2	8387	4193	2.828	0.0628 .
gender:position	2	2194	1097	0.740	0.4793
Residuals	128	189815	1483		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 32: pitch movement anova results for /ritaal/ under broad focus

For /ritaal/ no significance is reported for any of the variables according to gender, nor position. According to pairwise comparisons, the difference between the two words is not significant for (v8) (ritaal-laana: diff=5.809898, p=0.4). However, a significant difference between the two words is reported for (v9) (ritaal-laana: diff=1.263, p<0.05).

5.4.2.3 Conclusion

To conclude this section we have seen many qualitative and quantitative patterns that mark broad focus utterances. We have seen that under broad focus, the target can be different, but not necessarily greater in peak height, excursion size, or intensity from the secondary target word. It was also shown that neither accent type nor phrasing was used as a qualitative measure to mark this condition. This realisation is expected since /ˈlaa.na/ is not the intended answer to the contrastive focus question, but rather the whole sentences is a response to a more general question. Broad focus utterances showed significant within-utterance relationships regarding the quantitative details, which mirror the relevant accent distribution hierarchy and its phonetic implementation. Under this focus type, a nuclear accent is free to be placed on any constituent, and accent distribution and phonetic detail then take place accordingly.

5.4.2.4 Narrow Focus intonational patterns

Examination of the narrow focus utterances revealed a noticeable consistent behavior across speakers. In initial position and in a single intonational phrase [Figure 54], a nuclear accent H* or L+H* is placed on the target /ˈlaa.na/ which is followed by a monotonous stretch of low pitch until the end of the intonational phrase boundary L%. In many studies this behaviour is interpreted as evidence of deaccentuation or post focal compression (Chahal and Hellmuth, 2014 among others). Another method used by some speakers was to occasionally break the utterance and place the target in its own intermediate/ intonational phrase where the target bears a prominent nuclear accent- with or without an accompanying pause. This is confirmatory of the conclusion in the previous chapter 4 that in JA post-nuclear accents are prohibited in the same phrase. When a phrase/boundary break was inserted, speakers varied in regards to usage of pauses following this phrase/boundary tone, with the most common pattern being the insertion of a pause after the phrase. Phrase tone H-, and boundary tone H% are used to mark this phrase boundary, demonstrating a similarity to Lebanese (Chahal, 2001, 2014) and Spanish for example (Face, 2002), where mainly high boundaries were used to mark a contrastive focus word. The remainder of the string is realised as a separate phrase with its own prominence distribution. In the following table is a count report of the use of a phrase boundary to mark focus as seen in the current corpus. The raw count data reveals that there is both within and between- speaker variation regarding whether or not they use phrasing to mark narrow focus in the dialect. In the data, only four speakers (3 male, 1 female) showed no within- speaker variation. They were either consistent in not using phrasing to mark their NF utterances thus producing them in one IP, or they were consistent in producing all of their NF repetitions using phrasing to mark narrow focus targets:

Speaker	NF count including repetition	NF using phrasing	NF not using phrasing -realised as one IP
01_M	9	6	3
01_F	7	4	3
02_F	9	1	8
02_M	9	0	9
03_F	7	6	1
04_F	9	3	6
05_F	8	6	2
03_M	7	7	0
04_M	8	5	3
05_M	9	9	0
06_F	9	3	6
06_M	9	3	6
07_M	9	4	5
08_M	9	4	5
09_M	9	1	8
07_F	9	2	7
08_F	9	0	9
09_F	9	7	2
10_F	9	8	1
10_M	9	4	5
Total	172	83 (48.25%)	89 (51.74%)

Table 33: speaker variation in using phrasing to signal narrow focus

As can be seen in figure [55], when the breaking strategy is used, F0 rises at the beginning of the target stressed syllable, and this high F0 is also seen at the edge of the word. This high pitch level at the word edge is indicative of a high edge tone. There is also an observable long pause and lengthening of the last vowel immediately preceding the boundary. It can also be noted that the peaks of accented and nuclear accented words in second phrase are somewhat lower than the peak in the target word, a case that might indicate a long- distance association to the previous phrase in the form of gradient pitch manipulation.

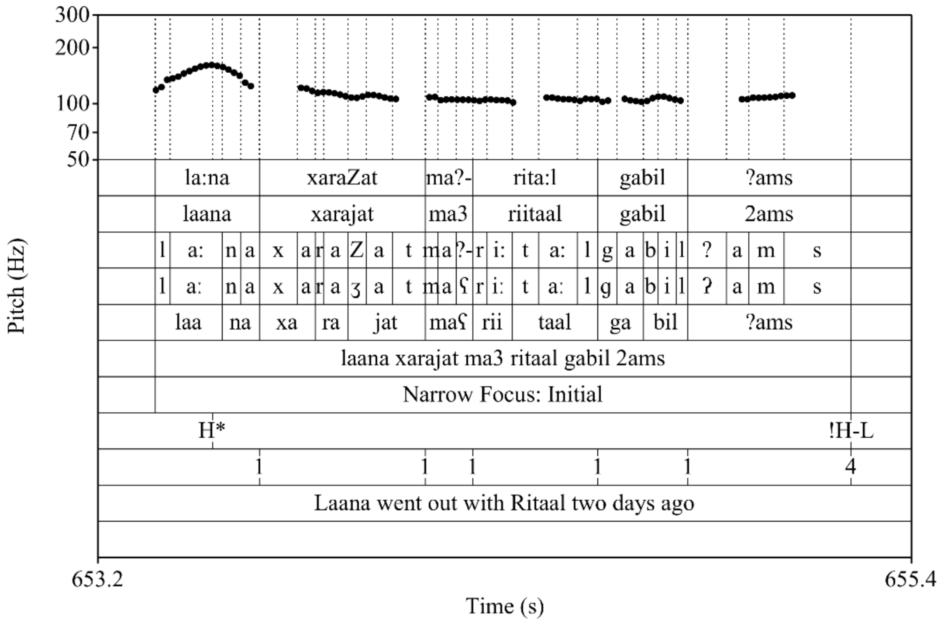


Figure 57: Narrow focus utterance with target in initial position. The utterance was realised as one IP. PFC is evident in the post- focal string. Speaker 10_M

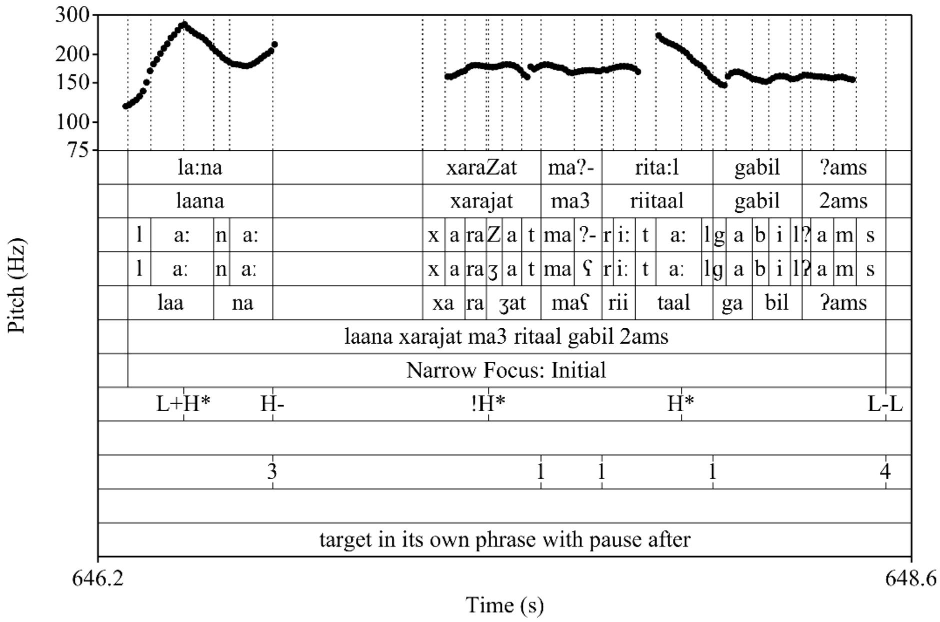


Figure 58: Narrow focus utterance with target in initial position. The utterance is realised as two prosodic phrases. Speaker 01_M

In medial position, again two patterns emerge. The most common pattern is where the target is nuclear accented and the preceding string shows evidence of pre-nuclear pitch accented words and post- focal deaccentuation, as can be seen in figure [56]. Another pattern is to divide the utterance into smaller prosodic phrases with the target in its own phrase. In this pattern the utterance would be realised as two phrases, with an edge tone occurring after the nuclear accented target, similar to figure [55] above.

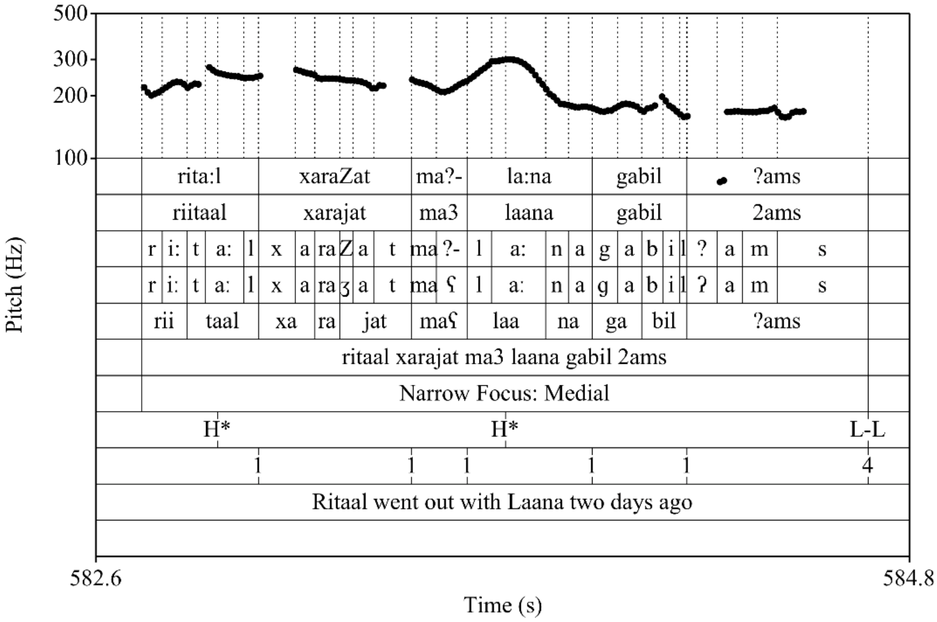


Figure 59: Narrow focus utterance with target in medial position. This utterance is realised as one IP. pre- nuclear accent can be noticed. Speaker 09_F

In final position, the utterance is mainly realised as a single intonational phrase where the target is nuclear accented and the preceding words bear prenuclear pitch accents [Figure 57]. The peaks of the pre- nuclear accents are lower than target. Another less common pattern is for the utterance to be broken into two prosodic phrases, where the adverbial phrase /gabil ?ams/ heads a separate phrase followed by a L- or H- phrase tone. This pattern was observed in less than 15% of the data.

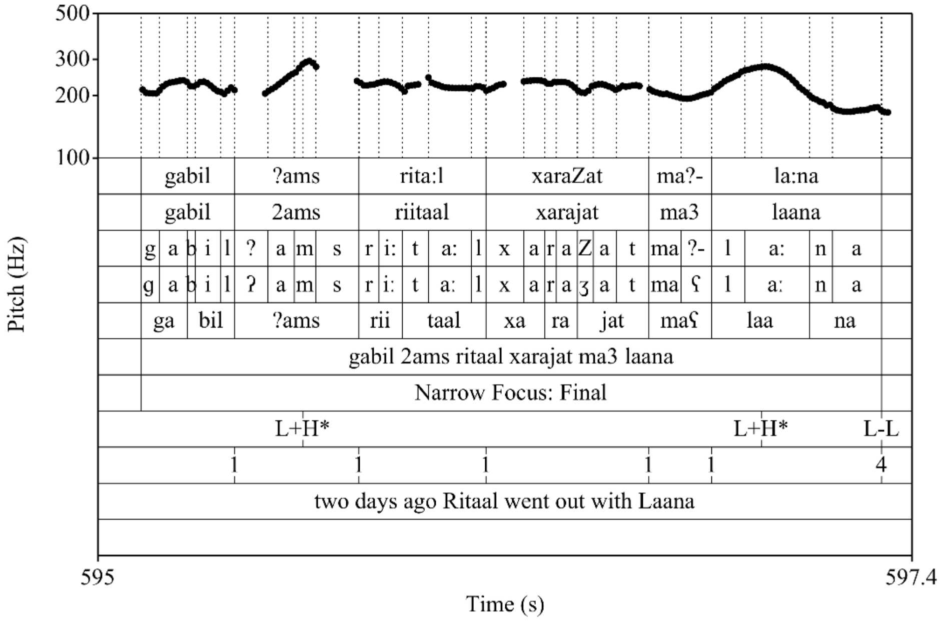


Figure 60: Narrow focus utterance with target in final position. This utterance is realised as one IP. A pre- nuclear accent can be noticed on /?ams/. speaker 03_F

5.4.2.5 Quantitative detail

Narrow focus utterances, as mentioned previously, are sometimes produced in one whole Intonational Phrase, and were additionally broken down into separate prosodic phrases in other cases. The relationships between the durations, intensity, peaks and excursions are reported accordingly.

1. <u>Duration</u>

The following are the results for when the targets are realised in one IP.

<u>Summary statistics</u>

Laana	Initial		Medial		Final	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
v1	170	35.0	159	39.5	175	36.2
Gender	**F 181.8**	**M 157.8**	**F 159.6**	**M 158.1**	**F 187.9**	**M 161.7**
v2	246	56.1	224	51.9	232	46.3
Gender	**F 261.5**	**M 230**	**F 225.7**	**M 222.3**	**F 252.8**	**210.5**
v1: Phonemedurationms, v2: Syllabledurationms						

Table 34: duration summary statistics for /laana/ under narrow focus in one IP

<u>Anova results</u>

Dependent variable: 'PhonemedurationMS (v1)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	3827	3827	6.684	0.0115 *
position	2	3400	1700	2.969	0.0570 .

	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender:position	2	1258	629	1.099	0.3382
Residuals	81	46379	573		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
Dependent variable: 'SyllabledurationMS (v2)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	15361	15361	6.310	0.014 *
position	2	6069	3035	1.246	0.293
gender:position	2	6159	3080	1.265	0.288
Residuals	81	197191	2434		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 35: duration anova results for /laana/ under narrow focus in one IP

The differences according to gender are reported to be significant. However, no significant differences are reported according to word position. For the secondary target word in the same phrase, please note in the following a similar effect of gender for vowel duration (v1) only:

Summary statistics

Ritaal	Initial		Medial		Final	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
v1	111	21.1	119	18.3	118	37.2
Gender	**F 106.3**	**M 114.7**	**F 110.8**	**M 126.4**	**F 112.5**	**M 124.1**
v2	230	33.9	218	36.3	221	50.0
Gender	**F 236**	**M 224.4**	**F 210.7**	**M 225.7**	**F 220**	**221.5**
v1: Phonemedurationms, v2: Syllabledurationms						

Table 36: duration summary sttaistics for /ritaal/ under narrow focus in one IP

Anova results

Dependent variable: 'PhonemedurationMS (v1)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	4223	4223	6.752	0.0111 *
position	2	1602	801	1.281	0.2832
gender:position	2	995	497	0.795	0.4549
Residuals	83	51909	625		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
Dependent variable: 'SyllabledurationMS (v2)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	138	138.5	0.210	0.648
position	2	1730	865.1	1.311	0.275
gender:position	2	2120	1060.1	1.607	0.207
Residuals	83	54749	659.6		

Table 37: duration anova results for /ritaal/ under narrow focus in one IP

As for the comparison between the two words in this case, according to pairwise comparisons, there is a significant difference between words (ritaal-laana: diff=18.543, $p<0.05$) for (v1). For (v2), there is not a significant difference according to word ($p=0.11$).

When the utterance was broken into two prosodic phrases, the target /ˈlaa.na/ formed the head of its respective phrase and so was nuclear accented. The results for this case are as follows:

Summary statistics

Laana	Initial		Medial		Final	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
v1	226	52.6	210	52.1	172	18.2
Gender	**F 244.1**	**M 209.8**	**F 220**	**M 199.7**	**F 171.7**	**M 171.5**
v2	311	60.5	274	58.7	240	27.2
Gender	**F 337.6**	**M 287**	**F 291**	**M 257.6**	**F 250.6**	**M 234.7**
v1: Phonemedurationms, v2: Syllabledurationms						

Table 38: duration summary results for /laana/ under narrow focus in a separate prosodic phrase

Anova results

Dependent variable: 'PhonemedurationMS (v1)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	4262	4262	11.989	0.000917 ***
position	2	6794	3397	9.555	0.000214 ***
gender:position	2	626	313	0.881	0.419114
Residuals	70	24887	356		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
Dependent variable: 'SyllabledurationMS (v2)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	5331	5331	15.453	0.000196 ***
position	2	6907	3453	10.010	0.000150 ***
gender:position	2	186	93	0.269	0.764841
Residuals	70	24150	345		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 39: duration anova results for /laana/ under narrow focus in a separate phrase

The differences according to both gender and position in this case are reported to be significant as can be inferred from the tables. For the secondary target word, no significant differences are reported according to gender, nor position as can be inferred from the following:

Ritaal	Initial		Medial		Final	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
v1	137	32.6	134	35.4	140	40.4
Gender	**F 142.7**	**M 132.2**	**F 139.6**	**M 129**	**F 135.4**	**M 143**
v2	262	61.2	247	61.3	262	70.0
Gender	**F 284**	**M 244.6**	**F 265**	**M 230.6**	**F 262.3**	**M 261.1**
v1: Phonemedurationms, v2: Syllabledurationms						

Table 40: duration summary statistics for /ritaal/ under narrow focus in a separate phrase

Anova results

				Dependent variable: 'PHONEMEdurationMS (v1)'	
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	373	372.8	0.751	0.389
position	2	125	62.3	0.126	0.882
gender:position	2	383	191.7	0.386	0.681
Residuals	69	34266	496.6		
				Dependent variable: 'SYLLABLEdurationMS (v2)'	
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	1657	1657.2	3.539	0.0641 .
position	2	930	465.2	0.993	0.3755
gender:position	2	255	127.3	0.272	0.7627
Residuals	69	32307	468.2		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 41: duration anova results for /ritaal/ under narrow focus in a separate phrase

As for the difference between the two words in this case, pairwise comparisons regarding the two words is significant (ritaal-laana: diff=-60.21456, p=0.00000<0.001) for both (v1) and (v2).

2. Pitch variables

The following are the results for when the targets are realised in one IP.

F0 Peak height

Summary statistics

Laana	Initial		Medial		Final	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
v3	223	59.0	211	57.6	207	49.8
Gender	**F 276.3**	**M 169**	**F 265.4**	**M 160.2**	**F 247.2**	**M 166**
v4	13.3	4.64	12.3	4.71	12.0	4.34
Gender	**F 17.5**	**M 9**	**F 16.81**	**M 8.15**	**F 15.5**	**M 8.5**
v3: phonememaxF0, v4: maxF0ST						

Table 42: peak height summary statistics for /laana/ under narrow focus in one IP

Anova results

				Dependent variable: 'phonememaxF0 (v3)'	
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	202138	202138	329.631	<2e-16 ***
position	2	3650	1825	2.976	0.0566 .
gender:position	2	3204	1602	2.612	0.0795 .
Residuals	81	49671	613		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
				Dependent variable: 'maxF0ST (v4)'	
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	1389.7	1389.7	334.258	<2e-16 ***
position	2	22.6	11.3	2.719	0.0719 .
gender:position	2	12.4	6.2	1.494	0.2306
Residuals	81	336.8	4.2		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 43: peak height anova results for /laana/ under narrow focus in one IP

The peak height means for the target in this case report significant differences according to gender but not position.

Summary statistics

Ritaal	Initial		Medial		Final	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
v3	180	66.4	204	69.6	212	64.6
Gender	**F 240.7**	**M 120**	**F 272.2**	**M 140.5**	**F 272**	**M 151.7**
v4	8.98	6.81	11.3	6.09	12.2	5.64
Gender	**F 15**	**M 2.7**	**F 17.2**	**M 5.79**	**F 17.3**	**M 7**
v3: phonememaxF0, v4: maxF0ST						

Table 44: peak height summary statistics for /ritaal/ under narrow focus in one IP

Anova results

					Dependent variable: 'phonememaxF0 (v3)'
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	43966	43966	382.317	< 2e-16 ***
position	2	5103	2551	22.186	1.91e-08 ***
gender:position	2	126	63	0.549	0.58
Residuals	83	9545	115		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
					Dependent variable: 'maxF0ST (v4)'
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	43966	43966	382.317	< 2e-16 ***
position	2	5103	2551	22.186	1.91e-08 ***
gender:position	2	126	63	0.549	0.58
Residuals	83	9545	115		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 45: peak height anova results for /ritaal/ under narrow focus in one IP

For /ritaal/ mean differences are reported to be significant according to both gender and position. As for the difference between the two words, according to pairwise comparisons, the difference is not significant for both (v3) and (v4) (ritaal-laana: diff=-11.99018, p=0.11).

When the utterance was broken down into smaller prosodic phrases, /ˈlaa.na/ was placed in a separate phrase as the case in initial position. The results for peak in this case are reported in the following:

Summary statistics

Laana	Initial		Medial		Final	
	mean	std. deviation	mean	std. deviation	mean	std. deviation

v3	221	51.5	207	47.2	186	47.8
Gender	**F 255.5**	**M 190.7**	**F 243.6**	**M 173.6**	**F 233.2**	**M 162.2**
v4	13.2	4.28	12.2	4.08	10.2	4.42
Gender	**F 16**	**M 10.7**	**F 15.2**	**M 9.2**	**F 14.5**	**M 8**
v3: phonememaxF0, v4: maxF0ST						

Table 46: peak height summary statistics for /laana/ under narrow focus in a separate prosodic phrase

Anova results

Dependent variable: 'phonememaxF0 (v3)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	90785	90785	70.059	3.8e-12 ***
position	2	7594	3797	2.930	0.060 .
gender:position	2	138	69	0.053	0.948
Residuals	70	90709	1296		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
Dependent variable: 'maxF0ST (v4)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	668.3	668.3	69.053	4.91e-12 ***
position	2	54.8	27.4	2.831	0.0657 .
gender:position	2	3.3	1.6	0.168	0.8454
Residuals	70	677.5	9.7		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 47: peak height anova results for /laana/ under narrow focus in a separate phrase

In this case, significant differences are reported according to gender only. In the following, results for the secondary target word are reported.

Summary statistics

Ritaal	Initial		Medial		Final	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
v3	213	53.0	219	54.9	189	50.0
Gender	**F 249**	**M 182.6**	**F 264**	**M 176**	**F 234**	**M 166**
v4	12.5	4.53	13.0	4.50	10.4	4.55
Gender	**F 15.6**	**M 10**	**F 16.7**	**M 9.5**	**F 14.5**	**M 8.3**
v3: phonememaxF0, v4: maxF0ST						

Table 48: peak height summary statistics for /ritaal/ under narrow focus in a separate phrase

Anova results

Dependent variable: 'phonememaxF0 (v3)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	17213	17213	70.492	3.77e-12 ***
position	2	790	395	1.617	0.206
gender:position	2	299	149	0.612	0.545
Residuals	69	16849	244		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
Dependent variable: 'maxF0ST (v4)'					

	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	17213	17213	70.492	3.77e-12 ***
position	2	790	395	1.617	0.206
gender:position	2	299	149	0.612	0.545
Residuals	69	16849	244		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 49: peak height anova results for /ritaal/ under narrow focus in a separate phrase

As can be seen, differences according to gender are reported for this /ritaal/ in the case where the utterance is broken down to separate prosodic phrases. As for the difference between the two words, according pairwise comparisons, the difference is not significant for both (v3) and (v4) (ritaal-laana: diff=-20.991, p=0.8).

Excursion size

In the following are the syllable excursion sizes for the targets under narrow focus when they are in the same IP.

Summary statistics

Laana	Initial			Medial			Final		
		mean	std. deviation		mean	std. deviation		mean	std. deviation
v5		65.0	24.0		48.7	20.6		34.9	18.0
Gender	**F**	**72.6**	**M 57.4**	**F**	**50.1**	**M 47.3**	**F**	**39.2**	**M 30.4**
v6		6.26	2.26		4.92	2.65		3.21	1.61
Gender	**F**	**5.3**	**M 7.2**	**F**	**3.5**	**M 6.1**	**F**	**3**	**M 3.4**
v5: SyllableRangeF0Hz, v6: SyllableRangeF0ST									

Table 50: excursion size summary statistics for /laana/ under narrow focus in one IP

Anova results

Dependent variable: 'SyllableRangeF0Hz (v5)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	1821	1821	4.108	0.046 *
position	2	16794	8397	18.942	1.78e-07 ***
gender:position	2	345	172	0.389	0.679
Residuals	81	35908	443		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
Dependent variable: 'SyllableRangeF0ST (v6)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	4797	4797	13.534	0.00042 ***
position	2	19428	9714	27.405	8.13e-10 ***
gender:position	2	1932	966	2.725	0.07153 .
Residuals	81	28711	354		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 51: excursion size anova results for /laana/ under narrow focus in one IP

Anova results indicate that the difference in excursion size according to sentential position are indeed significant, as well as differences according to gender. Now observe the results for the secondary target word in the case where it is realised in one IP with the target. A reported significance in the means according to gender and position can be seen in the following:

Summary statistics

Ritaal	Initial		Medial		Final	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
v5	37.8	25.9	50.5	26.2	49.3	21.7
Gender	**F 56**	**M 19.6**	**F 69.5**	**M 32.7**	**F 62.3**	**M 36.3**
v6	3.70	1.85	4.92	1.75	4.63	1.47
Gender	**F 4.5**	**M 2.8**	**F 5.2**	**4.5**	**F 4.5**	**M 4.6**
v5: SyllableRangeF0Hz, v6: SyllableRangeF0ST						

Table 52: excursion size summary statistics for /ritaal/ under narrow focus in one IP

Anova results

Dependent variable: 'SyllableRangeF0Hz (v5)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	26340	26340	76.711	1.99e-13 ***
position	2	3332	1666	4.852	0.0102 *
gender:position	2	569	284	0.828	0.4405
Residuals	83	28499	343		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
Dependent variable: 'SyllableRangeF0ST (v6)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	2665	2665	4.517	0.0365 *
position	2	5143	2572	4.359	0.0159 *
gender:position	2	1962	981	1.663	0.1959
Residuals	83	48969	590		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 53: excursion size anova results for /ritaal/ under narrow focus in one IP

As for the difference between the targets in this case, according to pairwise comparisons, the difference between the two words is not significant for neither (v5) nor (v6) (ritaal-laana: diff=-0.1932, p=0.2).

The following are results for targets when they are realised in separate prosodic phrases.

Summary statistics

Laana	Initial		Medial		Final	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
v5	63.8	36.5	41.4	19.0	36.0	18.3
Gender	**F 63**	**M 64.5**	**F 52.4**	**M 31**	**F 51.6**	**M 28.1**
v6	6.04	3.41	3.82	1.36	3.75	1.76
Gender	**F 5**	**M 7**	**F 4.2**	**M 3.4**	**F 4.5**	**M 3.3**
v5: SyllableRangeF0Hz, v6: SyllableRangeF0ST						

Table 54: excursion size summary statistics for /laana/ under narrow focus in a separate phrase

Anova results

				Dependent variable: 'SyllableRangeF0Hz (v5)'	
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	4445	4445	12.363	0.000773 ***
position	2	4813	2406	6.693	0.002189 **
gender:position	2	2148	1074	2.987	0.056887 .
Residuals	70	25169	360		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
				Dependent variable: 'SyllableRangeF0ST (v6)'	
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	179	179.0	0.423	0.51739
position	2	4483	2241.3	5.300	0.00719 **
gender:position	2	2312	1155.9	2.733	0.07194 .
Residuals	70	29602	422.9		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 55: excursion size anova results for /laana/ under narrow focus in a separate phrase

Again a significant difference is reported according to both gender and position for (v5), but only a significance according to position for (v6). For the secondary target word, only differences according to gender are significant for (v5). Note the following:

Summary statistics

Ritaal	Initial			Medial			Final	
		mean	std. deviation		mean	std. deviation	mean	std. deviation
v5		64.3	28.3		63.6	25.6	54.7	27.0
Gender	**F**	**71**	**M 58.6**	**F**	**80.2**	**M 48.1**	**F 69.4**	**M 47.3**
v6		6.11	2.09		5.87	1.94	5.66	2.06
Gender	**F**	**5.8**	**M 6.3**	**F**	**6.4**	**M 5.3**	**F 6**	**M 5.4**
v5: SyllableRangeF0Hz, v6: SyllableRangeF0ST								

Table 56: excursion size summary statistics for /ritaal/ under narrow focus in a separate phrase

Anova results

				Dependent variable: 'SyllableRangeF0Hz (v5)'	
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	9346	9346	15.254	0.000216 ***
position	2	498	249	0.406	0.667636
gender:position	2	1440	720	1.175	0.314961
Residuals	69	42275	613		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
				Dependent variable: 'SyllableRangeF0ST (v6)'	
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	2.12	2.120	0.516	0.475
position	2	1.99	0.994	0.242	0.786
gender:position	2	9.84	4.922	1.197	0.308
Residuals	69	283.78	4.113		

Table 57: excursion size anova results for /ritaal/ under narrow focus in a separate phrase

As for the difference between the two words, according to pairwise comparisons, it is significant for both (v5) and (v6) (ritaal-laana: diff=-21.57, p<0.05).

3. **Intensity**

The following are the results for when the targets are realised in one IP.

Summary statistics

Laana	Initial		Medial		Final	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
v7	72.6	3.85	69.8	3.93	70.6	5.08
Gender	**F 70.4**	**M 74.7**	**F 68.3**	**M 71.1**	**F 68.6**	**M 72.5**
v7: IntensitymaxdB						

Table 58: intensity summary statistics for /laana/ under narrow focus in one IP

Anova results

Dependent variable: 'IntensitymaxdB (v7)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	8189	8189	15.737	0.000156 ***
position	2	4251	2125	4.085	0.020412 *
gender:position	2	280	140	0.269	0.764650
Residuals	81	42148	520		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 59: intensity anova results for /laana/ under narrow focus in one IP

Intensity results show a significant difference according to both gender and position for the target word realised in one whole IP.

Summary statistics

Ritaal	Initial		Medial		Final	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
v7	65.7	4.11	70.5	4.35	71.2	4.25
Gender	**F 65.2**	**M 66.1**	**F 69.1**	**M 71.6**	**F 68.5**	**M 73.8**
v7: IntensitymaxdB						

Table 60: intensity summary statistics for /ritaal/ under narrow focus in one IP

Anova results

Dependent variable: 'IntensitymaxdB (v7)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	4432	4432	10.113	0.00207 **
position	2	14051	7025	16.031	1.3e-06 ***
gender:position	2	3885	1942	4.432	0.01483 *
Residuals	83	36373	438		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 61: intensity anova results for /ritaal/ under narrow focus in one IP

Intensity differences for /ritaal/ are significant according to both gender and position as can be seen in the previous tables. According to pairwise comparisons regarding the two target words, the differnce is significant (ritaal-laana: diff=-16.28619, p<0.05).

The following are intensity results for targets when they are realised in separate prosodic phrases:

Summary statistics

Laana	Initial		Medial		Final	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
v7	74.1	6.58	73.6	5.69	71.7	5.74
Gender	**F 69**	**M 78.5**	**F 70.5**	**M 76.5**	**F 68.7**	**M 73.1**
v7: IntensitymaxdB						

Table 62: intensity summary statistics for /laana/ under narrow focus in a separate prosodic phrase

Anova results

Dependent variable: 'IntensitymaxdB (v7)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	11409	11409	37.113	5.39e-08 ***
position	2	2182	1091	3.549	0.0340 *
gender:position	2	1465	732	2.383	0.0997 .
Residuals	70	21519	307		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 63: intensity anova results for /laana/ under narrow focus in a separate phrase

Intensity differences are significant according to both position and gender even when target form its own prosodic phrase.

Summary statistics

Ritaal	Initial		Medial		Final	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
v7	71.6	7.44	74.7	5.28	72.6	7.02
Gender	**F 66.5**	**M 75.7**	**F 72**	**M 77.2**	**F 69.8**	**M 74**
v7: IntensitymaxdB						

Table 64: intensity summary statistics for /ritaal/ under narrow focus in a separate phrase

Anova results

Dependent variable: 'IntensitymaxdB (v7)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	783.3	783.3	24.379	5.29e-06 ***
position	2	186.2	93.1	2.898	0.0619 .
gender:position	2	90.8	45.4	1.413	0.2504
Residuals	69	2216.9	32.1		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 65: intensity anova results for /ritaal/ under narrow focus in a separate phrase

For this word, differnces in means are statistically significant according to gender only. As for the differences between the two targets in this case, according to pairwise comparisons, it is not significant (ritaal-laana p= 0.8).

4. Pitch accent movement variables

The following presents results for the rise time and speed under narrow focus. Rise duration and rise speed are measures reported to signal different focus conditions. These are typically measured in rising tones L+ H* to assess how quickly or slowly the transition takes between the tones, in addition to the time. These too will be compared to their secondary target /ritaal/, and broken down according to whether they are realised in the same IP or in a separate prosodic phrase. First these are the results for when both are in the same IP:

Summary statistics

Laana	Initial			Medial			Final		
		mean	std. deviation		mean	std. deviation		mean	std. deviation
v8		153	60.1		118	53.8		86.5	34.0
Gender	**F**	**159**	**M 146.2**	**F**	**108.1**	**M 127.1**	**F**	**85.4**	**M 87.6**
v9		44.3	17.1		46.4	22.3		38.5	14.1
Gender	**F**	**37.2**	**M 51.2**	**F**	**39.5**	**M 52.7**	**F**	**37.2**	**M 39.7**
v8: SyllableRiseTimeMS, v9: RiseSpeedSTsec									

Table 66: pitch movement summary statistics for /laana/ under narrow focus in one IP

Anova results

Dependent variable: 'SyllableRiseTimeMS (v8)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	1177	1177	2.409	0.125
position	2	13273	6637	13.583	8.18e-06 ***
gender:position	2	841	420	0.860	0.427
Residuals	81	39577	489		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
Dependent variable: 'RiseSpeedSTsec (v9)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	3427	3427	5.632	0.020 *
position	2	1317	658	1.082	0.344
gender:position	2	835	417	0.686	0.507
Residuals	81	49289	609		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 67: pitch movement anova results for /laana/ under narrow focus in one IP

As can be seen, a significant result is reported for (v8) according to position, while an effect of gender is reported for (v9) when the target is realised in one IP. For the secondary target, no significant results are reported according to position, nor gender. Note the following:

Summary statistics

Ritaal	Initial		Medial		Final	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
v8	88.1	36.1	92.1	57.3	94.4	67.4
Gender	**F 98.2**	**M 77.8**	**F 78.5**	**M 104.6**	**F 75.3**	**M 113.4**
v9	58.6	65.5	82.8	108	78.0	76.9
Gender	**F 50.6**	**M 66.6**	**F 115.8**	**M 52.1**	**F 98.8**	**M 59**
v8: SyllableRiseTimeMS, v9: RiseSpeedSTsec						

Table 68: pitch movement summary statistics for /ritaal/ under narrow focus in one IP

Anova results

Dependent variable: 'SyllableRiseTimeMS (v8)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	1798	1798	2.806	0.0977 .
position	2	781	390.5	0.609	0.5461
gender:position	2	2968	1483.9	2.316	0.1050
Residuals	83	53186	640.8		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
Dependent variable: 'RiseSpeedSTsec (v9)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	2504	2503.5	3.891	0.0519 .
position	2	2323	1161.4	1.805	0.1709
gender:position	2	506	253.1	0.393	0.6761
Residuals	83	53406	643.4		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 69: pitch movement anova results for /ritaal/ under narrow focus in one IP

As for the comparison between the two words, according to pairwise comparisons the difference is significant for both (v8) and (v9) (ritaal-laana: diff=-23.19618, p=0.00).

Finally, the following are results for targets when they are realised in separate prosodic phrases.

Summary statistics

Laana	Initial		Medial		Final	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
v8	191	89.8	171	106	112	43.5
Gender	**F 222.4**	**M 163.8**	**F 208.1**	**M 136**	**F 140**	**M 97.5**
v9	37.6	27.3	30.3	14.9	35.5	16.8
Gender	**F 23.6**	**M 50**	**F 25.3**	**M 35**	**F 33**	**M 36.6**
v8: SyllableRiseTimeMS, v9: RiseSpeedSTsec						

Table 70: pitch movement summary statistics for /laana/ under narrow focus in a separate prosodic phrase

Anova results

Dependent variable: 'SyllableRiseTimeMS (v8)'					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	82691	82691	11.190	0.00132 **

	DF	Sum Sq	Mean Sq	F value	Pr (>F)
position	2	51499	25749	3.484	0.03609 *
gender:position	2	2086	1043	0.141	0.86864
Residuals	70	517282	7390		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
					Dependent variable: 'RiseSpeedSTsec (v9)'
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	6632	6632	16.554	0.000122 ***
position	2	73	37	0.092	0.912603
gender:position	2	1826	913	2.279	0.109913
Residuals	70	28044	401		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 71: pitch movement anova results for /laana/ under narrow focus in a separate phrase

When the target is in its own prosodic phrase, the differences according to both gender and position are reported for (v8), while only differences according to gender are reported for (v9). For the secondary target word, note that those variables do not report statistical significance as a result of gender, nor position:

Summary statistics

Ritaal	Initial			Medial			Final		
		mean	std. deviation		mean	std. deviation		mean	std. deviation
v8		130	94.7		101	77.0		161	111
Gender	**F**	**160.1**	**M 104.7**	**F**	**126**	**M 76.8**	**F**	**182**	**M 150.4**
v9		80.8	106		83.6	63.6		58.8	43.8
Gender	**F**	**83.5**	**M 78.5**	**F**	**88.4**	**M 79**	**F**	**60.2**	**M 58**
v8: SyllableRiseTimeMS, v9: RiseSpeedSTsec									

Table 72: pitch movement summary statistics for /ritaal/ under narrow focus in a separate phrase

Anova results

					Dependent variable: 'SyllableRiseTimeMS (v8)'
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	1258	1258.5	2.714	0.104
position	2	1868	933.9	2.014	0.141
gender:position	2	20	10.0	0.022	0.979
Residuals	69	32000			
					Dependent variable: 'RiseSpeedSTsec (v9)'
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
gender	1	929	928.6	1.999	0.162
position	2	1365	682.5	1.470	0.237
gender:position	2	811	405.3	0.873	0.422
Residuals	69	32046	464.4		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 73: pitch movement anova results for /ritaal/ under narrow focus in a separate phrase

As for the comparison between the two words, according to pairwise comparisons the difference is significant for both (v8) and (v9) (ritaal-laana: diff=-24.47789, p<0.05).

5.4.2.6 Conclusion

As can be inferred from the previous results, a number of qualitative and quantitative changes take place under narrow focus. The changes affect the duration, intensity, rise size, rise time and speed, and peak height. The results also differed according to the target position, gender and according to whether or not the target heads its own prosodic phrase.

When the target is in the same prosodic phrase as the rest of the string, compared to the string, the target has the longest duration and rise time, the highest peak and intensity, and the fastest rise speed in L+H* tones. An effect of position also shows some interesting patterns regarding the target itself, as well as its relationship to the pre/post focal word /riˈtaal/. The target /ˈlaa.na/ in initial position has a higher peak, larger excursion size and intensity peak, a longer rise time and a faster rise speed, compared to the same word in medial and final positions. As for the secondary target word /riˈtaal/, when /ˈlaa.na/ is in initial and medial positions, this word shows a suppressed peak, smaller excursion and lower intensity. On the other hand, when /ˈlaa.na/ is in final position, /riˈtaal/ shows a higher peak, and a somewhat different excursion size.

When the target is in a separate prosodic phrase from the rest of the string, the target duration is longer in initial position followed by medial, then final. The peak height differences between the target and /ritaal/ are non-significant in all positions. As for excursion size, this quantitative measure is different in /riˈtaal/ in all positions, compared to the target /ˈlaa.na/ in a separate phrase. The intensity of /ˈlaa.na/ in initial position is higher, but when it is in medial or final position, the intensity of /riˈtaal/ is higher. The rise time and speed in L+ H* tones in this condition for /laana/ is compared to the behaviour of its counterpart in the one intonational phrase condition. It was found that when the tone is in a separate phrase, its rise time is longer and the transition is slower than when it is in the same prosodic phrase as the rest of the utterance.

The results of the narrow focus patterns show that there are indeed changes in the signal as a response to the intended focus. The phonological relationships between the words in a string are implemented phonetically by changes in durations, peak heights, excursion sizes, intensity, and tone transitions. These changes take place accordingly depending on the hierarchical structure within each prosodic phrase. Essentially, the changes can be said to mirror the accent distribution mechanism that operates in the post-lexical constituents.

5.5 Chapter discussion and conclusion

The controlled experiment in this chapter shows that intonation is indeed a procedure to mark information structure of the sentences in this dialect, i.e. focus can be encoded intonationally in Jeddah Arabic. The experiment reports intonational patterns as well as intricate phonetic details to mark focus in the dialect. Similar to Lebanese Arabic (Chahal, 2001), Kuwaiti, Moroccan and Yemeni Arabic (Yeou et al., 2007a, b) and Taif Arabic (Al-Zaidi, 2014), JA Arabic demonstrates its own quantitative and qualitative details to mark focus.

Regarding quantitative measures, the dialect systematically employs pitch (Acoustic F0) to differentiate broad and narrow focus utterances. Pitch is used in the form of peak height and excursion size manipulations. There seems to be a systematic relationship between the excursion sizes and the height of peaks in the same prosodic phrase, as well as under different focus conditions. It was found that duration, intensity, rise speed are also phonetic details used to signal the difference between the two focus conditions.

As for qualitative detail, JA demonstrates occasional use of phrasing to mark narrow focus. Phrasing indicates that accent distribution takes place in a systematic manner after each phrase, and that speakers re-arrange the pitch configuration of words after each application to mark a specific focus condition. Accent distribution also shows that in this dialect nuclear pitch accents are obligatory, alongside non-nuclear pitch accents and unaccented syllables. Another qualitative measure used in the dialect is the compression of off-focus words. In an utterance where the target is in the same intonational phrase as the rest of the string, we see a suppression of tonal events pre- and post- focally in the form of lower or absent peaks and small excursions. This means that phonologically the nuclear accent is the head of its prosodic phrase, and the fact that it is the final accent in this phrase suggests that prominence distribution is right-headed in this dialect.

Alongside these conclusions, it is perhaps worth shedding light on another dimension that may contribute to the marking of narrow and broad focus utterances in Jeddah Arabic. Parrallel to the work cited previously by El-Zarka (2011, 2013), further analysis may reveal potential alignment and/or scaling differences that encode different focus interpretations depending on speaker. As reported in her cited studies, Elzarka presents the fact that even though peak alignment was earlier under focus for all the three speakers, the results were highly significant for only one speaker. That is, peak alignment modification was seen as a speaker-dependent strategy to signal focus in her data. Bearing this in mind, a further rigorous investigation is needed in JA to account for the differences in alignment and scaling across identical narrow/broad focal accents, i.e. alignment and or/scaling of H* under narrow and under broad, or L+H* under narrow and under broad focus, that incorporates speaker variation could reveal interesting results. A future investigation would also aim to determine how and whether these intricate details between focus conditions are captured perceptually in this dialect.

Chapter 6. Tonal Alignment

6.1 Introduction

As was discussed in chapter two, one of the main aims of AM theory is to uncover the link between text and tune- the link that gives rise to the different melodies in spoken language. It was also discussed how it is taken in the theory that the association between accents and specific, linguistically relevant positions in the segmental tier is formally observed through surface tonal alignment. Alignment is therefore concerned with the coordination between those tiers, which could indicate the most important landmarks in the segmental tier (the Tone Bearing Units TBU) that are relevant for the expression of intonation in languages. Moreover, it was also mentioned that there might be phonetic factors, as well as prosodic as will be shown shortly, that would cause the alignment of the peak to be re-adjusted in time with respect to the TBU.

Exploring and observing the alignment patterns with regards to the different aforementioned factors in JA is hence one of the main aims in this chapter that will link the theory and the practical evidence. In the current study, the models of tonal alignment will all be tested against the patterns of alignment that JA demonstrates. Their hypotheses, and cross-linguistic implications will also be evaluated, all to observe where JA falls within the phonetic alignment spectrum.

6.2 Experimental Literature

An experimental study in peak alignment patterns following the leading observations brought about by Bruce (1977), is Prieto, van Santen, and Hirschberg (1995) on tonal alignment in Mexican Spanish. The study aimed to investigate whether/how accented syllable segmental makeup and duration, stress 'accent' clash, as well as prosodic boundary (Intonational Phrase: IP, intermediate phrase: ip and phrase-mid: PWd) affect peak alignment in the dialect. Stress clash, occurs where two accented syllables occur in close vicinity to one another without intervening phrasal boundaries. The study thus combines both nuclear and pre-nuclear accents, i.e. a target word occurring before an IP boundary is nuclear accented, whereas pre-nuclear accented in the other two cases. The corpus for their study was 810 declarative sentences produced by two male Mexican Spanish speakers. Target syllables were open syllables of the type CV, varying in vowel identity. Target words vary in stress position (initial- medial- final), and were placed in different structural positions in the sentence. In order to examine stress clash and prosodic boundary effects, the researchers varied the distance (in syllables) between adjacent accents, and between accents and prosodic boundaries. They measured peak location as an absolute value in Milliseconds from the location of the F0 peak to the onset of the accented syllable. Additionally, they investigated the temporal distance between the L valley preceding the peak, and the onset of the syllable in what they call *rise delay* to observe how the valley is syncronised with the accented syllable onset. They also report F0 maximum values in Hertz to observe the difference in height according to position and prosodic boundary- similar to what has been done in the previous focus chapter [5] in this thesis.

The study finds that the onset is the reference point for peaks to align and associate in Mexican Spanish. Moreover, there was found to be a significant positive correlation between peak delay and accented syllable duration, whereby peaks shift rightward as syllable duration increases. As for prosodic boundary effects, the study reports that IP and ip boundaries affect peak location by causing the peak to align earlier in the syllable than in phrase-medial contexts. They found this effect to be consistent regardless of word stress position. They also found that in one of speakers' productions, the peak was retracted to a greater degree before an IP boundary than before an ip boundary, which they explain as a consequence of speech rate. In phrase-mid contexts (Pwd boundary) the peak was found to be located further to the left the more final an accented syllable is in a word. Thus in stress clash, the study reports a re-adjustment in peak location as a result of adjacency of accents, where peaks were relocated as an effect of the number of intervening syllables. However, there was a large variation in resolving tonal clash between the two speakers in their study. One speaker made large re-adjustments to the peak location in stress clash contexts by lengthening and retracting the peak in the target syllable and thus further away from the successive peak, while the second speaker merges the two successive peaks into one accent gesture.

In her study on Lebanese Arabic prosody, Chahal (2001) replicates the previous study in Prieto et al. (1995) and investigates tonal alignment patterns as conditioned by syllable duration, stress clash and different prosodic boundaries. Her targets were words varying in word stress position (initial-medial-final), vowel identity, and consequently in Arabic these have different syllable structures and durations. The syllable types include CV, CVC and CVVC. Her design does not include CVV syllable type, which she does not clearly explain why they have been left out of the analysis. She places the target words and syllables in different positions in the sentence, and in the vicinity of prosodic boundaries of different strengths. The prosodic boundaries are IP, ip and Pwd, and the study therefore combines nuclear and pre-nuclear accents. As for stress clash , she places target accented syllables within a fixed distance from the next accent-if any- separated by a controlled number of intervening unstressed and unaccented syllables. The distance from one accent to the other is either 0: when target word is pre-nuclear and has final stress, while the following nuclear accented syllable has initial stress, 1: when pre-nuclear word is medially stressed and following word is initially stressed, and so on. She also examines the contexts where there are 2, 3 and maximally 4 intervening syllables all by manipulating word stress position. The corpus for her experiment was a total of 2250 utterances obtained from 4 Lebanese Arabic speakers (3 females, 1 male). The researcher measures peak delay as the absolute distance in Ms from the location of the peak to the onset of the accented syllable. She also reports correlation results between syllable duration and peak location relative to the onset.

The results from LA show that there is a positive correlation between peak location and syllable duration indicating that the location of the peak is not static and varies in relation to syllable duration. She finds that- everything else being equal- the longer the syllable, the later (rightward) the peak aligns. The researcher also confirms that in this dialect, according to peak patterns, the onset of the accented syllable is the anchor point to which a peak associates and aligns. Examining the stress clash environment, the researcher reports that overall, the 0 intervening syllables context shows the earliest peak alignment. The degree of which also varies according to syllable duration, all in order to maximise the space between two adjacent accents. The other stress clash contexts (2, 3, 4) display late peak alignment- as late as permitting in the syllable. Finally, as for the prosodic boundary strength effects, regardless of word stress position (initial- medial- final), peaks occurring in the vicinity of an IP, are aligned the earliest in comparison to the other two boundary types. In this case, the peaks are thus located as left as possible within the accented syllable- and away from the edge of the phrase. Following and also displaying early alignment is the ip boundary and finally the Pwd, which shows the least peak retraction. The researcher therefore concludes that in LA, the canonical syllable is indeed the TBU and that prosodic boundary has an effect on peak alignment.

Other studies have also found that, all else being equal, the segmental composition and duration has an impact on the temporal alignment of tones. The effect of syllabic structure on tonal alignment is evident in many studies as we can see. Among them is Yeou's (2004) analysis of the alignment patterns in the Moroccan dialect spoken in Casablanca. This study sets out to investigate in detail how syllabic composition, focus condition (narrow/broad), and position within a sentence can affect tonal alignment. Data for the study was collected from 5 speakers of Moroccan Arabic (3 males, 2 females). The material consisted of words with target syllables of the shape CVVC. A follow up experiment included the syllable types CVC and CV, whose results are compared and reported in this study. The words were placed in three different positions in the sentences (initial- medial- final) and were produced under the two different focus conditions. The researcher measures the distance from the L to the onset of the target syllable, while the H distance was measured from the onset of the syllable and from the end of the target vowel.

The findings of particular relevance to the discussion here, are the tonal alignment patterns found in the dialect. The study shows that in Moroccan Arabic, the valleys are consistently aligned around the onset of the target accented syllables. The alignment of the peak was found to be highly variable and affected by the duration of the syllable. When the vowel is short in open syllables, the H is aligned with the following consonant, but when the vowel is long in an open syllable, the H is realised at the end of the second vowel. Moreover, when a vowel is long in a closed syllable, the H is mainly realised in the middle of the syllable, halfway through the vowel. Thus the study reports a positive correlation between syllable duration and peak alignment, whereby H location varies according to syllable duration. Yeou also concludes that the onset of the syllable is the most reliable anchor point for the H alignment in the dialect.

Replicating the previous study in an attempt to widen the scale and scope of alignment studies in Arabic is Yeou, Embarki, Al-Maqtari, and Dodane's (2007) study on the Moroccan, Kuwaiti and Yemeni dialects. The research carried out a study to investigate tonal alignment patterns in these Arabic dialects. The study examines both peak alignment and valley alignment in relation to specified segmental anchors. The corpus for their study consists of 150 declarative sentences produced by 15 speakers- 5 speakers (3 males, 2 females) from each dialect. Target words had stressed syllables of the shape CVVC finally and CVV medially, focusing mainly on long vowels. These were placed non-finally in a sentence and would thus attract pre-nuclear accents. For peak alignment, they measured absolute distance in Ms from peak location to the end of the stressed vowel/rhyme, for valley alignment, the obtained absolute distance from valley location to the onset of the stressed syllable. Notice the variation in the choice of anchor points used to measure peak alignment in Arabic, here (peak to rhyme) and in Chahal's study (peak to onset). In this thesis both methods are used in an attempt to best predict the anchor point for JA tonal alignment.

The peak alignment patterns showed significant differences in Moroccan according to syllabic structure, but not in Yemeni or Kuwaiti Arabic. In Moroccan, the peak is aligned early in closed syllables but just after the vowel in open ones, as opposed to the other dialects where the peak was always aligned within the vowel in both syllable types. As for the valley, on the other hand, they found that it was consistently aligned at the onset of the accented syllable across syllable types and dialects. This in their view shows that a segmental anchoring analysis of the peak to the rhyme of accented syllable can account for Kuwaiti and Yemeni, but not for the variation shown in Moroccan Arabic. In Moroccan, as was shown in Yeou (2004) previously, the onset of the accented syllable is the decided segmental anchor for peak alignment in the dialect. Unfortunately, this small-scale study does not further include the investigation of the other factors shown to affect tonal alignment (stress clash and prosodic boundary effects) in previous studies.

In a series of instrumental studies on Egyptian Arabic, Hellmuth (2006a, 2006b, 2005, 2004, 2007, 2011) examines tonal alignment behaviour and the implication it brings about for the classification of languages in prosodic typology. Her corpus for the (2006a) thesis included open and closed syllables of the types: CV, CVC, CVV, located in different positions within a word. She measures alignment relative to a number of segmental landmarks. The author obtains absolute temporal distance in Ms from the location of the start of the rise L to both the syllable onset and vowel onset of the stressed syllable. The temporal distance for the end of the rise H is measured between the peak location and: stressed vowel offset, onset of following unstressed vowel, as well as to the onset of the stressed syllable. The corpus for the study consisted of 24 target sentences produced by 15 Cairene Arabic speakers. The study also incorporates a relative 'proportional' measure of peak alignment as a percentage of the stressed syllable duration. The study finds that for this Arabic dialect the L is consistently aligned within the onset of the target syllable across all syllable types. On the other hand, the H was found to be variable in its alignment according to the syllable type. In open syllables the peak aligns outside the short vowel well into the following consonant in CV types, while aligning at the end of the long vowel in CVV syllables. For CVC syllables, the peak aligns just within the coda consonant. The author therefore concludes that the peak in EA aligns in the second mora in a foot, and the rise gesture as a whole is structurally anchored to the stressed foot.

Adding to the Arabic intonational patterns regarding alignment is the Bedouin Hijazi variety spoken in Taif (Al-Zaidi, 2014). In agreement with Egyptian and Lebanese, the researcher shows that in this dialect the two most common pitch accents are L+H* and H*. Al-Zaidi makes a general observation that the L in the bitonal accent is aligned with the onset of the lexically stressed syllable and the H consistently aligns within the syllable as well. The study does not attempt further rigorous testing of this claim as a function of accentual prominence, syllable duration, or stress clash, nor provides additional phonetic detail.

A series of studies by Ladd and colleagues were carried out to best understand the tonal alignment behaviour cross-linguistically. The study by Ladd, Faulkner, Faulkner, and Schepman (1999) set out to further explore the segmental anchoring hypothesis in British English pre-nuclear accents. This was done by observing and comparing peak alignment patterns under different speech rates. The latter goal of their study is outside the scope of the current thesis however their discussion and analysis of alignment behaviour is highly relevant. They investigated the segmental anchors for rising pre-nuclear accents, i.e. the alignment of the beginning (L) and end (H) of the rise with the segmental string. Corpus for their study was comprised of 14 target sentences produced by six RP speakers (4 males, 2 females). Target words varied in vowel identity in the stressed syllable, and in the number of preceding/ following unaccented/unstressed syllables. Target stressed syllables were open syllables of the type CV with sonorant (m, n) as onsets, and were placed in different positions within a word. For their measurement points, they marked the onset of the stressed syllables, the offset of the vowel/rhyme, and the onset of the following unstressed vowel. These were taken as absolute measures (in Ms) and relative measures (proportion of accent in relation to whole syllable duration).

The findings indicate a positive correlation between the duration of the accent and the syllable duration, which they take as evidence for a free duration between the target tones in bitonals, as opposed to the claims in the fixed duration models. Another finding was that regardless of changes in speech rate, the alignment of peaks and valleys was true to the specified anchor points. This was especially seen in the consistent alignment of L at the onset of the accented syllable. Moreover, and although the alignment of H was variable in their data, the peak always aligned at some distance from the accented syllable- usually at the start of the following unstressed syllable. They conclude that the end of the rhyme might not be the true anchor point for H in this dialect as was predicted, but the consonant following the rhyme. They take these findings as support for the SAH: *Segmental Anchoring Hypothesis*.

Another study by Atterer and Ladd (2004) further examines the SAH and alignment in two German dialects (northern and southern). Similar to what has been done in the aforementioned study on British English, the alignment of peaks and valleys of pre-nuclear rising accents in German were investigated. Additionally in this study, they investigate whether the native German patterns in alignment are transferred to the speakers' production of English rising accents, which they find to be the case. To determine the start and end of the German pre-nuclear accents, the corpus for the study consists of 13 target sentences with the target accented syllable preceded and followed by two unstressed syllables. The target syllables contained short vowels and sonorant onsets. This material was produced by 16 German speakers, 8 of each variety controlled for gender in each group. For the measurements, they obtained the absolute temporal distance in Milliseconds between the location of L and the onset of the syllable, as well as to the onset of the stressed vowel. For the alignment of H, the distance was measured between the location of the peak and the onset of the following unstressed vowel.

Their findings show that speakers from both dialects aligned the peak at the unstressed vowel following the accented syllable, i.e. outside the target syllable but in relation to it. As for the alignment of L, authors find that it was consistently aligned within the onset of the target syllable- sometimes spreading well into the target vowel. As for the differences in alignment between the two dialects, the authors find there to be a significant difference only in regards to the alignment of the valley. The onset of the stressed vowel was a consistent segmental point for the southern variety, whereas the onset of the consonant was a reference point for the northern variety. They conclude that languages do indeed possess established segmental landmarks that tones anchor to (TBUs), and that there are indeed cross-linguistic differences- in the form of a 'phonetic continuum of alignment' -regarding this mechanism (p. 192).

Another cross-linguistic comparative study is Ladd, Schepman, White, Quarmby and Stackhouse (2009) analysis of Scottish and RP English. The study looks into the tonal alignment patterns in the two dialects with the goal of analysing nuclear alignment patterns, pre-nuclear alignment patterns, as well as the effect of vowel length on alignment. They include this aspect since languages have been reported to demonstrate different alignment patterns for pre-nuclear accents in the vicinity of ips, and PWds, and others for nuclear accents in the vicinity of IPs, as was discussed earlier (Arvaniti, Ladd & Mennen, 2006, D'Imperio, 1997a, among others). The studies show that the type of pitch accent (nuclear, pre-nuclear) is aligned differently in respect to an upcoming boundary or accent and would therefore modify the alignment of the peak sometimes beyond its respective TBU. Pre-nuclear accent modification was shown to be caused by the upcoming nuclear (final) accent, and the nuclear accent's modification is caused by the upcoming edge tone, all to avoid *Tonal crowding*. The studies also show that this alignment is subject to language-specific implementation rules and speech rate.

For the pre-nuclear accents experiment, data consisted of 66 sentences with target syllables produced with a rising pitch accent L+H. Those syllables were followed by 1 to 3 unstressed syllables and another accentable word before a major phrase boundary. All were produced by 8 speakers (4 RP speakers, 4 Scottish speakers). They measured the alignment of the peak relative to the offset of the target accented vowel (distance between H- and end of vowel), where negative values indicate alignment before end of vowel, and positive values indicate alignment following the vowel. Their results show that there is indeed a difference in alignment between dialects and according to vowel length. Scottish peaks were aligned systematically later than English peaks, however no significant difference was observed in the alignment of the valley between the dialects. Again, it was found that in English, peaks were aligned during the vowel if it is long and following the vowel (on the following consonant) if it is short.

For nuclear accents, two experiments were carried out, consisting of 477 target sentences all together. Target words were embedded in short sentences and placed in absolute final position where they would bear the last pitch accent in the intonational phrase. Target syllables were in initial and final positions in the word and were produced by 6 speakers (3 of each dialect). For these experiments they measure peak alignment both relative to the offset and onset of the target accented vowel. Their results indicate a difference in alignment between pre-nuclear and nuclear peaks, where the nuclear peaks were aligned 32 ms after onset, 100 ms before offset, while pre-nuclear peaks were aligned 75 ms after onset, 6 ms before offset. They also report a difference in alignment between the dialects parallel to what was found in the first experiment. Additionally, they observe an effect of vowel length on alignment where the peak is aligned earlier (away from the onset of the vowel and towards the vowel offset) in vowels of longer duration. A similar effect of stress position was also reported where the peaks were aligned earlier in final position and later in penultimate position. These results in their view support a SAH analysis of the alignment patterns observed in the dialects.

In her comparative study of the production of Canadian English prosody by German migrants to Canada in comparison to monolinguals from both languages, de Leeuw (2008) follows the approach in Atterer and Ladd (2004). Relevant to the present thesis is her methodology for tonal alignment measurement. In addition to obtaining the absolute distance in Ms between peak, valley, and different segmental landmarks as was done in previous studies, the author further includes a relative alignment measurement. As a response to the observation that absolute alignment could be affected by intra/inter-speaker variation in speech rate and segmental durations, relative alignment aims to control for different speech rates in larger corpora as well as variable segmental durations in target speech. She therefore obtains the relative duration of an accent as a ratio based on the distance between L, H and the onset of the target syllable divided by the duration of the syllable as a whole. In the following is a schematic representation of the relative measurement of L, and H from her thesis (p. 145):

$$MIN_{rel} = \frac{MIN - C0}{\mathrm{C1} - \mathrm{C0}}$$

$$MAX_{rel} = \frac{\mathrm{MAX} - \mathrm{C0}}{C1 - C0}$$

Figure 61: Relative measurement of tonal alignment in de Leeuw (2008). C0 is the onset of the target syllable, while C1 marks either the onset of the consonant following the target vowel or the onset of the consonant in the following unstressed syllable.

Based on this, the author concludes that the larger the MIN_{rel} and MAX_{rel} values are, the later the tone is aligned relative to the onset of the target syllable. Indeed, her findings corroborate those in Atterer and Ladd (2004) that German monolingual speakers align L earlier within the target syllable, while H is aligned late (outside) with respect to the target syllable. Her findings also support the idea that intricate phonetic detail regarding prosody- tonal alignment in particular- is highly language- specific or language demarcating, so that it could be carried over from L1 to L2 in bilinguals.

In a study on tonal alignment of pre-nuclear peaks in modern Greek, Arvaniti, Ladd and Mennen (1998) study bitonal accents L+H patterns to determine which one in this sequence is the starred tone (i.e. the one that is primarily associated with the accented syllable). Corpus for this study was 22 sentences produced by five Greek speakers (3 females, 2 males). Target syllables were open and closed carrying antepenultimate lexical stress and were followed by the number of 1, 2 or 3 unaccented syllables. Vowel identity was also variable in target syllables. The study measures the interval from H to the onset and offset of the accented syllable, and the L from the onset of the accented syllable. The study also measures the distance from the H to the onset of the post-accentual vowel, and the rise duration between the L and H.

They tested three hypotheses in their study: 1) in a bitonal accent, the trailing tone occurs at a fixed distance from the leading tone, regardless of the segmental composition of the syllable, 2) alignment of the peak is at a fixed distance from the end of the word (right edge of the word), thus is variable and affected by segmental composition that is limited by word edge, and 3) alignment of the peak is variable and modified due to an upcoming accent to resolve stress clash or tonal crowding. Their reported findings suggest that the H tone is not affected by segmental composition, and all else being equal, H is aligned after the onset of the unstressed/post-accentual vowel, while the L is aligned on or before the onset of the accented syllable in this language. They conclude that the alignment of the peak is not influenced by syllable duration. They also find that the distance between the two tones in this bitonal accent is not fixed and is affected by segmental composition. However, their results also clearly indicate that although this distance is variable, the two tones consistently take the accented syllable onset and the post-accentual vowel as anchor points for alignment. On the other hand, they find inconsistent results for the general effect of an upcoming accent in stress clash cases on the alignment of the H tone. For some speakers the number of intervening syllables between the accents did not influence the location of the H at all, while for others there was an effect on the peak location that was retracted depending on whether there was 1 syllable or maximally three between the accents.

In light of the aforementioned studies and analyses we can see that tonal alignment patterns and behaviour is reported to: 1) cue and mark prosodic constituents (especially IPs and ips), 2) cue and demonstrate the canonical TBU *Tone Bearing Unit* in a given language or variety, and 3) reveal the rules governing the implementation of tones in a language as a function of the different contextual factors. Accordingly, this study will aim to analyse tonal alignment patterns in JA and its theoeretical implications for the classification of the dialect in prosodic typology. As mentioned, there are no instrumental studies on this dialect, and the current study will benefit from the studies on the other Arabic dialects reviewed here. Moreover, although Al-Zaidi's (2014) study does not report on Jeddah Arabic specifically, nor carry out detailed analyses, it does make an indication of the alignment behaviour to expect among Hijazi dialects. Across the information structure experiments carried out in his thesis (reviewed in section 5.2), Al-Zaidi reports that target peaks are always aligned within the accented syllable, starting at the onset of the syllable and reaching the peak within that syllable (p. 93). The current thesis will therefore use this observation along with the results of previous Arabic studies as working hypotheses.

6.3 Experiment design and recordings

To investigate tonal alignment patterns, the design of the experiments must control for a number of contextual and microprosodic factors. As was discussed in the methodology chapter [3], studies have shown that the segmental composition and identity of the string (duration, manner, voicing, length) affect the peak location. Therefore, sonorants are chosen for alignment studies as they maintain the F0 contour and minimise tracking errors (Prieto et al., 1995).

More importantly in this regard, is to explore the relationship between tonal alignment patterns and tone bearing-units. The stressed syllable in intonational literature and across languages is regarded as the canonical tone-bearing unit. However, some studies found that constituents smaller than the syllable can be tone-bearing units. The main hypothesis driving this TBU investigation is as follows: if alignment varies according to syllable moras and foot internal structures, then the canonical syllable isn't the TBU in the language (Reported as the foot in Egyptian Arabic; Hellmuth, 2006a). Conversely, if patterns of alignment are consistently within the confines of the syllable and do not show significant variation according to type and number of moras, then it is the TBU in the language (Reported as the syllable for Lebanese Arabic; Chahal, 2001). Defining a TBU has implications for the concept of association in the theory- the association of accents to positions in the metrical structure; thus revealing information about the language's metrical configuration. The exact identity of TBU is additionally reported to be language-specific according to the language's syllable structure (Arvaniti, 2017). Observing the alignment behaviour in JA is one of the main goals of the experiment.

Bearing all these aspects and studies in mind, the scripted material to investigate tonal alignment was designed as follows. Four tri-syllabic target words with varied target stressed syllable locations (initial-medial-final) and constant segmental compositions. All the syllables had the sonorant /l/ as onset, long and short /a/ vowel as nucleus and, /t/ as coda in closed syllable type. Yielding target syllables of the shapes: /la/, /laa/, /lat/ and /laat/, that vary in syllable type and mora count respectively. In order to be able to reach an informed conclusion, the design in this study expands the Arabic syllable structure inventory to include further syllable types (open syllables varying in mora count) excluded from previous Arabic studies in relation to tonal alignment. Further methodological inadequacies are avoided in order to get the best outcome. For example, in Chahal's (2001) thesis, for the purpose of varying stress position, the author uses morphologically inflected words, which further contain a geminated consonant. This consequently yielded target syllables with different segmental identities making it difficult to maintain comparable target contexts. Furthermore, studies show that tonal location is indeed sensitive to the presence of geminated consonants in target syllables, demonstrating an effect of gemination on tonal alignment (D'Imperio et al., 2007 for Italian). The target words for the current study are shown in the table below:

Initially stressed	Medially stressed	Finally stressed
ˈla.ha.fu 'he/they nicked it'	mu.ˈ**laa.**zim 'lieutenant' riH.ˈ**lat**.hum 'their journey'	lil.Haf.ˈ**laat** 'for the parties'

Table 74: Alignment Experiment Target words (stressed syllables in bold)

Varying lexical stress location within a word (initial-medial-final) allows for investigating the effect of proximity to various prosodic units on tonal alignment. By doing this we serve two purposes: first, we observe how the stressed and accented syllable interacts when they co-occur with a phrase boundary tone marking ips and IPs; second we observe how these syllables behave (in terms of alignment) when they are placed further to the left from these right-edge boundaries. That is, when the accented syllable is initial in tri-syllabic word in sentence final position; for example, it is separated from the IP boundary by 2 syllables, if it is final it is separated by 0 syllables, and so forth. The experimental designs in Chahal (2001) and Prieto et al. (1995) further attempt to observe the effect of the prosodic word boundary PW on tonal alignment, by observing the tonal alignment within each word in each boundary-final and non-boundary final syllables. This is also done in this thesis. Varying stress location also helps examine the stress clash context, as explained earlier, whereby target syllables occur successively with no intervening phrasal (IP, ip) boundary. To achieve this, the target words above are placed preceding control words that also vary in stress location. By virtue, the two words are accented; the target word receives a pre-nuclear accent, while the control word receives a nuclear accent. The intervening unstressed syllables between the two words vary from 0, 2, 3 and 4 syllables.

Accordingly, following the designs in Chahal (2001), Jun and Fletcher (2014) and Prieto et al. (1995), the target words are embedded in three structural positions in the sentences at three different prosodic boundaries; prosodic word, intermediate phrase, Intonational Phrase.

These elicit three phrasing levels of different strengths, sentence medial position, end of ip, and end of IP. In the following is a detailed example of the contexts in which the initially stressed word /la.ha.fu/ is placed, and the level they encode- the same structure goes for the remaining words in table 3 above. The full list of target sentences is provided in Appendix [D].

End of IP (nuclear): She should have said **Y** 'control'. But she said **X** 'target'

Kaan laazim tiguul lafaHu. bas gaalat ***lahafu***

End of ip (prenuclear): She should have said **X** but she couldn't

Kaan laazim tiguul ***lahafu,*** bas maa gidrat

Sentence medial/stress clash (prenuclear): She should have made clear that **X** 'target' **Y** 'control' (unstressed syllables underlined):

Kaan laazim tiwaDDiH 2innu ***ˈla.<u>ha.fu</u> ˈriz.gu*** **(2 intervening unstressed syllables)**

Kaan laazim tiwaDDiH 2innu ***ˈla.<u>ha.fu</u> <u>ma</u>.ˈkaa.ni*** **(3 intervening syllables)**

Kaan laazim tiwaDDiH 2innu ***ˈla.<u>ha.fu</u> <u>Ta</u>.<u>la</u>.ˈbaat*** **(4 intervening syllables)**

Each speaker produced 60 sentences for this experiment. As seen above, each word has 5 different structural positions (IP, ip, stressclash 1, stressclash 2, stressclash 3) and each case was repeated 3 times, which totals as 15 productions for each case. An overall number of 1200 data points were collected for this experiment:

4 target words x **5** structures x **3** repetitions x **20** speakers = **1200**

6.3.1 Recordings

Participants were instructed to read the scripted sentences carefully and as naturally as possible yielding target declarative sentences read at normal speech rate. Speech rate related effects on peak alignment are not the goal of this experiment (Cf. Ladd et al., 1999). Target sentences were randomised and included filler sentences in between. A production was excluded from the analysis if:

1- Target produced incorrectly or uncarefully, or when produced with hesitations (disfluencies)
2- Target sentence produced differently from the rest of the productions and from the rest of the group.
3- Target word was unaccented (not realised with a pitch accent).

As a result, **73** productions were excluded, leaving **1127** utterances for analysis.

6.3.2 Quantitative measurements

Speech material for this experiment was transcribed both quantitatively and qualitatively. Qualitatively in terms of pitch accents, and boundary tones following the ToBI system notation. That is, sentences were examined for the occurrence of pitch accents, phrase and boundary tones. In addition to the shape/ type of pitch accent that occurred on target words. Accordingly, target word, syllable, and sentence were the domain for acoustic analysis. Maxima H and minima L F0 were labelled in the target word and syllable. The position of the peaks and valleys was determined with reference to the target syllable. Maxima are identified where the highest F0 point occurs within each target syllable. The minima are identified with reference to the maxima and start of the syllable, where the lowest point within the domain preceding the located Max F0 occurs. The max and min were obtained after correcting for each speaker range via the two- pass method (Hirst, 2011, same method used in the previous focus experiment).

Quantitatively, relative and absolute alignment measurements were carried out. Following previous studies, the potential segmental landmarks in target words were identified as 1) for the alignment of valley L: onset of stressed syllable, and onset of stressed vowel, 2) for the alignment of peak H: onset and offset of the stressed vowel, onset of the stressed syllable, and offset of the stressed rhyme. Because an established reference point for tones in this dialect is still unknown, this study explores the aforementioned anchor points in order to reach an informed result that would aid the comparison of the current data with published data. Additionally, the alignment of the peak and valley will be observed across syllable structures in order to identify the most reliable segmental anchor. Note the following schematic representation of the reference points in each target syllable:

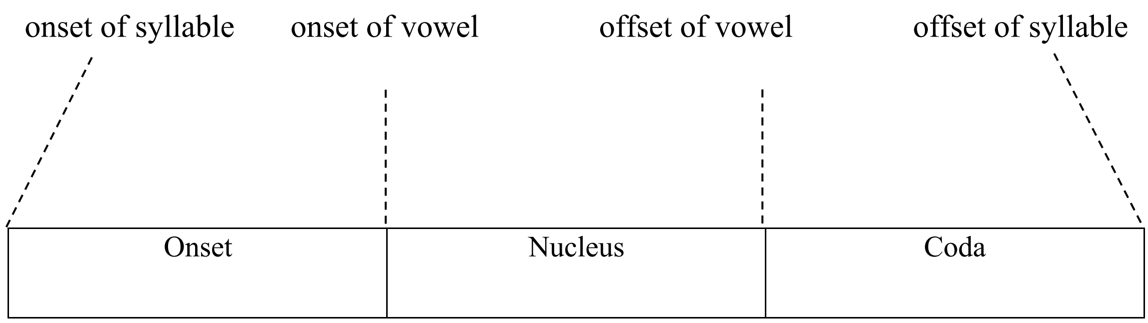


Figure 62: Segmental landmarks explored.

Durational measurements were obtained for target accents, vowels and syllables. Following Ladd et al. (1999) the duration of the target accent (rise time: temporal interval between the start and end of the rise) was obtained to be correlated with the duration of the syllable to observe the extent of the effect of segmental/syllabic duration on the duration of the accent. In other words, if there is a positive correlation between the two, then this shows that the duration of the accent is anchored to specific landmarks regardless of segmental identity, and vice versa.

For absolute measures, the absolute distance in milliseconds was measured between the reference points and the location of L. Similarly, the distance in milliseconds was measured between the reference points and the location of H. The larger the difference (Ms), the later the peak aligned with the anchor points. Also negative values would indicate alignment before the established anchor points for L and H. Note the following formulas used to capture the distances:

H – σ ONSET (distanceMaxSyllableToOnsetMs)

H – V ONSET (distanceMaxSyllableToRhymeMs) vowel onset

H – V OFFSET (MaxToEndOFVowelMs)

H – σ OFFSET (maxToEndOFSyllableMs)

L – σ ONSET (distanceMinsyllableToOnsetMs)

L – V ONSET (MinToVowelMs)

As was mentioned earlier, because speaking rate across speakers/tokens is subject to high variability, a calculation of F0 location would need to take into account the variability in segmental durations among speakers. That is, to be able to say that L or H occur at some distance from a landmark, we would need to account for the individual segmental durations within that landmark. Therefore the literature suggests (Hellmuth, 2006a, Atterer & Ladd, 2004, de Leeuw, 2008) a proportional measure of tonal alignment, which allows us to state that a particular F0 point is located at some percentage of the whole duration of the syllable. Thus, in addition to the absolute measures (in Ms) above, relative alignment measures were calculated. The ratios were calculated by the method in figure [62] to obtain the temporal difference between L, H and the onset of the syllable in relation to the overall duration of the syllable Again, the larger the ratios are, the later the alignment.

Throughout the statistical analyses in this chapter, the assumption of normal distribution is considered. The assumption of normal distribution was tested by using the Shapiro-Wilk's test. The assumption that homogeneity of variance within the population is equal was also considered in order to conduct relative tests. Tests of homogeneity were conducted by using the Levene test. When satisfied, sets of Analysis of variance tests were carried out accordingly. When violated, firstly the outlier values were removed. Then the assumptions were checked again. If the assumptions were still violated, rank transformations of these variables were taken into consideration. In the analyses gender and prosodic boundary are included as independent variables. Alongside those detailed tests, logistic regression tests were carried out taking into consideration variations along the many dimensions included in the experiment in an attempt to construct a more generlised vision of the data.

6.3.3 Experiment questions

The tonal alignment experiment was carried out to investigate the following questions:

1. How does proximity to prosodic boundaries of different strengths (ProsodicWord, ip, and IP) affect tonal alignment?
    - Where does the tone align with respect to the segmental landmarks in each case?
    - Does peak alignment behaviour around these constituents contribute to their marking?
    - Having analysed the results for each, which segmental landmark can be said to constitute the best anchor for the alignment of respective L and H in this Arabic dialect?
    - Which of the prosodic boundaries has the most effect on peak alignment?
2. How does syllable duration/weight, and stress clash affect tonal alignment?
    - Where does the peak align with respect to the onset of the syllable for each syllable type according to the number of intervening syllables (0, 1, 2, 3, and 4)?
    - Is there a difference in alignment with segmental landmarks according to syllable type [open/closed]?
    - How much of the variation in accent duration can be explained by overall accented syllable duration? Is there a correlation between accent duration and syllable duration?
    - How much of the variation in the location of the peak relative to the onset can be explained by overall accented syllable duration? Is there a correlation between the two?

6.4 Results

In the following are mean alignment results in proportional distance from the onset (maxalign, minalign) and in absolute distance in milliseconds between the detected L, H and the specified segmental landmarks included in the corpus. These are presented according to intonational level, as well as combining all levels. Negative values indicate alignment before the respective landmark:

	H-TO-SYLLABLE-ONSET	H-TO-VOWEL-ONSET	H-TO-VOWEL-OFFSET	L-TO-SYLLABLE-ONSET	L-TO-VOWEL-ONSET	MAX ALIGN	MIN ALIGN
Overall	78.7	33.4	-63.7	12.5	-32.7	0.46	0.06
ip	85.44	35.54	-79.1	51.41	1.57	0.42	0.21
IP	49.73	1.14	-97.46	-24.54	-73.12	0.29	-0.11

Table 75: H and L alignment patterns relative to the segmental landmarks in the corpus

The means indicate that generally in the corpus, L is aligned just under 13 Milliseconds into the onset consonant, while the H is aligned approximately 34 Milliseconds into the vowel. These means also differ according to intonational level, as can been seen. More specifically, a tone occurring in the vicinity of an IP boundary is pushed earlier to the left with respect to the segmental landmark than a counterpart tone in the vicinity of an ip. Across intonational levels, H and L are realised later with respect to the syllable onset in ips, as compared to an earlier location in IP vicinity. This suggests that there is a potential interaction between the H location and syllable duration and makeup. This aspect will be further investigated in the upcoming section on alignment variation according to syllable type.

More generally, proportionate distances (maxalign, minalign) show that L aligns earlier in the syllable, while H aligns later in the syllable. These are reflected in the overall results, as well as in the results according to intonational level and syllable type in table [76] below. The L consistently aligned relative to, on/before the onset consonant of the target syllable, whereby the H aligned relative to the rhyme of the syllable. Additionally, as can be inferred from table [77] it can be said that in the general tonal makeup of the syllable, L is realised proportionately earlier at less than 10% through the duration of the syllable, while the H gesture is realised proportionately later at almost 50% through the duration of the syllable. This is the general result across speakers, intonational levels, stress clash and segmental landmarks. This comes in accordance with the findings discussed throughout the quantitative analyses here, whereby maintaining that the peak is realised somewhere relative to the rhyme of the syllable. This alignment pattern could potentially support the claim that only one rising accent of the shape L+H* exists in JA as showed in chapter [4]. Based on the current data, there is no motivation for a categorical difference according to alignment (e.g. L+H* vs. L*+H) in the dialect at this stage.

	IP		ip	
Syllable type/proportionate measure	Max Align	Min Align	Max Align	Min Align
Heavy	0.25	0.05	0.35	0.12
closed	0.22	0.05	0.35	0.18
open	0.29	0.04	0.36	0.00
Light	0.42	-0.32	0.62	0.12
open	0.42	-0.32	0.62	0.12

Table 76: Proportionate measures by intonational level and syllable type.

	Max align		Min align	
	Mean	SD	Mean	SD
Ratio	0.46	0.29	0.06	0.90
Percentage	46.80%		6.14%	

Table 77: The alignment of L, H as a proportion of the syllable duration across corpus.

To check the patterns for each word and syllable type, the analyses proceed as follows: first on each word individually according to intonational level, then analyses proceed according to syllable type and stress clash context. Individual alignment results with reference to each segmental landmark for each male and female in the corpus are presented in Appendix [B].

6.4.1 /ˈlahafu/

Analyses were run on the most important segmental anchors mentioned in the literature: onset of syllable, onset of vowel (rhyme), and end of vowel. Those anchors were analysed to observe whether the mean differences we find in H, L alignment according to the word being in ip, or IP are in fact statistically significant. Please refer to the figures in section 6.4.5 for a schematic representation of the alignment of the tones relative to these segmental landmarks and according to the durations of the onsets and rhymes.

Summary statistics for H

	v1		v2		v3	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
ip	22.4	25.7	80.3	31.0	-45.3	23.2
IP	-4.96	33.4	47.5	31.4	-65.3	33.4

v1: distanceMaxSyllableToRhymeMs, v2: distanceMaxSyllableToOnsetMs,
v3: maxToEndOFVowelMs

Table 78: /lahafu/ peak summary statistics

Summary statistics for L

	v4		v5	
	mean	std. deviation	mean	std. deviation
ip	15.6	16.3	-42.3	15.0
IP	-48.4	302	-101	306

v4: distanceMinSyllableToOnsetMs, v5: minToVowelMs

Table 79: /lahafu/ valley summary statistics

These results indicate that in ip, across the chosen segmental landmarks, L is realised within the onset syllable and before the onset in IP. For the alignment of the peak, it is realised later in the vowel in ip, and earlier than this landmark in IP. It is in fact pushed outside the vowel and realised in the syllable onset in this case. Regarding statistical analyses, two-way ANOVAs were conducted. The independent variables are gender (F, M) and condition (ip, IP). The dependent variables are (v1-5). For the alignment of H in this word, Anova results report a significant difference based on intonational level (v1: $F=24.575$, $p=2.56e-06<0.001$, v2: $F=28.046$, $p=6.21e-07<0.001$, v3: $F=13.936$, $p=0.00<0.001$). There is also a significant difference based on gender for (v1) and (v2) but not for (v3): $F=4.11$, $p>0.05$). For L, there is not a statistically significant difference according to the intonational levels nor gender (v4: $F=0.395$, $p=0.5$, v5: $F=3.874$, $p=0.5$). The results for both H and L indicate that there is no significant interaction between intonational level and gender.

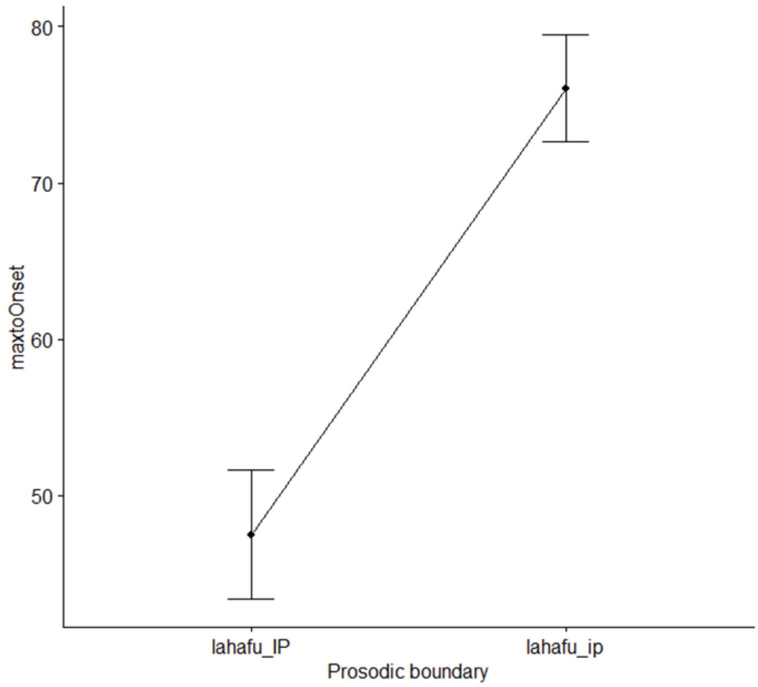


Figure 63: /lahafu/ peak alignment relative to syllable onset. Middle dot represents the mean, upper and lower whiskers represents the standard deviation.

6.4.2 /riH`lathum/

Summary statistics for H

	v1		v2		v3		v6	
	mean	std. deviation	mean	std. deviation	mean	std. deviation	mean	std. deviation
ip	19.1	47.0	66.5	48.3	-55.2	44.7	-126	45.5
IP	-4.77	41.2	47.2	42.1	-73.3	42.3	-142	48.8

v1: distanceMaxSyllableToRhymeMs, v2: distanceMaxSyllableToOnsetMs, v3: maxToEndOFVowelMs; v6: maxToEndOFSyllableMs

Table 80: /riHlathum/ peak summary statistics

Summary statistics for L

	v4		v5	
	mean	std. deviation	mean	std. deviation
ip	41.8	166	-5.57	164
IP	-0.03	182	-52.0	184

v4: distanceMinSyllableToOnsetMs, v5: minToVowelMs

Table 81: /riHlathum/ valley summary statistics

The results indicate that across the landmarks in this word, L is realised into the onset of the syllable, while it is pushed outside the syllable in the IP vicinity. For the alignment of the peak, it is realised into the vowel in the ip vicinity, and outside the vowel and within the onset in the IP vicinity. Regarding statistical analyses, two-way ANOVAs were conducted. The independent variables are gender (F, M) and condition (ip, IP). The dependent variables are (v1-6). Anova results show that for H, there is a significant difference based on intonational level and gender (v1: F=9.119, p=0.00<0.01, v2: F=4.989, p=0.02<0.05, v3: F=5.465, p=0.02<0.05, v6: F=4.201, p=0.04<0.05). For L, there is a significant difference based on intonational level only for (v5) (F=6.446, p=0.01<0.05). For (v4) however, there is not a statistically significant difference according to gender nor intonational level (F=3.028, p=0.08<0.1). The results for both H and L indicate that there is not significant interaction between gender and intonational level.

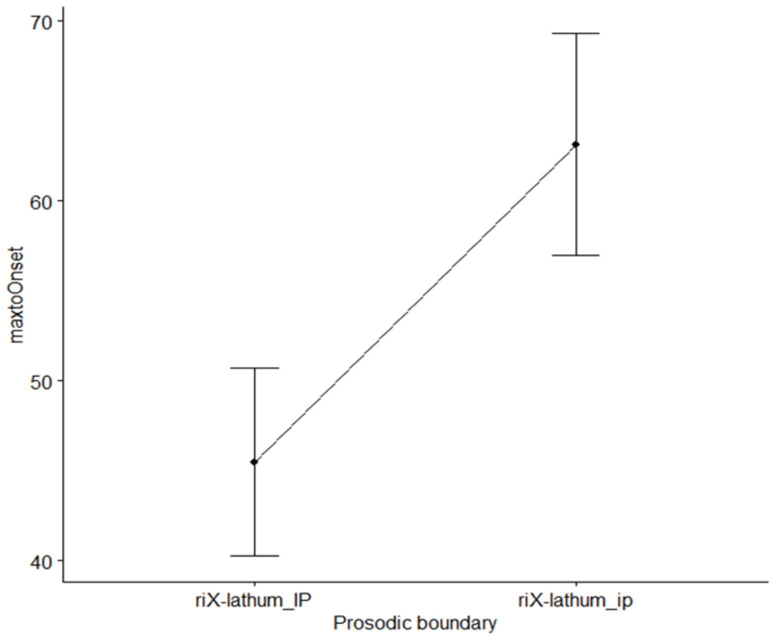


Figure 64: /riHlathum/ peak alignment relative to syllable onset. Middle dot represents the mean, upper, and lower whiskers represents the standard deviation.

6.4.3 /mu`laazim/

<u>Summary statistics for H</u>

	v1		v2		v3	
	mean	std. deviation	mean	std. deviation	mean	std. deviation
ip	34.2	56.0	74.0	58.2	-122	48.9
IP	14.3	31.5	52.8	30.7	-129	39.2

v1: distanceMaxSyllableToRhymeMs, v2: distanceMaxSyllableToOnsetMs,
v3: maxToEndOFVowelMs

Table 82: /mulaazim/ peak summary statistics

<u>Summary statistics for L</u>

	v4		v5	
	mean	std. deviation	mean	std. deviation
ip	1.73	5.22	-38.1	8.09
IP	8.77	28.9	-29.8	28.2

v4: distanceMinSyllableToOnsetMs, v5: minToVowelMs

Table 83: /mulaazim/ valley summary statistics

For this word, L is located within the onset in both ip and IP. However it is earlier in the onset in ip, and later in the IP vicinity. As for the alignment of the peak, it is realised earlier into the vowel in IP, and later into the vowel in ip. Regarding statistical analyses, two-way ANOVAs were conducted. The independent variables are gender (F, M) and condition (ip, IP). The dependent variables are (v1-5). For H, there is not a statistically significant difference between the intonational levels for this word (v1: F=3.813, p=0.05<0.1, v2: F=2.293, p=0.13, v3: F=0.026, p=0.8). Likewise, for L, there is not a statistically significant difference between the levels (v4: F=0.035, p=0.8, v5: F=1.456, p=0.2). Differences according to gender are reported significant for (v1), (v3) and (v4) (p<0.05). As for the interaction between gender and intonational level, only the results for the (v1) and (v3) indicate that there is a significant interaction (v3: F=6.481, p=0.0123<0.05).

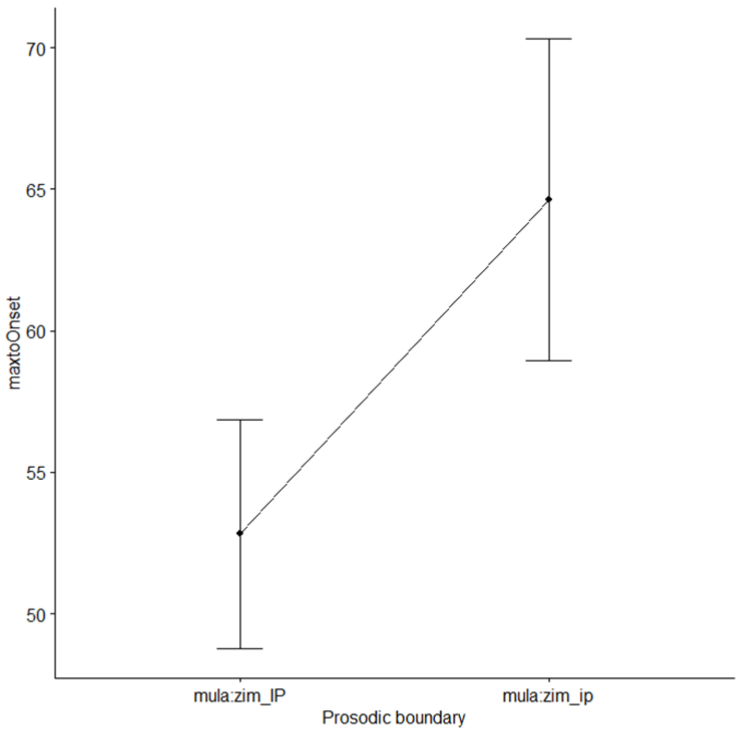


Figure 65: /mulaazim/ peak alignment relative to syllable onset. Middle dot represents the mean, upper and lower whiskers represents the standard deviation.

6.4.4 /lilHafˁlaat/

Summary statistics for H

	v1		v2		v3		v6	
	mean	std. deviation	mean	std. deviation	mean	std. deviation	mean	std. deviation
ip	67.6	103	122	108	-94.2	75.8	-184	79.9
IP	-0.04	61.2	51.8	62.6	-128	62.7	-207	70.8

v1: distanceMaxSyllableToRhymeMs, v2: distanceMaxSyllableToOnsetMs,
v3: maxToEndOFVowelMs; v6: maxToEndOFSyllableMs

Table 84: /lilHaflaat/ peak summary statistics

Summary statistics for L

	v4		v5	
	mean	std. deviation	mean	std. deviation
ip	15	365	-39	365
IP	-67.1	361	-119	363
v4: distanceMinSyllableToOnsetMs, v5: minToVowelMs				

Table 85: /lilHaflaat/ valley summary statistics

Finally, the alignment of L in this word is realised within the onset in ip, and outside the onset in the vicinity of an IP. The peak is realised within the target vowel in ip, and outside the vowel in IP vicinity. The H is in fact pushed earlier into the onset in the IP vicinity. Regarding statistical analyses, two-way ANOVAs were conducted. The independent variables are gender (F, M) and condition (ip, IP). The dependent variables are (v1-6). For H, there is a significant difference based on intonational levels for (v1), (v2), and (v3) (v1: $F=12.405$, $p=0.00<0.001$, v2: $F=11.71$, $p=0.00<0.001$, v3: $F=5.957$, $p=0.01<0.05$), but not (v6) ($F=3.219$, $p=0.07$). For L, a significant difference is reported for v4 ($F=4.041$, $p<0.05$), but not for v5 ($F=1.609$, $p=0.208$). According to gender, a significant difference is reported for (v1) and (v2) only. As for the interaction between gender and intonational level for H, the results for (v1, v2, and v3) indicate that there is a significant interaction between level and gender ($p<0.01$). As for the interaction between gender and intonational level for L, the results indicate that there is no significant interaction between level and gender ($p>0.05$).

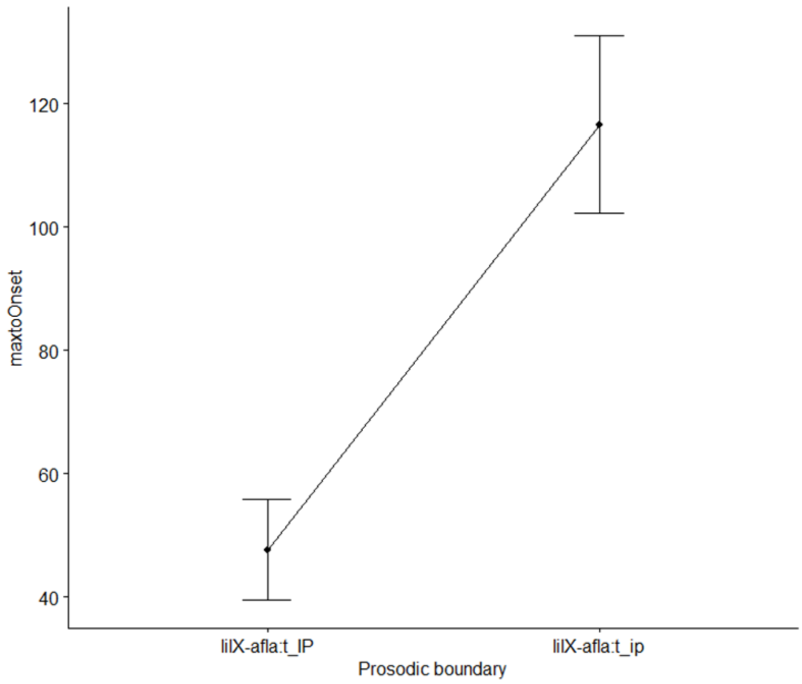


Figure 66: /lilHaflaat/ peak alignment relative to syllable onset. Middle dot represents the mean, upper and lower whiskers represents the standard deviation.

6.4.5 Schematic representations

In the following are schematic representations of the short vowels and the approximate location of the tone according to intonational level and segmental landmarks in figure [59] (approximate locations are based on the averaged durations for each syllable type reported in previous tables):

1. **End of intermediate phrase ip:**

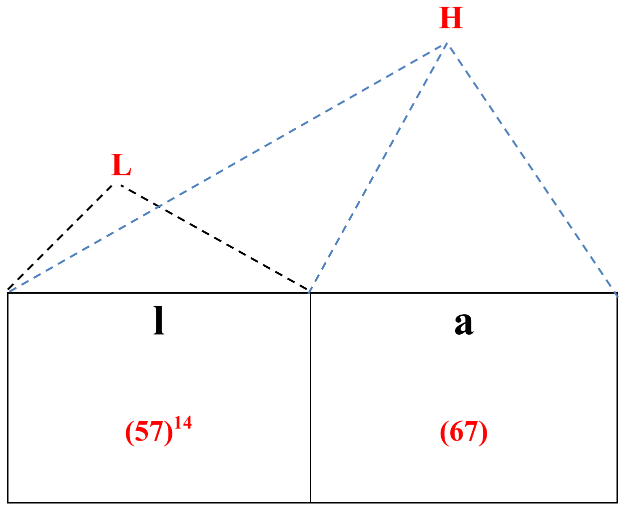


Figure 67: short vowel tone alignment patterns in intermediate phrase

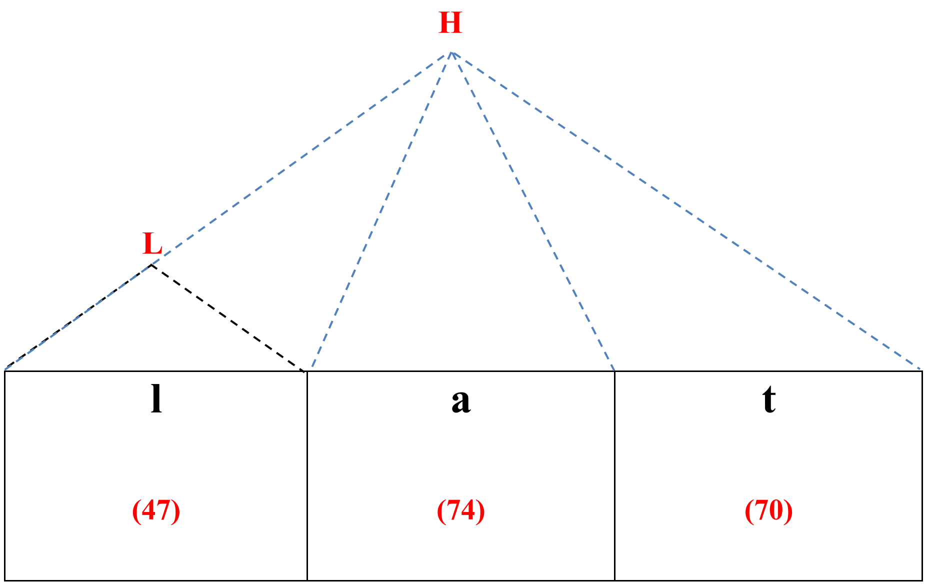


Figure 68: short vowel tone alignment patterns in intermediate phrase

[14] Mean segment durations in brackets.

2. End of intonational phrase IP:

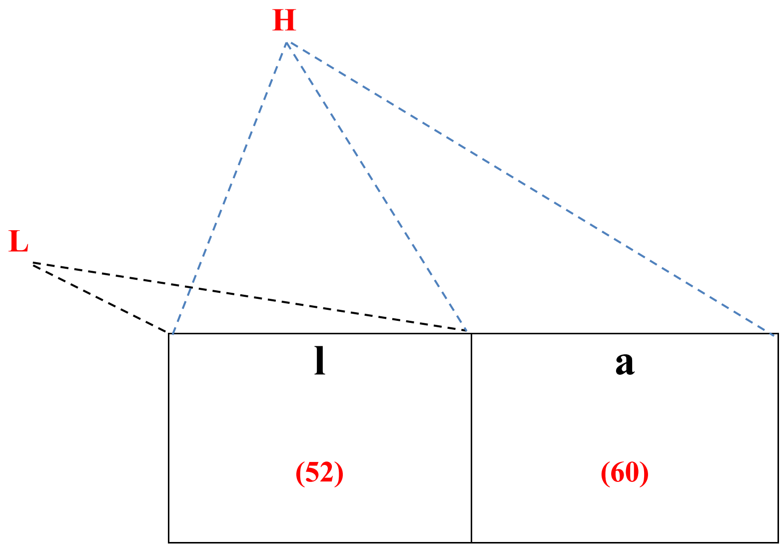


Figure 69: short vowel tone alignment patterns in intonational phrase

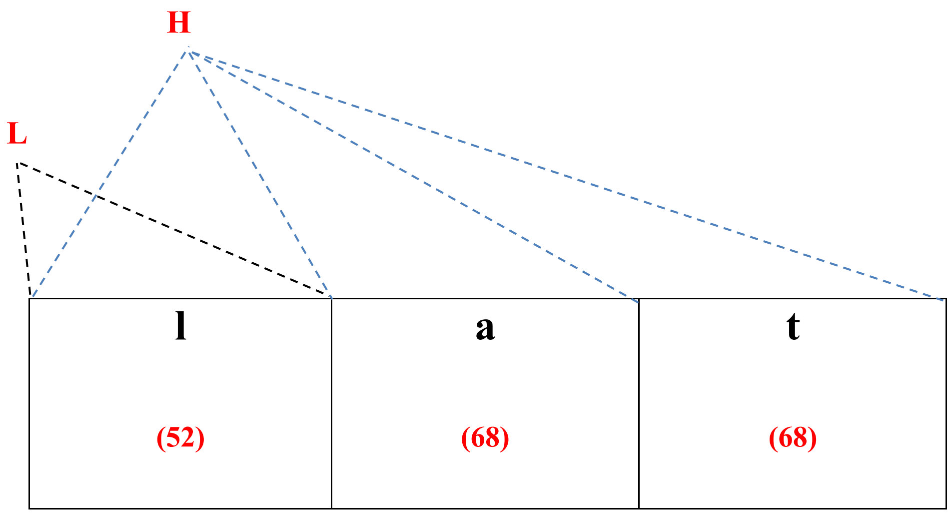


Figure 70: short vowel tone alignment patterns in intonational phrase

In the following are schematic representations of the long vowels and location of the tone according to intonational level and segmental landmarks in **Figure** ***62*** (mean segment durations in brackets):

1. **End of intermediate phrase ip:**

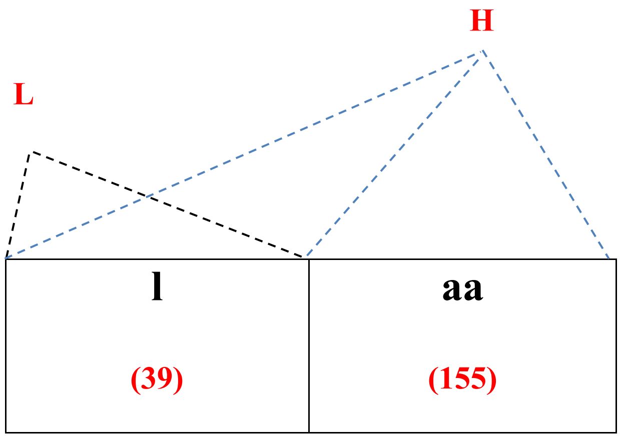


Figure 71: long vowel tone alignment patterns in intermediate phrase

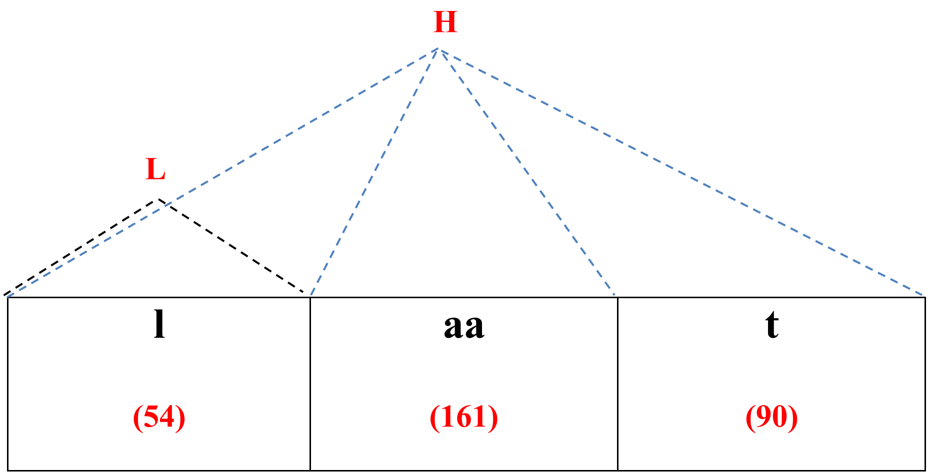


Figure 72: long vowel tone alignment patterns in intermediate phrase

2. **End of intonational phrase IP:**

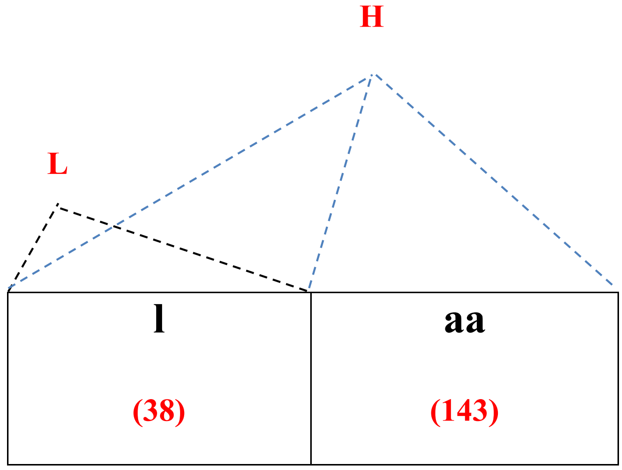


Figure 73: long vowel tone alignment patterns in intonational phrase

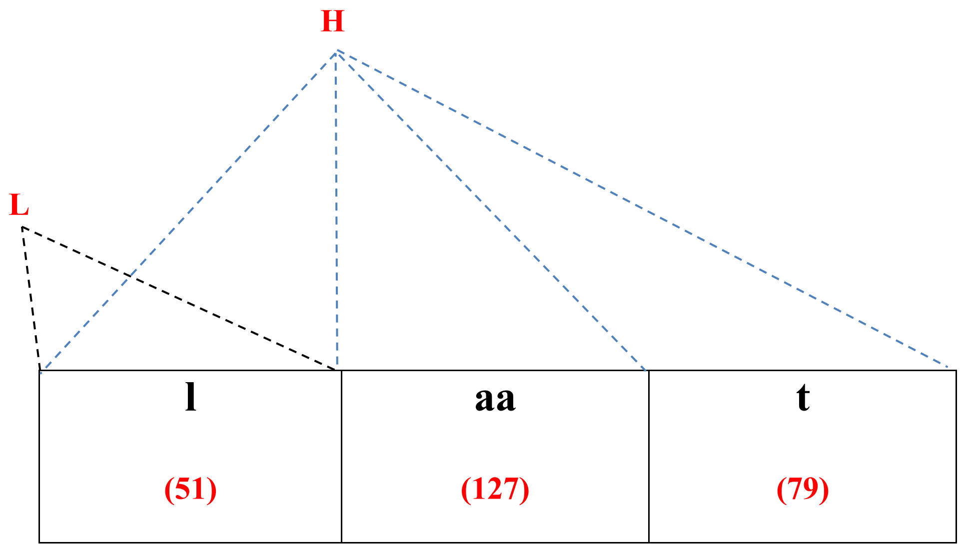


Figure 74: long vowel tone alignment patterns in intonational phrase

6.4.6 Interim discussion

As can be seen, the target words show differences in peak alignment location according to the intonational level they occur in. It is reported across segmental landmarks that the highest intonational level- the Intontaional Phrase- exerts a leftward push to the peak and valley, in comparison to the intermediate phrase. This may indicate that there is a variation in alignment as an effect of prosodic boundary. This effect would be taken as a tonal correlate for the occurrence of these intonational phrases in JA. Statistical evaluation of this claim of a prosodic boundary effect that takes into consideration the potential confounding fixed and random effects is carried out in the final step of the tonal alignment analysis. Related to the discussion of effects, variation according to gender and syllabic factors is also observed in the previous results. Throughout the analysis in this section, it is reported that according to syllable structure, long and short vowels, open and closed syllables do show some variation regarding the exact location of the peak. The analysis in the following section is thus dedicated to exploring the direction and significance of this observed variation along with the possible correlations between the different variables in alignment realisation.

6.5 Alignment variation according to syllable type

In the following are the results for analyses carried out in order to investigate the significance of the peak alignment variation according to syllable structure. Recall the targets for this study include four syllable types varying in whether the syllable is open or closed, heavy or light and varying in vowel length. These syllable types contain variation in the duration of the rhyme and therefore distance between the location of the peak and end of the rhyme is chosen for analysis.

First to be reported are the results according to syllable type grouping lahafu/mulaazim as open syllables, while riHlathum/lilHaflaat constitute the closed syllable group. Negative values indicate leftward alignment from the segmental landmark, while positive numbers indicate rightward alignment.

	ip	
	mean	std. deviation
maxToEndOFSyllableMs (open)	-83.5	54.0
maxToEndOFSyllableMs (closed)	-155	71.0

Table 86: Average distance in ms of the peak location to syllable offset in intermediate phrases.

The means show that the peak in ip vicinity takes a shorter distance to align relative to the end of the syllable when it is open, while it takes longer time to align in closed syllables. The observed longer time takes into consideration the duration of the coda consonant. Analysis of Variance and pairwise comparisons report that there is a significant difference in “maxToEndOFsyllableMs” between the two syllable types (v6: F=68.46, p>0.05).

	Dependent variable: ‘maxToEndOFSyllableMs (v6)’				
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
syllable type	1	236399	236399	68.46	1.07e-14 ***
Residuals	229	790753	3453		
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

Table 87: anova results for peak alignment relative to offset of syllable according to syllable type in intermediate phrases

Similarly, in Intonational Phrases, there is a significant difference between the average of the variable ‘v6’ between the two syllable types (v6: F=87.01, p<2e-16). The peak takes a shorter distance to align from the end of the syllable when it is open and takes a longer distance in closed syllables.

	IP	
	mean	std. deviation
maxToEndOFSyllableMs (open)	-97.5	48.4
maxToEndOFSyllableMs (closed)	-171	67.6

Table 88: Average distance in ms of the peak location to syllable offset in intonational phrases.

	Dependent variable: 'maxToEndOFSyllableMs (v6)'				
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
syllable type	1	261046	261046	87.01	<2e-16 ***
Residuals	221	663066	3000		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 89: anova results for peak alignment relative to offset of syllable according to syllable type in intonational phrases

Pearson correlation tests were carried out in order to further estimate the strength and direction of the observed relationship between peak location and the duration of the syllable under question. The variables named 'durationSyllableMS' and 'maxToEndOFSyllableMs' are found to be significantly correlated in both ips and IPs (correlation coefficients all significant at the p= 0.05 level or better). Please refer to Appendix [C] for correlation graphs and gender effects.

Second to be reported are the results according to vowel length grouping mulaazim/lilHaflaat as the long vowel group, and lahafu/riHlathum as the short vowel group. In intermediate phrases, there is a significant difference between the average of the variable 'v3' across the two vowel types (v3: F=59.34, p=3.97e-13). According to the pairwise comparisons, the differences between long and short (diff (short-long) = 60.50, p<0.001) is significant. The means show that the peak takes a shorter distance to align from the end of the vowel when it is short, while it takes a longer distance to align when the vowel is long.

	ip	
	mean	std. deviation
maxToEndOFVowelMs (long)	-108	64.6
maxToEndOFVowelMs (short)	-50.2	35.6

Table 90: Average distance in ms of the peak location to syllable offset in intermediate phrases

	Dependent variable: 'maxToEndOFVowelMs (v3)'				
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
vowel type	1	211399	211399	59.34	3.97e-13 ***
Residuals	229	815753	3562		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 91: anova results for peak alignment relative to offset of syllable according to vowel duration in intermediate phrases

A similar pattern is reported under intonational phrases, there is a significant difference between the average of the variable ‘v3’ between the two vowel types (v3: F=111.1, p= <2e-16). According to the pairwise comparisons, the difference between long and short (diff (short-long)= 74.56, p<0.001) is significant.

	IP	
	mean	std. deviation
maxToEndOFVowelMs (long)	-128	51.0
maxToEndOFVowelMs (short)	-69.4	38.3

Table 92: Average distance in ms of the peak location to syllable offset in intonational phrases

Dependent variable: ‘maxToEndOFVowelMs (v3)’					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
vowel type	1	309230	309230	111.1	<2e-16 ***
Residuals	221	614882	2782		
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

Table 93: anova results for peak alignment relative to offset of syllable according to vowel duration in intonational phrases

Pearson correlation tests confirm the observed relationship. The variables named ‘durationSyllableMS’ and ‘maxToEndOFVowelMs’ are found to be significantly correlated in both ips and IPs (correlation coefficients all significant at the p= 0.05 level or better). Please refer to Appendix [C] for correlation graphs and gender effects. In general, the patterns report that the longer the duration of the syllable is, the closer the peaks are aligned with respect to the end of the vowel, and the shorter the duration, the earlier the peak is aligned in the vowel. The scatter plots also demonstrate clear clustering according to syllable weight whose independent effect is discussed in the following.

Last to be reported is the complementary comparison according to syllable weight. As was presented earlier, statistical and correlation results report significant differences according to vowel and syllable structure. Thus, it is effective to present results for the individual effect of syllable weight. In this comparison and according to the stress assignment rules in the dialect, mulaazim/riHlathum and /lilHaflaat/ constitute the heavy syllable group, while /lahafu/ constitutes the light syllable group. The following tables and graphs represent the results according to prosodic boundary and syllable weight. Negative values indicate leftward alignment from the segmental landmark. The results of the alignment patterns relative to the end of the syllable are reported first.

Boundary	*maxToEndOFSyllableMs*	*Std. Deviation*
IP		
heavy	-156.12	62.36
light	-65.31	33.35
ip		
heavy	-143.41	66.04
light	-45.27	23.16

Table 94: Average distance in ms between peak location and target syllable offset.

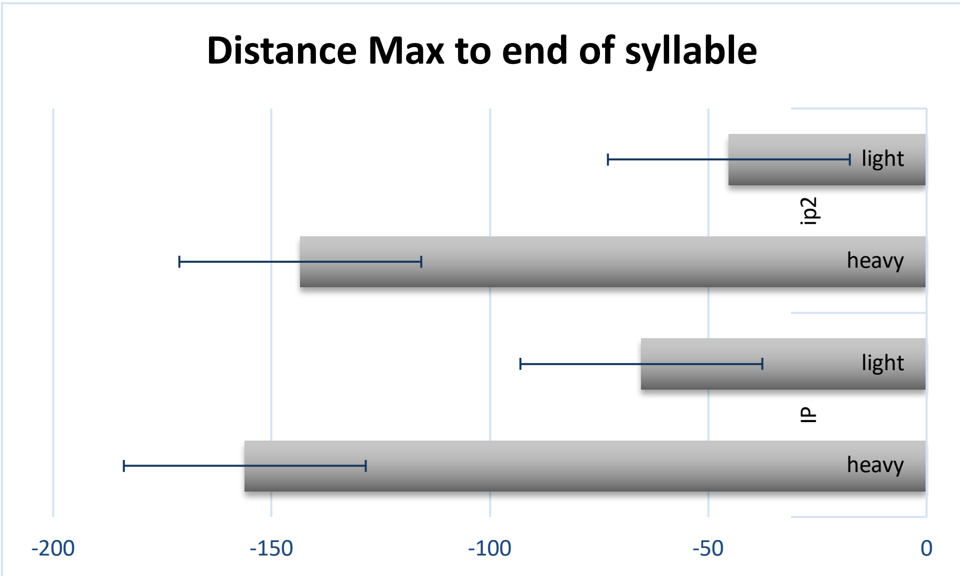


Figure 75: Peak distance relative to syllable offset according to syllable weight. IP= intonational phrase, ip2= intermediate phrase.

Second, the results of the alignment patterns relative to the end of the vowel are reported. The table and graph report results according to prosodic boundary, syllable weight and relative to the end of the target vowel.

Boundary	*maxToEndOFVowelMs*	*Std. Deviation*
IP		
heavy	-109.00	55.00
light	-65.31	33.35

ip		
heavy	-91.00	64.00
light	-45.27	23.16

Table 95: Average distance in ms between peak location and target vowel offset.

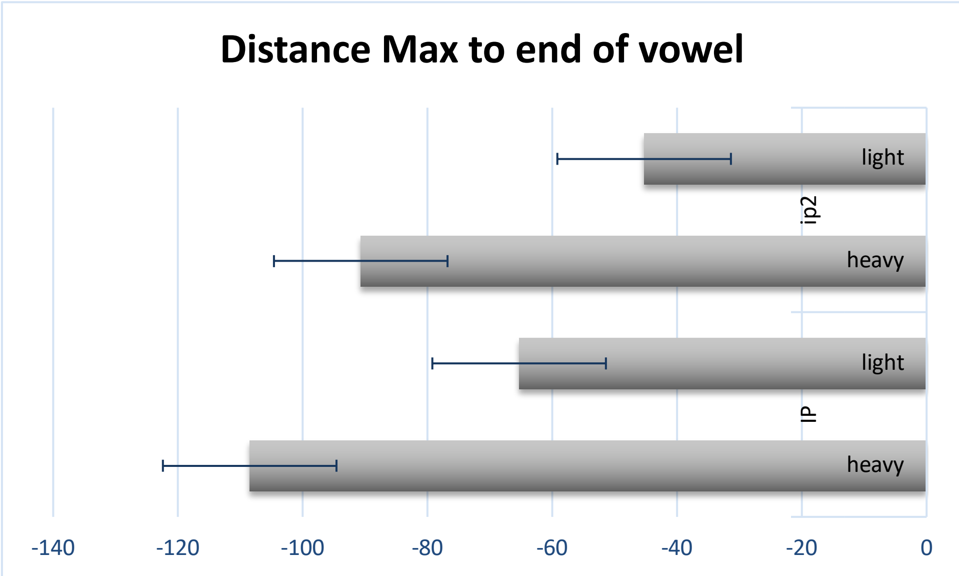


Figure 76: Peak distance relative to vowel offset according to syllable weight. IP= intonational phrase, ip2= intermediate phrase.

The means relative to syllable offset (maxToEndOFSyllableMs) and vowel offset (maxToEndOFVowelMs) show that overall, the peak in ip vicinity takes a shorter distance to align relative to the end of the syllable/vowel when it is light (i.e. is aligned later in the domain), while it takes longer time to align in heavy syllables (i.e. is aligned earlier in the domain).

Based on the results in this section, it can be said that in both prosodic domains, peak location does vary according to syllable structure in this Arabic dialect. That peak location does vary as a function of syllable duration and type. Closed syllables, heavy syllables and long vowels cause the peak to be aligned later relative to the offset target syllable. In this respect, comparing across tables [97] and [99], and across tables [101] and [103] it can also be observed that in the vicinity of ips and across vowel/syllable types, the peak takes a shorter amount of time to locate from the end of the rhyme. In IPs on the other hand, it takes a longer amount of time for the peak to locate away from the end of the rhyme. These patterns respectively indicate late and early alignment that can be explained relative to the rhyme duration. Recall in chapter [4], it was reported that ip boundaries cause significant lengthening of units, moreso than in IP boundaries. Based on this and since there is found to be a correlation between syllable duration and peak location in the corpus, it can be said that the later realisation of the peak coincides with the longer durations of syllables in the ip boundary. Conversely, the less lengthening of the units observed in IP boundaries results in earlier alignment of the peak. This conclusion comes in accord with the results reported for ips and IPs in table [75].

6.6 Correlation and association results

This step in the analysis is intended to further explore the possible correlation between accent duration and syllable duration. Having established an association between syllable duration and peak location as a function of the many contextual factors, it is now needed to establish the relationship between the tonal gesture and the syllable duration. That is, to establish whether or not tonal gestures are static and fixed in duration in this Arabic dialect. Recall the theory maintains a free variation in accent duration according to syllable duration, which comes in contrast to the other hypotheses claiming a fixed duration for accents (section 0). Pearson correlation reports a positive correlation between the two in this corpus ($p > 0.05$, $r = 0.38$). This correlation result confirms that the pitch movement is associated with the stressed syllable and varies in duration according to its internal makeup. This correlation along with the correlation between peak location and syllable duration in the previous section, supports a SAH analysis and disfavours claims of a fixed duration of tonal configurations (ladd et al., 1999, 2009, Prieto et al., 1995). Note the following graph:

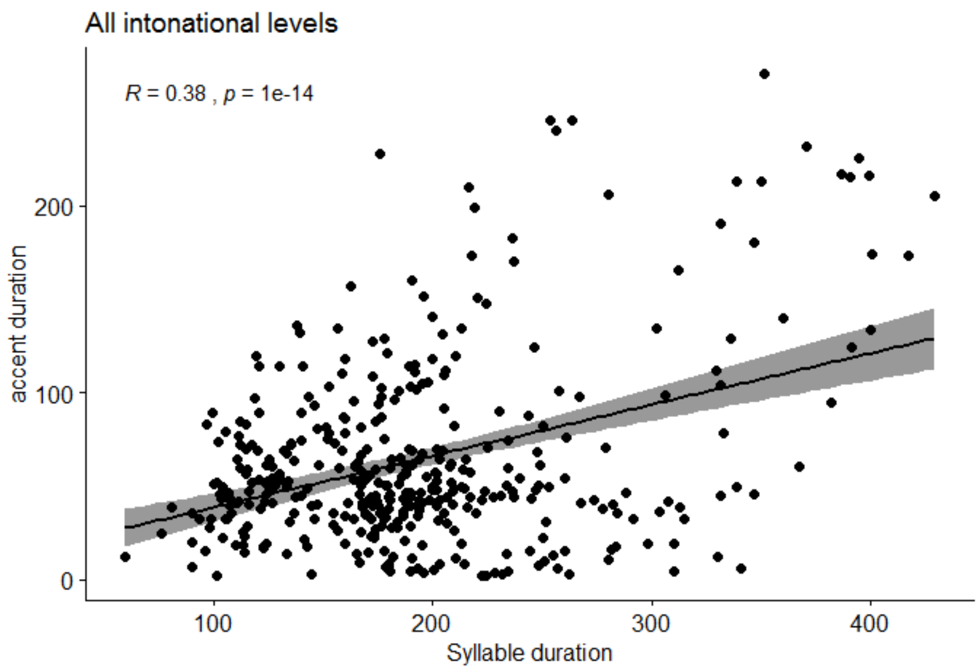


Figure 77: Scatter plot for the correlation between accent duration and syllable duration

6.7 Stress clash

Now it is time to observe the tonal alignment patterns when the effects of upcoming prosodic boundary are removed. These cases constitute the phrase-medial PW condition. In the following are the results obtained in order to observe how the dialect resolves stress clash of successive accents. To maintain comparable results to the published literature, peak location under the stress clash condition is measured relative to the onset of the syllable and in correlation with the syllable duration. The contexts relevant for the stress clash condition in the published literature are the zero and two intervening syllable cases. The zero intervening case would potentially demonstrate the local effects of the clash, as no syllables intervene between the two accents, while the two intervening syllables case is thought to eliminate the pressures observed in the zero intervening syllables case (Chahal, 2001, though not for Greek in Arvaniti et al. 1998).

Recall that each target word was placed preceding a control word varying in the location of the accent depending on stress location. Correlation tests were carried out to observe whether the differences in H alignment relative to onset, syllable duration and as a function of the number of intervening syllables. In the previous section it was seen that there is significant correlation between the location of the peak and the duration of the rhyme. In this section the segmental anchor will be the onset of the syllable and we will observe whether or not there is a correlation between peak location relative to onset and syllable duration. The following results and graphs show for each word the location of the observed tone relative to the onset of the syllable and according to the number of intervening syllables in between. First in the analysis is the overall comparison by the number of intervening syllables in the clash. Subsequently, the analysis is carried out on each stress location and syllable type individually, in order to eliminate any overriding effects of context.

Clash distance by no. of syllables	*SyllabledurationMS*	*distanceMaxSyllableToOnsetMs*
zero	220.21	88
one	181.12	88.45
two	163	89
three	142.41	84.47
four	102	71

Table 96: Results split by presence (zero) and absence (2-4) of stress clash.

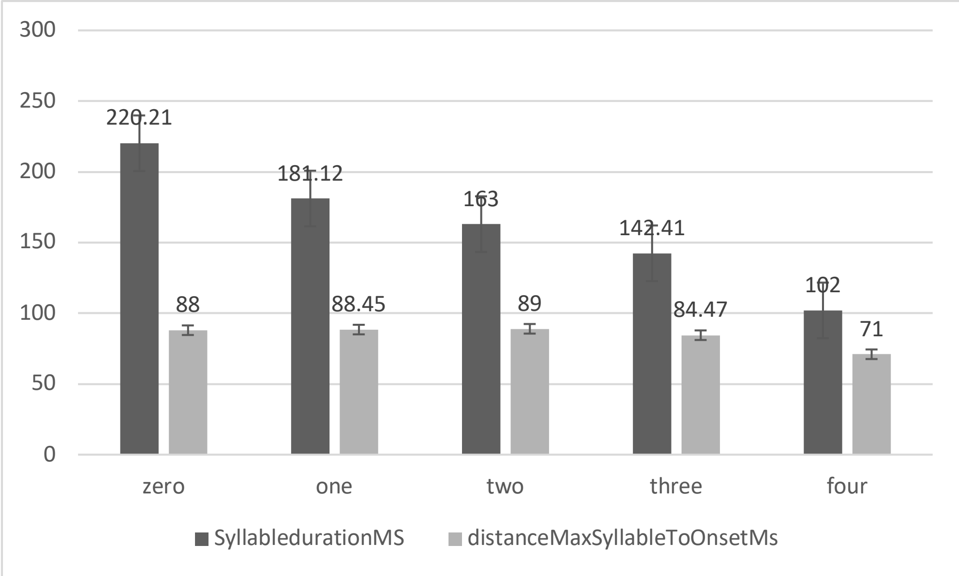


Figure 78: syllable duration and peak distance from onset according to the number of intervening syllables between accents.

1. /**la.ha.fu** riz.gu/, /**la.ha.fu** ma.kaa.ni/ and /**la.ha.fu** Ta.la.baat/.

The tables and graphs below show a difference in peak location according to syllable duration in this target word with initial stress. Moreover, syllable lengthening varies in degree according to the number of intervening syllables for this word. As the number of intervening syllables decreases, syllable duration increases, and as a result peak location is delayed relative to the onset. The results here are all significant ($p>0.05$). Thus, is could be said that stress clash resolution method in the dialect is lengthening the respective syllable. There is significant correlation between peak location and syllable duration.

	lahafu		
	mean	std. deviation	Syllable duration
distanceMaxSyllableToOnsetMs (four intervening syllables)	70.6	28.8	101.89
distanceMaxSyllableToOnsetMs (three intervening syllables)	74.7	25.7	102.00
distanceMaxSyllableToOnsetMs (two intervening syllables)	75.9	26.3	103.37

Table 97: Mean durations for distance of peak to onset in /lahafu/ as a result of three types of intervening syllables

		maxtoonset	syllableduration
maxtoonset	Pearson Correlation	1	.548**
	Sig. (2-tailed)		.000
	N	160	160
syllableduration	Pearson Correlation	.548**	1
	Sig. (2-tailed)	.000	
	N	160	160

**. Correlation is significant at the 0.01 level (2-tailed).

Table 98: /lahafu/ stress clash correlation results

2. /**riH.lat.hum** Hil.wa/, /**riH.lat.hum** ga.sii.ra/ and /**riH.lat.hum** li.lib.naan/.

This word has medial stress. Syllable duration is the longest in close proximity to another accent (1 intervening syllable). This lengthening also entails later peak alignment. The results here are significant (p=0.001).

	riHlathum		
	mean	std. deviation	Syllable duration
distanceMaxSyllableToOnsetMs (one intervening syllable)	78.3	46.9	169.00
distanceMaxSyllableToOnsetMs (three intervening syllables)	70.2	39.7	165.39
distanceMaxSyllableToOnsetMs (two intervening syllables)	65.1	42.0	163.86

Table 99: Mean durations for distance of peak to onset in /riHlathum/ as a result of three types of intervening syllables

		maxtoonset	Syllableduration
maxtoonset	Pearson Correlation	1	.401**
	Sig. (2-tailed)		.000
	N	169	169
Syllableduration	Pearson Correlation	.401**	1
	Sig. (2-tailed)	.000	
	N	169	169

**. Correlation is significant at the 0.01 level (2-tailed).

Table 100: /riHlathum/ stress clash correlation results

3. /**mu.laa.zim** taa.ni/, /**mu.laa.zim** taH.giig/ and /**mu.laa.zim** taH.gii.gaat/.

The word also shows similar patterns according to syllable duration. Longer syllable durations in close proximity (1 intervening syllable) entail later alignment of the peak. The results are all significant (p=0.01).

	mulaazim		
	mean	std. deviation	Syllable duration
distanceMaxSyllableToOnsetMs (one intervening syllable)	100	53.1	162.66
distanceMaxSyllableToOnsetMs (three intervening syllable)	99.7	51.2	157.76
distanceMaxSyllableToOnsetMs (two intervening syllables)	98.1	57.4	155.40

Table 101: Mean durations for distance of peak to onset in /mulaazim/ as a result of three types of intervening syllables

		maxtoonset	Syllableduration
maxtoonset	Pearson Correlation	1	.602**
	Sig. (2-tailed)		.000
	N	170	170
Syllableduration	Pearson Correlation	.602**	1
	Sig. (2-tailed)	.000	
	N	170	170

**. Correlation is significant at the 0.01 level (2-tailed).

Table 102: /mulaazim/ stress clash correlation results

4. **/lil.Haf.laat** 2aS.lan/, /**lil.Haf.laat** 2a.maa.kin/ and /**lil.Haf.laat** fil.gu.Suur/.

Finally, this is a word with final stress that leaves no intervening syllables until the upcoming accented syllable if the following word starts with initial stress. The other two cases are separated by one or two syllables. In this word too, the syllable duration is correlated with the peak location as expected, with longer durations entailing later alignment. And the longest duration is observed in the absolute proximity case (zero intervening syllables). The results here are all significant (p=0.00).

	lilhaflaat		
	mean	std. deviation	Syllable duration
distanceMaxSyllableToOnsetMs (one intervening syllable)	86.2	78.5	211.27

distanceMaxSyllableToOnsetMs (two intervening syllables)	80.1	60.7		217.00
distanceMaxSyllableToOnsetMs (zero intervening syllables)	88.00	71.2		220.21

Table 103: Mean durations for distance of peak to onset in /lilHaflaat/ as a result of three types of intervening syllables

		Maxtoonset	Syllableduration
Maxtoonset	Pearson Correlation	1	.767**
	Sig. (2-tailed)		.000
	N	174	174
Syllableduration	Pearson Correlation	.767**	1
	Sig. (2-tailed)	.000	
	N	174	174

**. Correlation is significant at the 0.01 level (2-tailed).

Table 104: /lilHaflaat/ stress clash effect results

6.7.1 stress clash discussion

As can be inferred from the previous results, peak alignment to the onset does vary according to the number of intervening syllables between accents, while staying true to the respective TBU (stressed syllable). Negative values in alignment would otherwise indicate alignment before the onset of the target syllable. Moreover, statistical results report that it does not matter if the accents were absolutely close (zero intervening syllables), or were as far as four syllables apart, the peak still aligns with respect to the onset of the target syllable. More importantly, the results point towards the fact that in JA stress clash triggers lengthening of the target syllable and subsequent peak modification. However, stress clash in Jeddah Arabic does not trigger early peak alignment, as tonal alignment seems to mainly interact with syllable duration. By lengthening the target syllable, speakers are able to realise the peak accordingly within. These JA results appear to differ from the findings in Lebanese Arabic in Chahal (2001) where stress clash triggered early peak alignment along with increased syllable duration in the zero intervening syllables case, and a later alignment accompanied with a decreased syllable duration in all the other cases. For Egyptian Arabic, Hellmuth (2006) found peaks to align on or before the onset of the syllable in stress clash contexts, however no investigation was carried out on the following clash regarding target syllable duration.

6.8 Intonational levels and alignment of tones

Recall that earlier results report that IPs exert a leftward push of the peak in their vicinity compared to ips and this was taken as evidence for an effect of prosodic boundary on tonal alignment. The final step in the analysis is to evaluate this prosodic effect taking into consideration the speaker, gender and the other contextual factors reported by the previous analyses in a unified model. The analysis below is set out to investigate which of the two intonational levels has the most effect on the push of the peak location and relative to which segmental landmark. It is carried out combining all the absolute variables for the alignment of H (MaxTovowel (v1), MaxToOnset (v2), maxToEndOFVowel (v3), and maxToEndOFSyllable (v4).

For this investigation, a mixed effects regression analysis of the relationship between those was carried out. In the model: the Outcome is the binomial prosodic levels: IP/ip. Variables (v1-4) + gender, repetition + are fixed effects, while speaker + word are random effects. Visual inspection of residual plots did not reveal any obvious deviations from homoscedasticity or normality.

Two models were tested for this enquiry. Model-1 included all of the aforementioned variables. However, values regarding the variables (v1) and (v2) were quite high. The correlation between the variable (v1) and (v2) were quite high and significant (corr=0.97, p=2.2e-16). Therefore, (v1) was removed from the model. Model-2 was used to account for the difference between the levels. According to Model-2, the only significant variable is (v2) (p=2.59e-07). The variables (v3) and (v4) are not as significant (v3: p=0.29, v4: p=0.85). According to this, each one-unit change in (v2) will decrease the log odds in favour of (IP) by (0.0151). Thus it can be concluded that indeed IPs cause the peak to significantly retract relative to the onset of the syllable in comparison to the peak in the vicinity of ips. This result is thus taken as evidence for an effect of prosodic boundary on peak location in the current study.

MODEL-2

Deviance Residuals:					
	Min	1Q	Median	3Q	Max
	-1.5834	-1.1732	-0.3078	1.0936	2.2424
Coefficients					
	Estimate	Std. Error	z value	Pr (>|z|)	VIF
(Intercept)	1.3158492	0.3836737	3.430	0.000604***	
MaxToOnset (v2)	-0.0151590	0.0029430	-5.151	2.59e-07***	1.876490
maxToEndOFVowel (v3)	0.0036328	0.0034336	1.058	0.290051	3.775031
maxToEndOFSyllable (v4)	0.0004269	0.0023227	0.184	0.854186	2.809426
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					
Null deviance=629.24, Residual deviance=582.83, AIC=590.83					

Table 105: Regression results for peak alignment

6.9 Chapter discussion and conclusion

The data in this corpus show that in Jeddah Arabic, the stressed syllable is a true segmental anchor for tone alignment. As was presented, across structures, words, intonational levels, syllable types, and conditions, the tones consistently aligned with the target accented syllable. Additionally, when peak alignment was used as a variable to distinguish between intonational levels, it was pushed as leftward as permitted within the target syllable (recall schematic representations and figures). This leads to the conclusion that the syllable is the *Tone Bearing Unit* in JA.

A clear distinction in terms of peak alignment patterns was observed in the data between intermediate phrases, and Intonational phrases. This behaviour serves as further evidence for the existence of intonational levels above the prosodic word, as was discussed in chapter [4]. In IPs, across long-short vowels, open-closed syllables, and across genders, the H was pushed leftward to the onset of the target syllable, leaving L to align on or just before the onset consonant. In ips on the other hand, the L consistently aligns with the onset consonant, while the H aligns within the confines of the vowel. Other segmental landmarks show collaborating patterns as well. The variables maxtoendofvowel and maxtoendofsyllable demonstrate consistent patterns according to intonational level. In ips, H takes shorter distance (in ms) from the end of the vowel aligning approximately towards the end of the rhyme, while in IPs, it takes a much longer time aligning earlier to the vowel.

An additional supporting evidence is the stress clash condition, which shows contrary behaviour to the languages reviewed in section 6.2. The data demonstrate a resistance to tone alignment outside the syllable, even with the time pressure of an adjacent accent. The respective tones always aligned within the syllable boundaries regardless of how many intervening syllables occurred between the target accent and a following accent. The duration of the target syllables constitutes the the main factor that varies according to the number of intervening syllables. This lengthening effect triggered later peak alignment in Jeddah Arabic.

As the duration of the syllable plays a large role in tonal alignment in this dialect, it was important to observe how it interacts with the the tonal targets L, and H. Tests for correlation reveal that the data does support a correlation between peak location, accent duration (the time between L and H in L+H accents) and syllable duration. The results suggest that in this dialect both the peak location varies according to syllable length, and the time it takes to transition from L to H varies according to syllable duration. While the duration is variable, the targets are fixed to the contents of the stressed syllable (Atterer & Ladd, 2004, Ladd at al., 1999, Arvaniti et al., 1998).

Chapter 7. Discussion, thesis contribution and future research

7.1 Introduction

The thesis provided an analysis for Jeddah Arabic intonation following an Autosegmental-Metrical approach. Properties of the suprasegmental structure for the dialect have been presented alongside detailed phonetic and phonological analyses. This chapter summarises all the findings in the previous chapters, the contributions of this thesis as well as some suggestions for future research.

7.2 Findings and implications

Chapters one and two provided a discussion of the theoretical motivation behind the analyses as well as the elements in the AM theory. The theory combines a phonological component in the form of both Metrical and Autosegmental information, as well as a phonetic component concerning the surface realisation of linguistically relevant F0 events. Under this theory there is variation in how languages express their intonational structure in addition to variation in which acoustic cues they employ in the expression of this structure. In this regard, experimental evidence from this dialect was in favour of classifying it as a stress- accent language. Under Beckman's (1986) definition, stress- accent languages generally rely on non-tonal correlates to mark prominence in the lexical level, and both tonal and non-tonal correlates are used in the postlexical level to express intonation, which involves prominence differences between some words relative to others in an utterance. This makes JA similar to English, Lebanese Arabic, and Egyptian Arabic; languages that express intonation via the alternation of prominent and non-prominent intervals with a tendency for prominent units to occur on the right edge of a domain. In such languages, the rhythmic alternation of units is said to be right-headed, whereby the stronger unit tends to occur on the end of utterances, usually followed by a juncture.

Chapter three presented details of the methodology used to extract the needed information from the dialect. The corpus was designed to include both controlled and semi-controlled data from a representative sample of speakers. ToBI style annotation was carried out on the corpus following the broad patterns reported in previous studies on Arabic dialects, and this would subsequently aid in a cross-linguistic comparison across the Arabic language spectrum. Alongside, special Praat scripts were used to extract the essential acoustic information for each experiment in the respective chapters. The decisions made throughout the process of corpus design stress the importance of choosing the correct methodological tools to analyse intonation.

Chapter four was concerned with the intonational categories observed in the dialect. These are in the form of accent types, different tunes/melodies in JA, and the prosodic components in the form of phrasing levels. It was observed that the dialect demonstrates a tonal inventory that is comprised of four pitch accents: H*, !H*, L+H*, L* which are phonotactically constrained to associate and align with lexically stressed syllables. The dialect also presents evidence for three phrase accents of the shape H-, L-, and !H-, and boundary tones of the shape H% and L% whose combinations yield different melodies. The composition of melodies in JA showed both cross-linguistic and cross-dialectal similarities and differences, an aspect that highlights universal and language-specific trends in intonation expression. The differences were mainly observed in regard to the range of intonation expression. There were cases of intonational patterns shared by two different pragmatic meanings in JA. Such cases would perhaps benefit more from further prosodic analyses, in addition to analyses that go beyond the widely accepted primitives. For example, it was seen in the analysis that different functions did not entail differences in intonation according to the nuclear contours alone. The request and imperative tunes included in this study would thus benefit from a focused study on the upper and lower pitch range trend lines observed in longer stretches of material, as they may then be distinguishable from their counterpart melodies in question and declarative tunes. Indeed, it is undeniable that pitch range would have an effect on intonation expression, however it was noted that semantic word choice in those cases matters too. Therefore, an investigation is needed that incorporates the other tools speakers resort to when their language has exhausted the possible range of intonation expression.

Chapter four also proposed a hierarchy of the prosodic constituents in the dialect. The observed units were Prosodic words PWs, intermediate phrases ips, and Intonational phrases IPs. The major boundaries are marked by a combination of tonal and non-tonal correlates. The phrase accents mark the right edge of intermediate phrases along with pre-boundary lengthening of the stressed syllable and vowel and tonal alignment, while boundary tones mark the right edge of Intonational phrases along with pre-boundary lengthening of the stressed syllable and vowel and early tonal alignment. The reported correlates in JA demonstrate a cross-dialectal variation in regards to the patterns a dialect demonstrates in order to differentiate those domains. Similar to a typologically- unrelated language like Spanish and contrary to a typologically- related language such as Lebanese Arabic for example, Jeddah Arabic was found to use greater pre-boundary lengthening of the target stressed syllable and vowel to mark intermediate phrases that is greater in degree compared to intonational phrases. It is worth noting that while peak retraction patterns in Jeddah Arabic are similar to the reported results in Lebanese, JA shows a contrary pre-boundary lengthening pattern to the one reported for Lebanese. This may be taken as evidence for a diversity in the implementation of prosodic features among languages in general and among dialects of the same language. Future research in the form of a comparative analysis would be beneficial in confirming whether this lengthening pattern in JA is a gradient attribute observed among Arabic dialects, whether intermediate phrases are stronger than Intonational phrases in a group of dialects but not others, or whether JA and similar dialects makes use of a further prosodic level not reported in Lebanese nor Egyptian Arabic. Of course, such studies would ultimately start by elucidating the difference between ips with and without pauses and the duration of those pauses in such analysis. Additionally, in this thesis, phrasing and relative prominence relationships among the number of accented words in an intonational phrase was analysed and reports the choice of one nuclear accent to head the most prominent ip. Jeddah Arabic as we have seen from prominence distribution according to focus structure in medium sized sentences, dictates that one prominent accent heads the ip, and according to sentential position, uses phrasing and gradient pitch range manipulation as tools to convey this hierarchical organisation. However, further investigation is needed on what unit constitutes the head of the Intonational phrase itself, as there are cross-linguistic reports of the existence of parallel relative prominence relationships among the nuclear accented words in an extended IP that collaborate in order to designate one as the head of that major prosodic constituent. Moreover, and in longer stretches of sentences, it would be interesting to observe the relationships among the many IPs in that respective string and how accents are distributed among them. This would consequently show whether or not those IPs in Arabic collaborate to make up a higher prosodic constituent such as the Utterance.

Finally, chapter four additionally presents details regarding the tonal implementation rules in the dialect that make up the surface F0 contour, these include: Interpolation, Downstep, Upstep, and Phonetic declination. Facts from the dialect support the notions assumed in AM theory (Pierrehumbert, 1980, Pierrehumbert & Beckman: 1986) regarding the realisation of the F0 contour as a sequence of H and L phonological targets whose surface realisation is a result of phonetic implementation. Indeed phonetic observations from the corpus show the variable realisations for these targets according to speaker pitch range and relative to other targets in the same F0 contour, while being faithful in associating and aligning with metrical structure. The flexibility captured by the notion of interpolation allows for the expression of this variability without the risk of disregarding phonetic reality nor phonological structure. In addition, the observations from focus structure and alignment patterns reviewed in detail below, point to the existence of a relative hierarchy of prominence essential to the expression of intonation as assumed in the theory. This hierarchy is observed from pitch distribution patterns, phonetic alignment patterns and edge-marking pitch configurations, within and across prosodic domains. Such findings support the interpretation of a substantial link between phonetic reality and phonological structure in the expression of intonation, as well as constitute cross-linguistic support for the validity of these structures.

Chapter five is focused on the prominence marking and prominence hierarchy observed in the dialect as conveyed by information structure. For this, a dialogue was elicited to observe the quantitative and qualitative details of the target word when it occurs under narrow focus, or broad focus alongside the interaction with sentential position and gender. The facts from focus realisation indicate that in Jeddah Arabic, cumulative quantitative detail is systematically used to signal the prominence status of a word. It was found that when a word occurs under narrow focus, its phonetic properties are enhanced compared to the rest of the string and compared to when it is under broad focus. The target is longer in duration, demonstrates a larger excursion size, peak height, intensity and rise speed. It was also reported that Jeddah Arabic realises the rest of the string in much smaller pitch excursions thus displaying Post-Focal compression. Indeed, the phonetic composition of an utterance can be said to matter in prosodic expression in JA. The observed pitch manipulations among words in the same stringin JA conspire to yield a specific focus type is parallel to what has been observed for English and Chinese in Xu and Xu (2005), Egyptian in Hellmuth (2005, 2006a, b) and Lebanese in Chahal (2001), among others. Pitch values in JA were enhanced according to sentential position and focus type. Peak F0 Hz/ST, range F0 Hz/ST, and the rise speed of tones were significantly higher in the nuclear accented word in initial and medial positions under narrow focus, under broad focus, and between narrow and broad focus utterances. Similarly reported pitch manipulations led Xu and Xu (2005) to conclude that speakers consciously adjust pitch range as an effect of focus, thus highlighting the word under focus and compressing the post focal material. Not only pitch manipulations were observed as a result of focus prominence in JA, other acoustic measurements also showed enhanced qualities under narrow focus and in comparison to broad focus. Among them are the durational measurements that were significant in the differentiation between the two focus types and in the realisation of the nuclear target in a given phrase. More importantly, the durational measurements are significant in the realisation of the nuclear accent in final position. While the tonal measurements were not as helpful in signalling the nuclear accent in this position, vowel duration, and syllable duration constitute pivotal markers of words before major boundaries. Individually in narrow and broad focus utterances where /laa.na/ is final, and across those focus types, the duration of units are greater in comparison to initial and medial positions. This behaviour mirrors the reported patterns of pre-boundary lengthening before major prosodic boundaries in JA. Units at the right edge of a boundary are lengthened to signal the rhythmic alternation maintained in the dialect. More generally, the findings from focus realisation constitute evidence that the dialect maintains a hierarchical organisation of accents within a contour thus regarding nuclear accents as the most prominent accents in a prosodic phrase as assumed in the AM theory. In terms of qualitative patterns, it was reported that this hierarchy affects within utterance accent distribution by modifying phrasing and post focal material. Phrasing was used as an occasional pattern that places the narrow focus target in the dominant intonational phrase making it the nuclear accent of that phrase. Though phrasing was used as a method to signal narrow focus in the current corpus, it cannot be said to systematically convey this type of focus. As was reported in table [33], there is a great deal of within and between-speaker variation regarding the use of phrasing. The cited experimental studies on focus report that their speakers vary regarding the use of qualitative measures with some measures being more systematic than others, e.g, deaccentuation, or focal accent type as qualitative measures were found to be more systematic in languages than other measures. On the other hand, variation in the quantitative measures is not commonly reported, thus speakers would be unified regarding increasing syllable duration under focus for example. Hellmuth's analysis for instance found an L- phrase tone to "occasionally" mark narrow focus utterances, hence she argues that this measure is not to be taken as a systematic reflex of focus- while excursion size/ syllable range was reported as a more systematic way to mark focus in her data. This preference and variation could potentially be a matter of languages- or speakers- giving increased priority to some qualitative measures over others or giving more priority to quantitative measures to increase prominence than qualitative measures. Predictive and evaluative statistical processing models parallel to the method reported in Baumann and Winter (2018) would bring more insight into those speaker or language preferences. Overall findings regarding focus support the existence of a prominence hierarchy that places higher phonetic value on nuclear accents as was presented earlier, compared to unaccented and pre- nuclear pitch accents. All these properties are used to deliberately highlight the nuclear accented word against the rest of the utterance.

As can be inferred from table [106] below, Arabic dialects share numerous similarities regarding the effect of accentual prominence on their intonational expression. The use of duration seems to be a significant feature that unifies Arabic dialects with respect to the expression of prominence. This is to be expected since it is reported that as a speech planning strategy, any phonetic feature including suprasegmentals would need sufficient space for it to be realised (Krivokapic, 2012). Those durational effects are also reported to be correlated with prosodic structure, since prosodic phrasing levels involve junctures and pauses of varying lengths. The durations of those pauses are taken as indicators of planning strategies for the upcoming syntactically affiliated units; speakers tend to lengthen units and insert pauses while planning the articulatory and acoustic properties of the upcoming unit(s) (Krivokapic, 2012, Rao, 2010). Indeed, often the reviewed Arabic dialects- including the findings of JA in this thesis- use embedded intermediate phrases to a) mark the focused target with a prominent nuclear accent that involves substantial lengthening and enhanced tonal detail, and b) indicate that this unit is incomplete and constitutes part of the meaning of the major Intonational phrase. As a result, the following phrase is planned to be uttered with compressed pitch range to demonstrate the prominence of the nuclear accented word. In essence, this behaviour would constitute an insight into the reported planning strategy for how prosodic structure is implemented based on the given syntactic information (Keating & Shattuck-Hufnagel, 2002, for example). In Keating and Shattuck-Hafnagel's (2002) model they suggest that prosodic structure is available prior to the level of phonological and phonetic encoding, even before the segments are encoded. That there is a "skeletal default prosody"; a prosodic template that is built on the given syntactic information (p.139). In their view, this level of representation includes information about the phrasing constituents that are influenced by the syntactic length of the preceding and following material, the relative prominence of the units, basic information on the pitch accents and information on the edge tones. Subsequently, this prosodic information feeds the phonetic encoding processes that translate into the surface temporal, acoustic and tonal phenomena, and the skeletal prosody pattern is reconstructed along the way. The reconstruction process also takes into account the external factors to the discourse, such as turn-taking, speech errors, etc. Such a model seems to constitute a step further in the analysis of intonation as it involves the many aspects essential to the planning and production of prosody. Indeed, further investigation into the adoption of these processes is needed in Arabic.

Finally, Arabic dialects in the table below also seem to all be unified in the disfavour of using specific accent types to mark focus, thus suggesting that they would be exploiting to a greater degree the other prosodic means to signal prominence. This may also reflect the wealth of accent types in these Arabic dialects. In regards to tonal means of prominence expression, Arabic dialects show a preference for enhancing the pitch range of a unit to signal prominence. This comes in accord with a durational expansion of the target syllable where this expanded range is realised. In addition to this, though primarily, Jeddah Arabic showed that the rise speed of the tones under this expanded range is faster. And since the alignment patterns of the dialect show a consistent anchor for the L at the onset of a syllable, the transition from this L to the H peak is faster in nuclear accents, as a way to enhance this accent type under focus (perhaps as a chosen alternative to using categorical accent type to mark focus as in some languages). It would be thus interesting to see how this rise speed feature is implemented in the other Arabic dialects as a function of accentual prominence.

In chapter five, only the melodic and dynamic correlates of pitch measurements, duration and intensity were analysed as influenced by focus in Jeddah Arabic. The analysis would therefore benefit from the inclusion of further acoustic measurements, such as formant frequencies and spectral balance that have reported significant results in other Arabic dialects. Alongside this, an expansion of the analysed vowel inventory would be needed to statistically support the effects of focus on the formant frequency measures. Lastly in this regard, an elaborate study on the implications of post focal material relative to the information structure they convey is needed. In the dialect it was seen that the question- answer dialogue included post focal material that is given or understood along the course of the dialogue. This material was realised with a compressed pitch range as was seen. While an advantage of the question-answer paradigm is demonstrating the one item that is contrastively focused at a time, it still does not tell whether the compression of the rest of the material is because it appears following the focused constituent as a systemic reflex of off-focus, or if this compression is due to the information being given and repeated in the dialogue. The experimental analysis in Hellmuth (2011) for example, found that in controlled Verb-Object-Object sentences, the pitch range of the indirect object was compressed as a result of its position following the focused item 'off-focus', but not as a result of it being given in the sentence since given material are not deaccented in EA. Therefore, a study on the prosodic encoding of givenness in JA is needed as a complement to the findings of the prosodic encoding of focus.

In the following is a comparison table of the different variables used to mark accentual prominence across languages and across Arabic dialects. A blank cell indicates that the correlate has not been studied in the language; a ✓indicate that the correlate is studied and reported for the language, while an x indicates that it is studied and not found for the language[15]:

Language	Duration	Peak F0	Excursion size	Intensity	Rise time/speed	Deccentuation	Accent type	phrasing
1. LA	✓	✓	✓	✓		✓	x	✓
2. TA	✓	✓	✓	✓		✓	x	
3. EA	✓		✓	✓			x	
4. MA	✓	✓	✓			✓	x	x
5. KA	✓	✓	✓			✓	x	✓
6. YA	✓	✓	✓			✓	x	x
7. TaifA	✓	✓	✓	✓		✓	x	
8. AEng	✓	✓	✓		Both	✓	x	x
9. Chinese	✓	✓	✓		Both	✓	x	x
10. Spanish	✓	✓			Rise duration		✓	✓
11. JedA	✓	✓	✓	✓	Rise speed	✓	x	✓

Table 106: Cross-linguistic focus marking strategies.

Chapter six presented details from the alignment patterns in Jeddah Arabic. Alignment was investigated to determine the Tone Bearing Unit TBU in the dialect, and similar to previous studies, alignment behaviour is taken as a marker for the different prosodic constituents. For this, extensive phonetic analyses were carried out on the target under a number of cases: when it is in the vicinity of different prosodic boundaries, according to the lexical stress location and syllable type, and according to the distance between the accent and a following accent (stress clash). Here the plan was to observe peak location in accents occurring in proximity to prosodic constituents, and as a result of different syllable structures and distances between each other. In these cases the peak timing was observed relative to different points in the target syllable: the onset of the syllable, the onset of the vowel, the end of the syllable and the end of the vowel. It was found that under all these cases, and regardless of distances between accents and syllable types, the peaks and valleys were faithful to their target syllable; L aligns at the onset of the accented syllable or just before, while H aligns within the rhyme of the syllable. The fact that peaks and valleys consistently aligned with the lexically stressed syllable, means that the syllable is the TBU in the dialect. This consistent behaviour was taken as evidence for the relationship between underlying and surface structure assumed in AM theory. The current

[15] 1. Lebanese Arabic as analysed in Chahal (2001, 2014). 2. Tunisian Arabic as analysed in Bouchhioua (2008). 3. Egyptian Arabic as analysed in Hellmuth et al. (2006, 2014). 4,5 &6 Moroccan, Kuwaiti and Yemeni Arabic as analysed in Yeou et al. (2007). 7. Taif Arabic as analysed in Al-Zaidi (2014). 8 & 9 American English and Chinese as analysed in Xu and Sun (2002), Xu and Xu (1999, 2005). 10. Spanish as analysed in Face (2002). 11. Jeddah Arabic as presented in this thesis.

thesis therefore maintains a phonological association of tones with heads of metrical/prosodic constituents that results in an observed timing coordination (alignment) between them on the surface.

Regarding proximity to prosodic boundaries, this was a case that posed a re-location of peak. The peak was pushed as leftward as possible in the target syllable in the vicinity of an IP boundary, less so in the vicinity of an ip boundary. This was taken as evidence for the existence and marking of these consitituents in JA, as well as the syllable being a TBU since peaks were never observed outside[16]. Given that the syllable is the TBU in the dialect, it is expected to be acoustically and perceptually enhanced when accented and relatively different from unaccented syllables. More generally, the finding that alignment was to an extent uninfluenced by the different experimental manipulations in this thesis supports the segmental anchoring hypothesis discussed in section 6.2 (Atterer & Ladd, 2004, Ladd at al., 1999, Arvaniti et al., 1998). Indeed, the tones were shown to be consistently aligned with the head of the prosodic word, thus showing where the dialect categorically stands within cross-linguistic prosodic typology regarding tonal alignment.

The other causes for alignment modification in JA were reported according to the contextual segmental and syllable structure effects. There was found to be a variation in peak location mainly as an effect of syllable duration. The target accent's duration was varied according to the duration of the syllable it occurs in. Furthermore, in the stress clash resolution patterns, the relocation of the peak was mainly in response to syllable duration. These findings therefore do not support the analysis of a fixed duration of tonal configurations. It must be noted that the results and analyses of stress clash resolution in the current thesis were carried out on the pre-nuclear target word as manipulated in distances from the following material. The study thus did not introduce results on the realisation of the following word in this case. The tonal alignment, syllable duration and other acoustic correlates of the following word would be interesting cases to investigate. According to the patterns observed in nuclear accents at IP boundaries in JA, those would demonstrate shorter durations and earlier alignment compared to the previous words in stress clash contexts. Finally in this regard, an expansion of the segmental anchors in a word would be more insightful in the further analysis of the exact alignment location within a stressed foot. The analysis in this chapter does not control for word internal foot count or structure. Controlled studies in this regard were able to report the exact

[16] In rising accents: L+H*, Peaks were always within the nucleus, while valleys were pushed outside onset if peak was pushed into onset in the vicinity of an IP.

level a tone associates with in the moraic structure. At the moment evidence points toward the tone preferring to be aligned with the head mora of the foot. Future analyses that include left and right boundaries of the prosodic word as anchors, in addition to controlling foot structure and duration may reveal further interesting results.

In general, it can be said that Jeddah Arabic belongs to the group of languages in prosodic typology where a pitch event is used postlexically to mark prominence as well as boundaries of specified prosodic constituents above the word (Jun, 2005, 2014). In such languages, the tonal marking of a unit is a reflection of its accentual prominence. Jeddah Arabic consequently displays contrasts according to prominence where some syllables are more prominent than others and some words are more prominent than others. These syntagmatic contrasts as we have seen translate into contrasting phonetic detail that includes both tonal and non-tonal correlates. Jeddah Arabic demonstrated the use of a further phonetic correlate to mark accentual prominence in Arabic, which is rise speed, as well as demonstrated a contrary lengthening pattern in the marking of the intermediate phrase boundary. Although comparison is limited to two other Arabic intonational systems, it was seen that JA adds a further feature to the variation in the marking of prosodic constituents among Arabic dialects. The variation is also seen in the general manner and number of correlates used to implement prominence in Arabic. In fact, an important implication to be drawn from this thesis' results is how important phonetic detail is to the expression of prominence in Arabic. More importantly it was evident there is indeed structure to this surface phonetic variation that takes into account language-specific implementation rules, as well as enable more cross-linguistic generalisations.

7.3 Contribution of the thesis

The study of Jeddah Arabic prosody adds further intonational evidence to the limited number of Arabic intonational studies. The study adds to the prosodic typology in general, and to the Arabic prosodic spectrum in particular. The facts from the TBU identity in the dialect, the consistent pitch alignment with this TBU despite syllabic/segmental and time pressure factors, focus structure, as well as the prosodic structure all show how JA is both similar and different to previously studies dialects. In addition to this, the size of the pitch accent inventory and the identity of the accents alongside the unique combinations of nuclear accents phrase and boundary tones via which the dialect expresses pragmatically establised tune melodies demonstrate how intonation varies cross-dialectally. Findings from alignment patterns support the existence of a Segmental Anchoring Hypothesis, and findings from focus structure support a phonological hierarchy of accents in Arabic.

A more general contribution relates to methodology employed in the thesis. Throughout the chapters it can be seen how the thesis expands on the methodological practices used in previous studies. The number of participants in an Arabic intonation based study is relatively different; a representative sample of 20 speakers controlled for gender was used in order to aid in the reliability of the reported results. Additionally, more segmental anchors and acoustic variables were used for alignment measurement, more syllable types were added all in order to reach a well-rounded conclusion. Moreover, the current thesis is conducted on a relatively large corpus that includes many hours of recordings. The corpus includes a wealth of material for future studies on the production of intonational categories incorporating different reading material, narration, and simple and complex sentential constructions. While the analyses relevant to the current thesis' research questions are reported, the less relevant information was not included though still analysed and would thus constitute a base for further intonational investigation on the dialect.

In addition to this, facts from this Arabic study along with previous Arabic studies help contribute to a better practice and understanding in the field of speech technology. The limited efforts regarding Arabic speech recognition despite the stark variability among Arabic dialects and the impractical goal of agreeing on a standard Arabic language means that speech recognition practices both in linguistics and outside the linguistic field are prone to imperfection. Via the use of prosodic information from modelling and processing pitch contours, in addition to lexical, grammatical and segmental information from a large number of Arabic dialects, we can help improve the segmentation, translation and pronunciation of Arabic on modern day software and tech platforms designed for different purposes. Moreover, it is the researcher's wish to have contributed, in particular, to future efforts of developing a specific ToBI notation system on Arabic that takes into account the variability among dialects, parallel to what has been done in American ToBI, Dutch ToBI and Chinese ToBI (Jun, 2005).

7.4 Future research

The findings from the thesis pave way for future perceptual and articulatory studies on Arabic prosody. Arabic experimental studies on intonation, while scarce, are mainly carried out on the production aspect of intonation. Thus, more studies are needed on the perception element necessary for intonation in these dialects, particularly investigations that tap into the ways the observed patterns in production are translated in the perception of listeners. The perceptual studies may seek to investigate the categorical distinctions between accent types and tune melodies taking into consideration target scaling and alignment properties. Moreover, regarding focus, a confirmatory study may be carried out in order evaluate the most important cues detected by listeners in the interpretation of narrow and broad focus. Of course, studies on Arabic prosody would also benefit considerably from a rigorous analysis incorporating discourse structure and meaning, prosodic structure planning and implementation, as well as syntactic and semantic aspects in the expression of prosody in each variety.

References

Abdoh, E., 2010. *A study of the phonological structure and representation of first words in Arabic.* PhD thesis. University of Leicester.

Abu-Mansour, M., 2008. Meccan Arabic. In: K. Versteegh, ed. *Encyclopedia of Arabic language and linguistics IV (Q-Z).* Brill.

Al-Essa, A., 2009. When Najd meets Hijaz: Dialect contact in Jeddah. *Arabic dialectology,* Issue 53, pp. 203-222.

Al-Mohanna, F., 2009. *Positional syllable maximality: Syllabification in Hejazi.* [Online] Available at: http://repository.ksu.edu.sa/jspui/handle/123456789/5637. [Accessed 2013].

Al-Shehri, A., 1993. *Urbanization and linguistic variation and change: A sociolinguistic study of the impact of urbanization on the linguistic behavior of urbanized rural imiigrants in Hijaz, Saudi Arabia.* PhD thesis. University of Essex.

Al-Zaidi, M., 2014. *Information structure and intonation in Hijazi Arabic.* PhD thesis. University of Essex.

Al-Zaidi, M., Xu, Y., & Xu, A. 2019. Prosodic encoding of focus in Hijazi Arabic. *Speech communication,* Issue: 106, pp. 127-149.

El Zarka, D. 2011. Leading, linking and closing tonal contours in Egyptian Arabic: What a simple intonation system tells us about the nature of intonation. In *E. Broselow & H. Quali,* eds. *Perspectives on ArabicLinguistics, Papers from the Annual Symposium on Arabic Linguistics, Volume XXII–XXIII :College Park, Maryland, March 2008 and Wisconin 2009 [Current Issues in Linguistic Theory 317].* Issue: *57–73. Amsterdam: John Benjamins.*

El Zarka, D. 2013. Pragmatic Functions and the Biological Codes: Evidence from Sentence Topic and Focus in Egyptian Arabic Declaratives. In: S. Hancil & D. Hirst, eds. *Prosody and Iconicity*, *Iconicity in Language and Literature 13*. 109–126. Amsterdam: Benjamins.

Arvaniti, A., 2012. Segment to tone association. In: A. Cohn, C. Fougeron & M. Huffman, eds. *The oxford handbook of laboratory phonology*.

Arvaniti, A., 2016. Analytical decisions in intonation research and the role of representations: Lessons from Romani. *Laboratory phonology: Journal of the association for laboratory phonology,* Issue 7, pp. 1-43.

Arvaniti, A., 2017. The autosegmental metrical model of intonational phonology. In: S. Shattuck-Hufnagel & J. Barnes, eds. *Prosodic theory and practice.* MIT press.

Arvaniti, A., Ladd, D. & Mennen, I., 1998. Stability of tonal alignment: The case of Greek prenuclear accents. *Journal of phonetics,* Issue 26, pp. 3-25.

Ashby, M., 1978. A study of two English nuclear tones.. *Language and speech,* Volume 21, pp. 326-336.

Astruc, L. & Prieto, P., 2006. Stress and accent: Acoustic correlates of metrical prominence in Catalan. *TRW: Tutorial and research workshop on experimental linguistics*, pp. 28-30.

Atterer, M. & Ladd, D., 2004. On the phonetics and phonology of 'segmental anchoring' of F0: evidence from German. *Journal of phonetics,* Issue 32, pp. 177-197.

Avesani, C. & Vayra, M., 2005. *Accenting, deaccenting and information structure in Italian dialogues,* Wisconsin: Omnipress.

Baeshen, L., 2008. *Atabaat wa nabaat: Hikayaat sha3biya Hijaziya (Hijazi anecdotes).* Saudi Arabia.

Barakat, M., Ohala, J., Pellegrino, F., 1999. *Prosody as a distinctive feature for the discrimination of Arabic dialects.* Budapest, the 6th European conference on speech communication and technology (EUROSPEECH '99).
Baumann, S. & Winter, B., 2018. What makes a word prominent? Predicting untraine german listeners perceptual judgments. *Journal of phonetics,* Issue 70, pp. 20-38.
Beckman, M., 1986. *Stress and non-stress accent.* De Gruyter.
Beckman, M. & Edwards, J., 1990. Articulatory evidence for differentiating stress categories. In: P. Keating, ed. *Phonological structure and phonetic form: paper in laboratory phonology III.* Cambridge University Press.
Beckman, M. & Edwards, J., 1990. Lengthenings and shortenings and the nature of prosodic constituency. In: J. Kingston & M. Beckman, eds. *Papers in laboratory phonology I.* Cambridge: Cambridge university press, pp. 179-200.
Beckman, M. & Hirschberg, J., 1994. *The ToBI annotation conventions,* Manuscript, Ohio university.
Beckman, M. & Pierrehumbert, J., 1986. Intonational structure in Japanese and English. *Phonology Yearbook,* pp. 255-309.
Boersma, P. & Weenink, D., 1992-2019. *Praat: Doing phonetics by computer.* Version: 6.0.4.3.
Bouchhioua, N., 2008. *The acoustic correlates of stress and accent in Tunisian Arabic: A comparative study with English.* PhD thesis. University of Carthage.
Broselow, E., 1992. Parametric variation in Arabic dialect phonology. In: E. Broselow, M. Eid & J. McCarthy, eds. *Perspectives on arabic linguistics IV.* John Benjamins.
Broselow, E., Chen, S. & Huffman, M., 1997. Syllable weight: convergence of phonology and phonetics. *Phonology,* Issue 14, pp. 47-82.
Broselow, E., Chen, S. I. & Huffman, M., 1997. Syllable weight: Convergence of phonology and phonetics. *Phonolgy,* Issue 14, pp. 47-82.
Bruce, G., 1977. *Swedish word accents in sentence perspective.* Lund: Gleerup.
Chahal, D., 1999. *A preliminary analysis of Lebanese Arabic intonation.* Proceedings of the 1999 conference of the Australian linguistic society.
Chahal, D., 2001. *Modeling the intonation of Lebanese Arabic using the autosegmental-metrical framework: A comparison with English.* PhD thesis. University of Melbourne.
Chahal, D., 2003. *Phonetic cues to prominence in Lebanese Arabic.* Barcelona, Proceedings of the 15th international congress of phonetic sciences.
Cole, J., 2015. Prosody in context: A review. *Language, cognition and neuroscience,* Issue 30, pp. 1-31.
D'Alessandro, C. & Mertens, P., 1995. Automatic pitch contour stylization using a model of tonal perception. *Computer speech and language,* Issue 3, pp. 257-288.
De Jong, K., 1995. The supraglottal articulation of prominence in English: LInguistic stress as localized hyperarticulation. *Journal of the acoustical society of America,* Issue 97, pp. 491-504.
De Jong, K., Beckman, M. & Edwards, J., 1993. The interplay between prosodic structure and coarticulation. *Language and speech,* Issue 36, pp. 197-212.
De Jong, K. & Zawaydeh, B., 1999. Stress, duration and intonation in Arabic word-level prosody. *Journal of phonetics,* Issue 27, pp. 3-22.

De Jong, K. & Zawaydeh, B., 2002. Comparing stress, lexical focus and segmental focus: patterns of variation in Arabic vowel duration. *Journal of phonetics,* Issue 1, pp. 53-75.
De Leeuw, E., 2009. *When your native language sounds foreign: A phonetic investigation into first language attrition.* PhD thesis. Queen Margaret University, Edinburgh.
D'Imperio, M., 1997a. *Narrow focus and focal accent in the Neapolitan variety of Italian.* Athens, Intonation: theory and applications.
D'Imperio, M., 1997b. Breadth of focus, modality and prominence perception in Neapolitan Italian. *OSU working papers in Linguistics,* Issue 50, pp. 19-39.
D'Imperio, M., 2000. *on defining tonal targets from a perception perspective.* PhD thesis. Ohio state university.
D'Imperio, M., 2012. Tonal alignment. In: A. Cohn, C. Fougeron & M. Huffman, eds. *The Oxford handbook of laboratory phonology*.
D'Imperio, M. & House, D., 1997. *Perception of questions and statements in Neapolitan Italian.* Eurospeech '97.
D'Imperio, M., Petrone, C. & Nguyen, N., 2007. Effects of tonal alignment on lexical identification in Italian. In: C. Gussenhoven, ed. *Tones and Tunes.* Mouton de Gruyter, pp. 79-106.
Donzel, M. & Beinum, F., 2000. Pitch movements and information structure in spontaneous Dutch discourse. In: A. Botinis, ed. *Intonation: Analysis, modelling and technology.* Dordrecht: Kluwer academic publishers, pp. 145-161.
Face, T., 2002. Local intonational marking of Spanish contrastive focus. *Probus,* Issue 14, pp. 71-92.
Flemming, E. & Cho, H., 2017. The Phonetic Specification of Contour Tones: Evidence from the Mandarin Rising Tone. *Phonology,* Volume 34, pp. 1-40.
Frota, S., 2000. *Prosody and focus in European Portuguese: Phonological phrasing and intonation.* New York & London: Garland Publishing.
Fujisaki, H., 1983. Dynamic characteristics of voice fundamental frequency in speech and singing.. In: P. MacNeilage, ed. *The production of speech.* New York: Springer-Verlag, pp. 39-55.
Ghazali, S., Hamdi, R. & Barkat, M., 2002. *Speech rhythm variation in Arabic dialects.* Aix-en-provence, Proceedings from speech prosody.
Grabe, E. & Low, E., 2002. Durational variability in speech and rhythm class hypothesis. *papers in laboratory phonology,* Issue 7.
Gussenhoven, C., 1983. Testing the reality of focus domains. *Language and speech,* 26(1), pp. 61-80.
Gussenhoven, C., 2004. *The phonology of tone and intonation.* Cambridge University Press.
Hellmuth, S., 2004. *Prosodic weight and phonological phrasing in Cairene Arabic.* Chicago, Proceedings of the Annual meeting of the Chicago linguistic society.
Hellmuth, S., 2005. No de-accenting in (or of) phrases: Evidence from Arabic for cross-linguistic and cross-dialectal prpsodic variation. In: S. Frota & M. F. M. Vigario, eds. *Prosodies (with special reference to Iberian languages).* Mouton de Gruyter, pp. 99-121.
Hellmuth, S., 2006a. *Intonational pitch accent distribution in Egyptian Arabic.* PhD thesis. University of London: school of oriental and african studies.

Hellmuth, S., 2006b. Focus-related pitch range manipulation (and peak alignment effects) in Egyptian Arabic. *Proceedings of speech prosody,* pp. 410-413.
Hellmuth, S., 2007. The relationship between prosodic structure and pitch accent distribution: evidence from Egyptian Arabic. *The linguistics review,* Issue 24, pp. 291-316.
Hellmuth, S., 2011. Acoustic cues to focus and givenness in Egyptian Arabic. In: B. Heselwood & H. Zeki, eds. *Instrumental Studies in Arabic Phonetics.* John Benjamins Publishing.
Hellmuth, S., 2013. Phonology. In: J. Owens, ed. *The Oxford Handbook of Arabic Linguistics.* Oxford University Press.
Hellmuth, S., 2014. Investigating variation in Arabic intonation: The case for a multi-level corpus approach. In: S. Farwaneh & H. Ouali, eds *Percpectives on Arabic linguistics XXIV-XXV*. John Benjamins publishing.
Hellmuth, S., Bouchhioua, N. & Almbark, R., 2014. *Acquiring the phonetics and phonology of English word stress: comparing learners from different L1 backgrounds.* Concordia working papers in Aplied linguistics.
Hellmuth, S. & Chahal, D., 2014. The intonation of Lebanese and Egyptian Arabic. In: S. A. Jun, ed. *Prosodic Typology II: The phonology of intonation and phrasing.* Oxford University Press.
HTK Toolkit:.[Online]. Available at: http://htk.eng.cam.ac.uk/
Hirst, D., 2015. ProZed: A speech prosody editor for linguists, using analysis-by-synthesis. In: K. Hirose & J. Tao, eds. *Speech prosody in speech analysis: Modeling and generation of prosody for high quality and flexible speech synthesis.* Springer Berlin Heidelberg.
Hulst, H., 2014a. The study of word accent and stress: past, present and future. In: H. Hulst, ed. *Word stress: theoretical and typological issues.* Cambridge University Press.
Hulst, H., 2014b. Representing rhythm. In: H. Hulst, ed. *Word stress: theoretical and typological issues.* Cambridge University Press.
Hyman, J., 1985. *A theory of phonological weight.* Dordrecht: Foris publications.
Hyman, L., 2006. Word-prosodic typology. *Phonology,* Issue 23, pp. 225-257.
Ingham, B., 1971. Some characteristics of Meccan speech. *Bulletin of the school of Oriental and African studies,* Issue 34, pp. 273-297.
Ingham, B., 2009. Saudi Arabia. In: K. Versteegh, ed. *Encyclopedia of Arabic language and linguistics IV (Q-Z).* Brill, pp. 123-130.
Jun, S. A., 2005. *Prosodic typology I: The phonology of intonation and phrasing.* Oxford University Press.
Jun, S. A. & Fletcher, J., 2014. Methodology of studying intonation: from data collection to data analysis. In: S. A. Jun, ed. *Prosodic Typology II: The phonology of intonation and phrasing.* Oxford University Press.
Jun, S. A. & Fougeron, C., 2000. A phonological model of French intonation. In: A. Botinis, ed. *Intonation: Analysis, modelling, and technology.* Dordrecht: Kluwer academic publishers, pp. 209-242.
Keating, P. and Shattuck-Hufnagel, S., 2002. A prosodic view of word form encoding for speech production. *UCLA working papers in phonetics,* Issue 101, pp. 112-156.
Khattab, G. & Al-Tamimi, J., 2014. Geminate timing in Lebanese Arabic: The relationship between phonetic timing and phonological structure. *Journal of the association for laboratory phonology,* Issue 5.

Krivokapic, J., 2012. Prosodic planning in speech. In: S. Fuchs, M. Weihrich, D. Pape, P. Perrier, eds. *Speech planning and dynamics.* Peter Lang.

Ladd, D., Faulkner, D., Faulkner, H. & Schepman, A., 1999. Constant segmental anchoring of F0 movement under changes in speech rate. *Journal of the acoustical society of America,* Issue 106, pp. 1543-1554.

Ladd, D. R., 2008. *Intonational Phonology.* Cambridge University Press.

Ladd, D. et al., 2009. Structural and dialectal effects on pitch peak alignment in two variaties of British English. *Journal of phonetics,* Issue 37, pp. 145-161.

Lubbers, M. & Torreira, F., 2016. *Praatalign: an interactive Praat plug-in for performing phonetic forced alignment.* Retrived from: https://github.com/dopfishh/praatalign.

McCarthy, J. & Prince, A., 1990. Foot and word in prosodic morphology: the Arabic broken plural. *Natural language and linguistic theory,* Issue 8, pp. 209-286.

Mennan, I., Schaeffler, F. & Docherty, G., 2012. Cross-language differences in fundamental frequency range; A comparison of English and German. *Journal of the acoustical society of America,* Issue 131, pp. 2249-2260.

Mertens, P., (to appear. The prosogram model for pitch stylization and its applications in intonation transcription. In: S. Shattuck-Hufnagel & J. Barnes, eds. *prosodic theory and practice.* MIT press.

Moussa, H., 2013. *A phonological description of the urban Hijazi dialect spoken in Jeddah.* MA thesis. University of Manchester.

Pierrehumbert, J., 1980. *The phonology and phonetics of English intonation.* PhD thesis. MIT.

Prieto, P., 2011. Tonal alignment. In: M. Van Ootendorp, C. Ewen, B. Hume & K. Rice, eds. *Companion to phonology.* Wiley-Blackwell, pp. 1185-1203.

Prieto, P., 2011. Tonal Alignment. In: M. Van Oostendorp, C. Ewen, B. Hume & K. Rice, eds. *Companion to phonology.* Wiley-Blackwell, pp. 1185-1203.

Prieto, P., Van Santen, J. & Hirschberg, J., 1995. Tonal alignment patterns in Spanish. *Journal of phonetics,* Issue 23, pp. 429-451.

Rao, R., 2010. Final lengthening and pause duration in three dialects of Spanish. In: M. Ortega-Llebaria, ed. *Selected Proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology.* Somerville, MA: Cascadilla Proceedings Project, pp. 69-82.

Schiel, F., 1999. *Automatic phonetic transcription of non-prompted speech.* Proceedings of the 1999 ICPhS.

Sluijter, A. & Van Heuven, V., 1996a. Spectral balance as an acoustic correlate of linguistic stress. *Journal of the acoustical society of America,* Issue 100, pp. 2471-2485.

Sluijter, A. & Van Heuven, V., 1996b. *Acoustic correlates of linguistic stress and accent in Dutch and American English.* Philadelphia, Proceedings of ICSL 96.

SpeCT: The speech corpus toolkit for Praat. [Online]. Available at: https://lennes.github.io/spect/

Thomas, E., 20111. *Sociophonetics: An introduction.* Palgrave Macmillan.

Turk, A., Satsuki, N. & Mariko, S., 2006. Acoustic segment durations in prosodic research: A practical guide. In: S. Sudhoff, et al. eds. *Methods in empirical prosody research.* De Gruyter.

Venditti, J., 2005. The J-ToBI model of Japanese intonation. In: S. A. Jun, ed. *Prosodic typology I: the phonology of intonation and phrasing* . Oxford University Press.

Wagner, M. & Watson, D., 2010. Experimental and theoretical advances in prosody: A review. *Language and cognitive processes,* Issue 25, pp. 905-945.
Wang, B. & Xu, Y., 2011. Differential prosodic encoding of topic and focus in sentence-initial position in Mandarin Chinese. *Journal of phonetics,* Issue 39, pp. 595-611.
Watson, J., 2002. *The phonology and morphology of Arabic.* Oxford University Press.
Watson, J., 2011. Word stress in Arabic. In: M. Oostendorp, C. Ewen, E. Hume & K. Rice, eds. *The Blackwell companion to phonology.* Oxford: Wiley-Blackwell.
Wichhmann, A., House, J. & Rietveld, T., 2000. Discourse constraints on F0 peak timing in English. In: A. Botinis, ed. *Intonation: Analysis, modelling and technology.* Kluwer academic publishers.
Wichmann, A., 2004. The intonation of please requests: A corpus based study. *Journal of pragmatics,* Issue 36, pp. 1521-1549.
Wightman, C., Shattuck-Hufnagel, S., Ostendorf, M. & Price, P., 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *The journal of the acoustical society of America,* Issue 91, pp. 1707-1717.
Xu, Y., 2010. In defense of lab speech in prosody research. *Journal of phonetics,* Issue 38, pp. 329-336.
Xu, Y., 2011. *Post-focus compression: Cross-linguistic distribution and historical origin.* Proceddings of the 17th International congress of phonetic sciences.
Xu, Y. & Sun, X., 2002. Maximum speed of pitch change and how it may relate to speech. *Journal of the acoustical society of America,* Volume 3, pp. 1399-1413.
Xu, Y. & Xu, C., 2005. Phonetic realization of focus in English declarative intonation. *Journal of phonetics,* Issue 33, pp. 159-197.
Yeou, M., 2004. *Effects of focus, position and syllable structure on F0 alignment patterns in Arabic.* Arabic Language Processing 2004.
Yeou, M., Embarki, M. & Al-Maqtari, S., 2007b. Contrastive focus and F0 patterns in three Aarabic dialects. *Nouveaux cahiers de linguistique francaise,* Issue 28, pp. 317-326.
Yeou, M., Embarki, M., Al-Maqtari, S. & Dodane, C., 2007a. *F0 alignment patterns in Arabic dialects.* Paper presented at the 16th ICPhS.

Appendix A: Transliteration

IPA	Transliteration	SAMPA (singleton/geminate)
ʔ	2 or `	?/??
b	b	b/bb
t	t	t/tt
θ	th	T/TT
ʒ	j	Z/ZZ
ħ	7 or H	X-/xx-
x	x or kh	x/xx
d	d	d/dd
ð	dh	D/DD
r	r	r/rr
z	z	z/zz
s	s	s/ss
ʃ	sh	S/SS
sˤ	S	s`/ss` (actually should be something else), e.g., s_j/s_js
dˤ	D	d`/dd` or d_j/ dd_j
tˤ	T or 6	t`/tt` or t_j/t_jt
ðˤ	Dh	D`/DD` we do not have it, but I can use: D/DD
zˤ	Z	z`/zz`
ʕ	3 or @	?-/ ??-
ɣ	gh	G/GG
f	f	f/ff
q	q	q/qq
g	g	g/gg
k	k	k/kk
l	l	l/ll
m	m	m/mm
n	n	n/nn
h	h	h/hh
w	w	w/ww
j	y	j/jj
ʧ	ch	tS/ttS
ʤ	dj	dZ/ddZ
v	v	v/vv
ɫ	L	5/55
p	p	p/pp
i	i	i
ɪ	I	I
e	e	e
ɛ	E	E
æ	ae	{
a	a	a
ɑ	A	A
ɔ	O	O
o	o	o
u	u	u
ʊ	U	U
i:/ii	i:/ii	i:
ɪ:/ɪɪ	I:/II	I:

e:	e:/ee	e:
ɛ:	E:/EE	E:
æ:	ae:/aeae	{:
a:	a:/aa	a:
ɑ:	A:/AA	A:
ɔ:	O:/OO	O:
o:	o:/oo	o:
u:	u:/uu	u:
ʊ:	U:/UU	U:
I don't know	?/french	<usb>

Appendix B: individual speaker alignment means

(Speakers are labelled numerically, and according to gender. M, F indicates speaker gender)

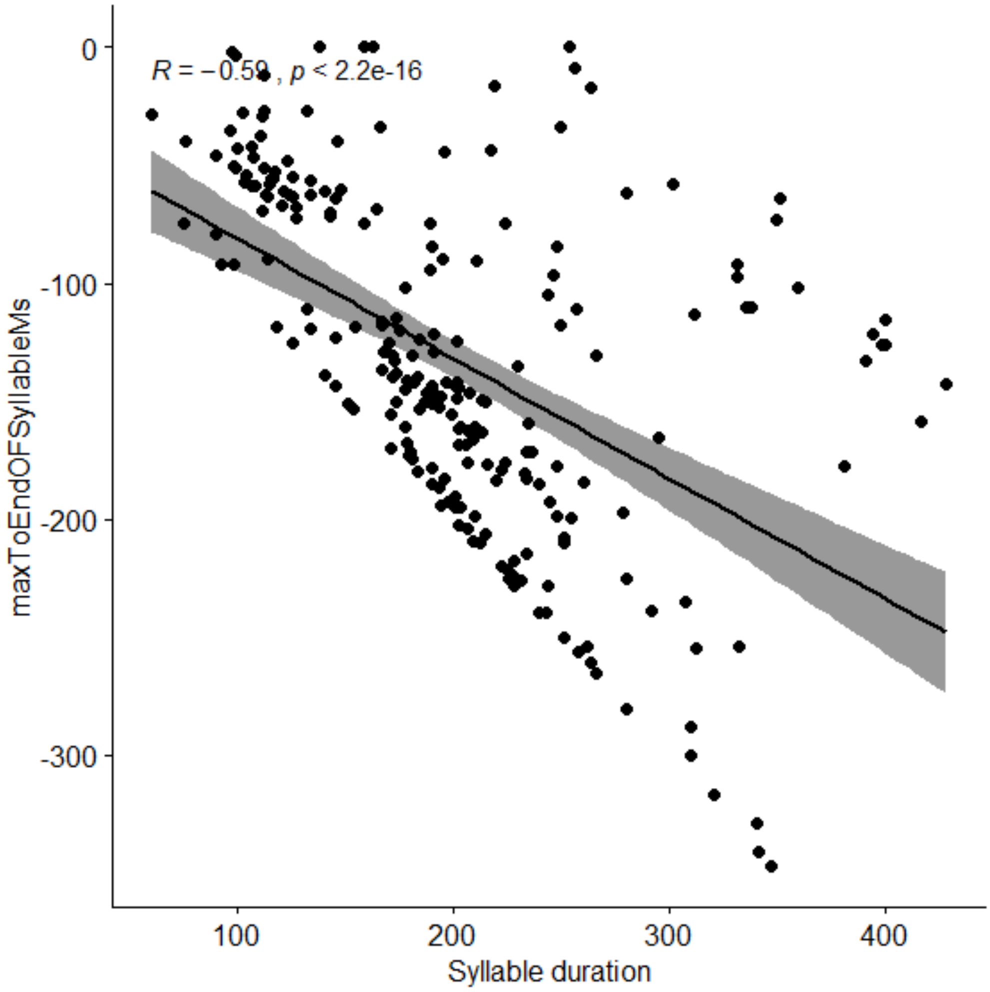
SPEAKER	H-TO-ONSET	H-TO-START-VOWEL	H-TO-END-VOWEL	L-TO-ONSET	L-TO-START-VOWEL
01_M	**68.9(28)**	**26.3(26.6)***	**-67.5(48)**	**3 (6)***	**-39.5 (11)**
01_F	**57.6(32)**	**16.8(31)***	**-71.5(50)**	**53.5(274)***	**12.7(271)***
02_F	**68.3(50)**	**33.8(50)***	**-65.8(56)**	**7.2(10)***	**-27.2(12)**
03_F	**69.4(64)***	**19.5(67)***	**-83.4(68)**	**34.5(319)***	**-15.3(321)***
04_F	**73.6(60)**	**25.3(60)***	**-70.1(51)**	**17.5(55)***	**-30.7(54)***
05_F	**135.8(89)**	**73.9(86)***	**-45.5(63)***	**-57(323)***	**-118.9(323)***
02_M	**59.3(49)**	**23(47)***	**-67.9(44)**	**10.8(19)***	**-25.4(19)**
03_M	**68.8(63)**	**19.5(59)***	**-77.6(50)**	**38.8(237)***	**-10.4(236)***
04_M	**108.3(63)**	**57.2(62)***	**-56.7(48)**	**22(52)***	**-29(52)**
06_F	**116.3(61)**	**60.5(60)***	**-47.6(60)***	**20.1(29)***	**-35.5(27)**
05_M	**68.2(42)**	**21.9(44)***	**-66.2(49)**	**22.5(94)***	**-23.7(95)**
06_M	**71.7(24)**	**33.1(25)**	**-49.4(37)**	**-23.2(203)***	**-61.9(204)***
07_M	**49.8(35)**	**4.2(33)***	**-83.5(48)**	**26.3(135)***	**-19.2(134)***
08_M	**130.8(73)**	**79.8(70)***	**-39.3(66)***	**-69.4(318)**	**-120.5(319)**
07_F	**73.4(66)**	**26.8(64)***	**-74.8(64)**	**21.6(132)***	**-25(134)***
09_M	**43(33)**	**8.7(34)***	**-73.2(39)**	**19.4(84)***	**-14.8(83)***
08_F	**79.4(52)**	**27.9(53)***	**-61.5(55)**	**54.3(206)***	**2.8(204)***
09_F	**63.5(55)**	**28.5(52)***	**-51(37)**	**48.2(122)***	**13.1(122)***
10_F	**77.8(53)**	**37.2(51)***	**-59.7(55)**	**10.4(26)***	**-30.1(29)***
10_M	**80.4(38)**	**37.5(37)***	**-65.5(43)**	**-3.4(106)***	**-46.3(107)***

Appendix C: syllable duration and peak location correlation

<u>1a. Correlation between ‘durationSyllableMS’ and ‘maxToEndOFSyllableMs’ combining intonational levels</u>

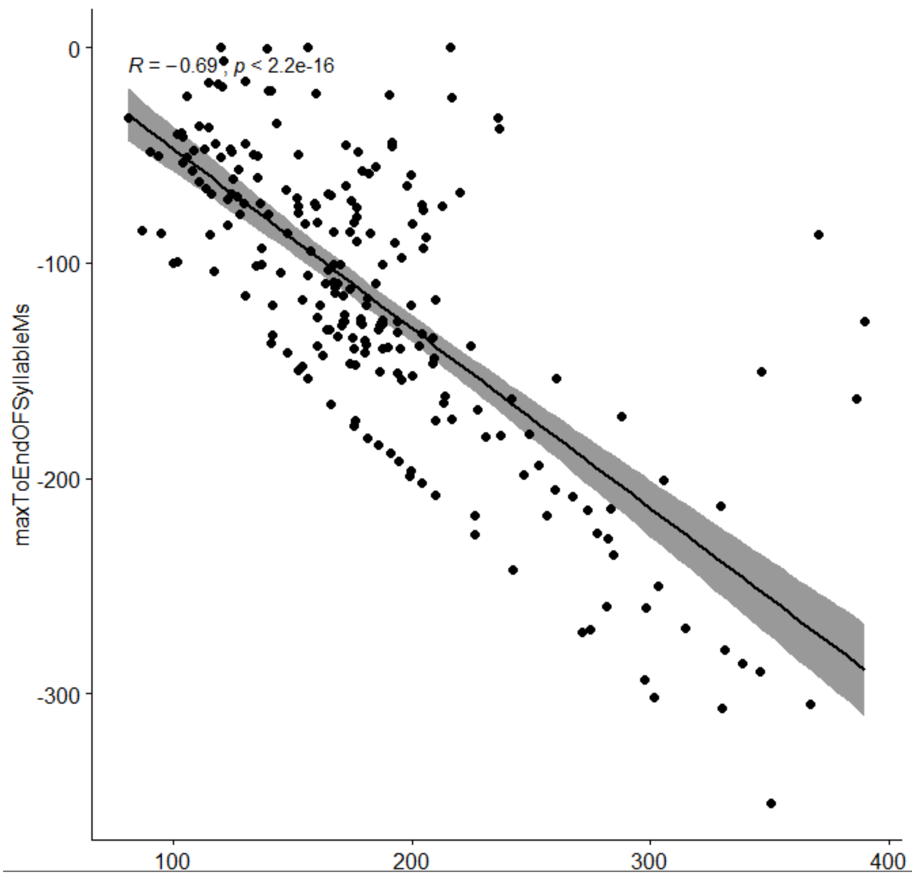
<u>For females:</u>

The variables named ‘durationSyllableMS’ and ‘maxToEndOFSyllableMs’ are significantly correlated (r=-0.589732, $p < 2.2e-16$). The relationship is negative and medium level.



<u>For males:</u>

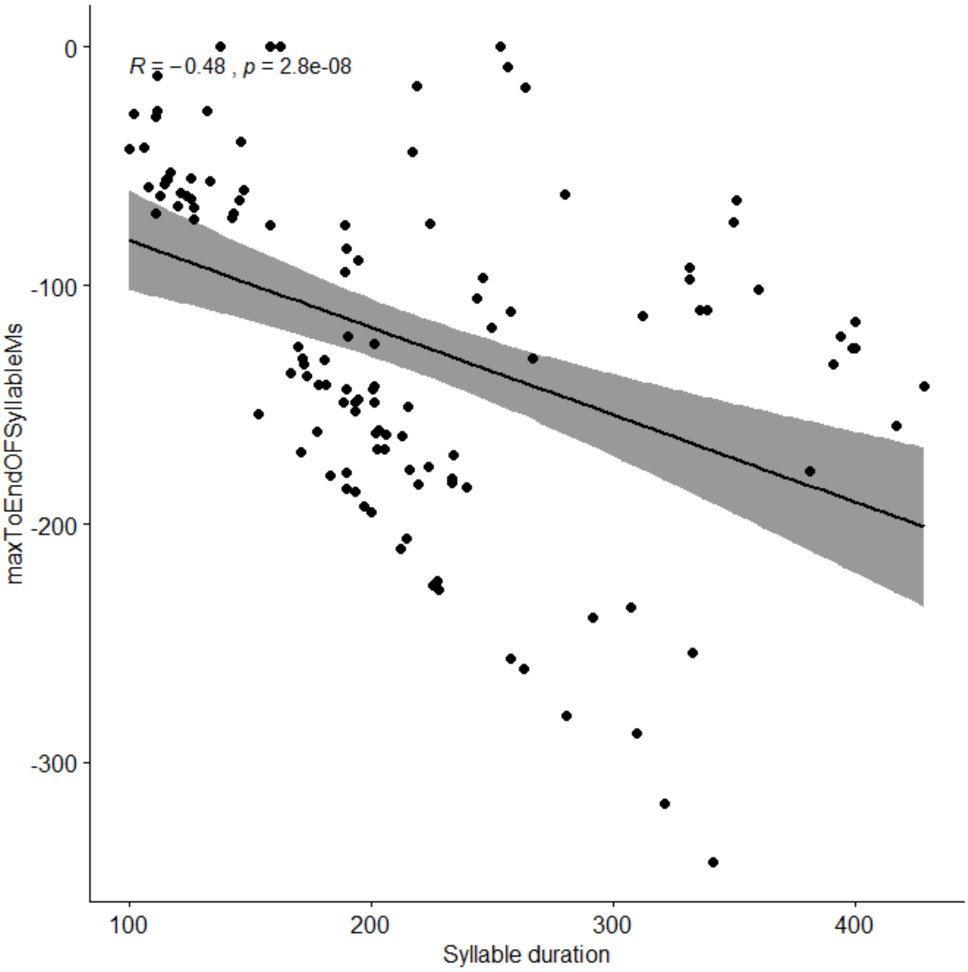
The variables named ‘durationSyllableMS’ and ‘maxToEndOFSyllableMs’ are significantly correlated (r=-0,6873926, p < 2.2e-16). The relationship is negative and medium level.



1b. Correlation between 'durationSyllableMS' and 'maxToEndOFSyllableMs' for intermediate phrases

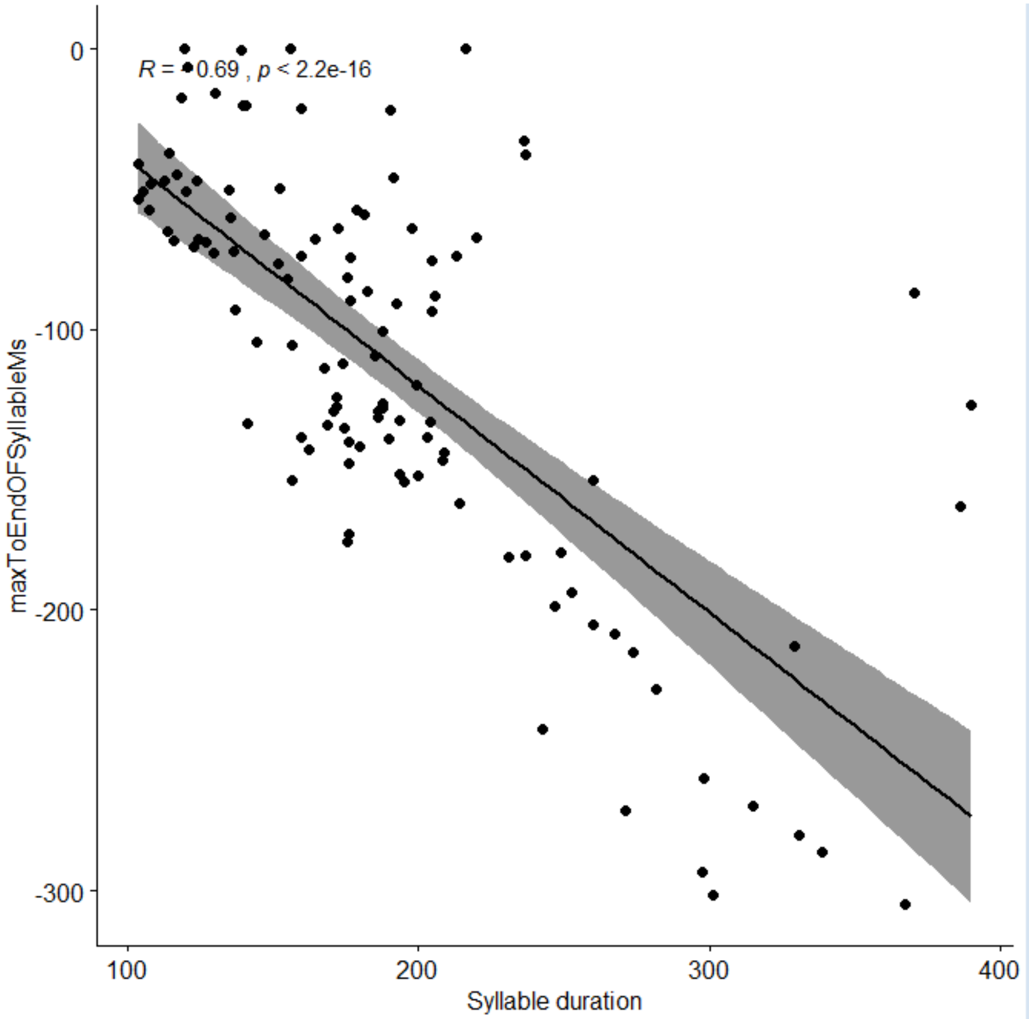
For female:

The variables named 'durationSyllableMS' and 'maxToEndOFSyllableMs' are significantly correlated (r=-0.483864, p=2.844e-08). The relationship is negative and medium level.



For male:

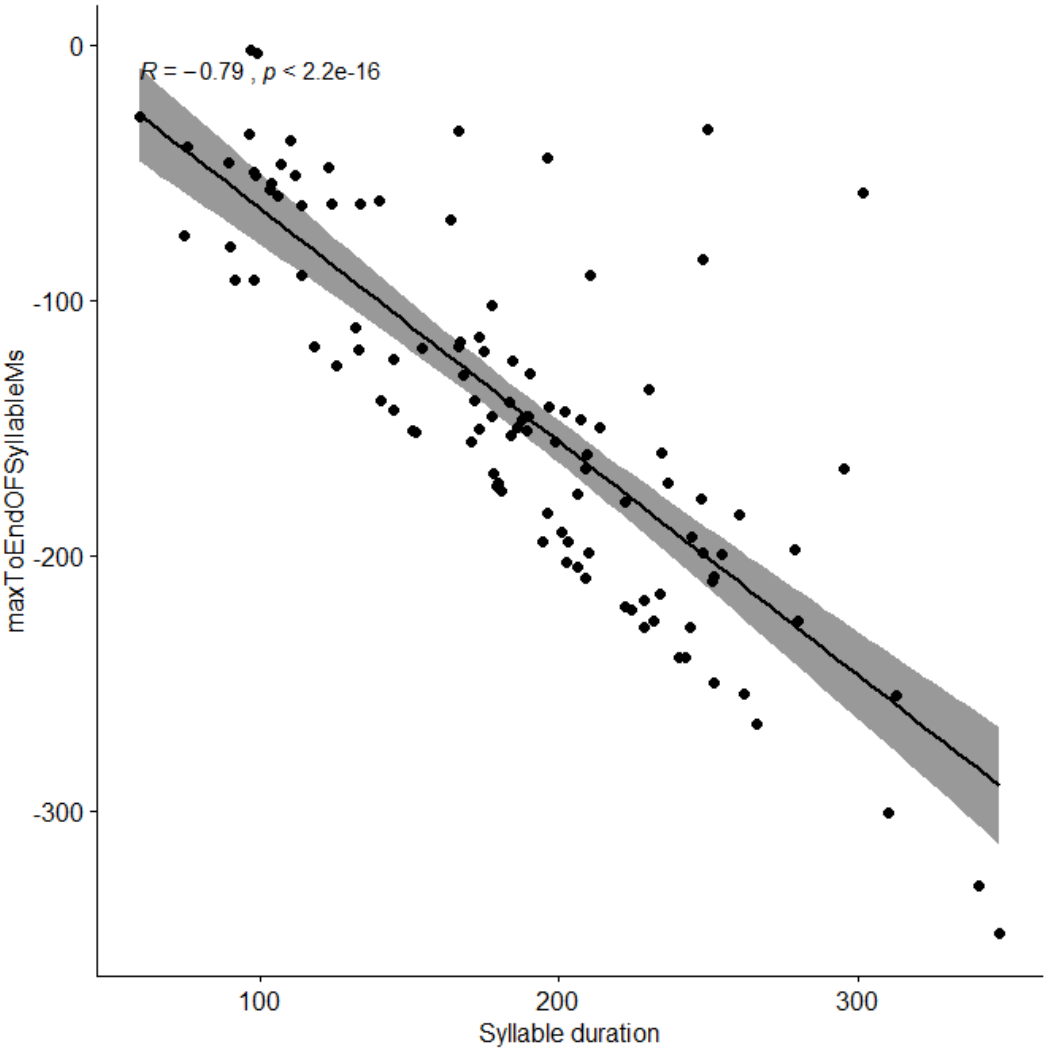
The variables named 'durationSyllableMS' and 'maxToEndOFSyllableMs' are significantly correlated (r=-0.6874604, p < 2.2e-16). The relationship is negative and medium level.



1c. Correlation between ‘durationSyllableMS’ and ‘maxToEndOFSyllableMs’ for intonational levels

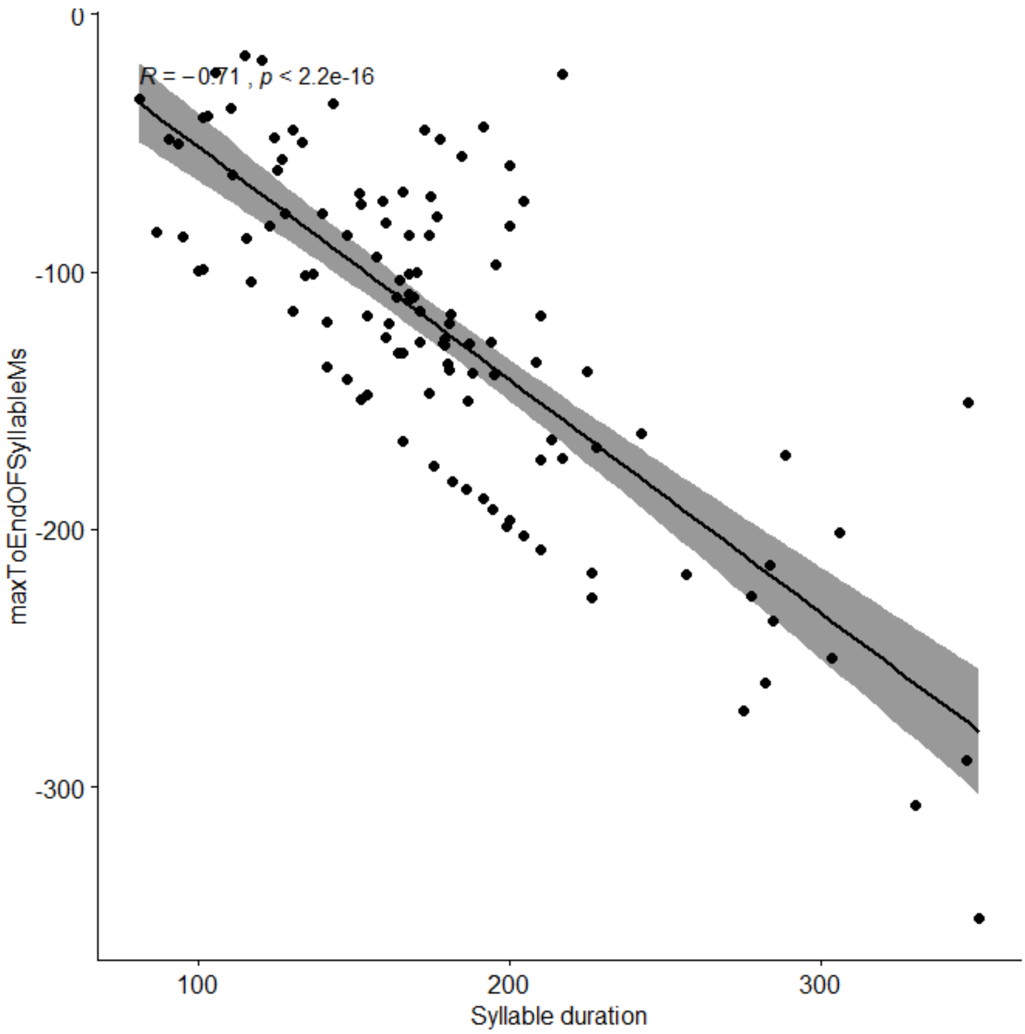
For females:

The variables named ‘durationSyllableMS’ and ‘maxToEndOFSyllableMs’ are significantly correlated (r=-0.7859972, p < 2.2e-16). The relationship is negative and strong.



For males:

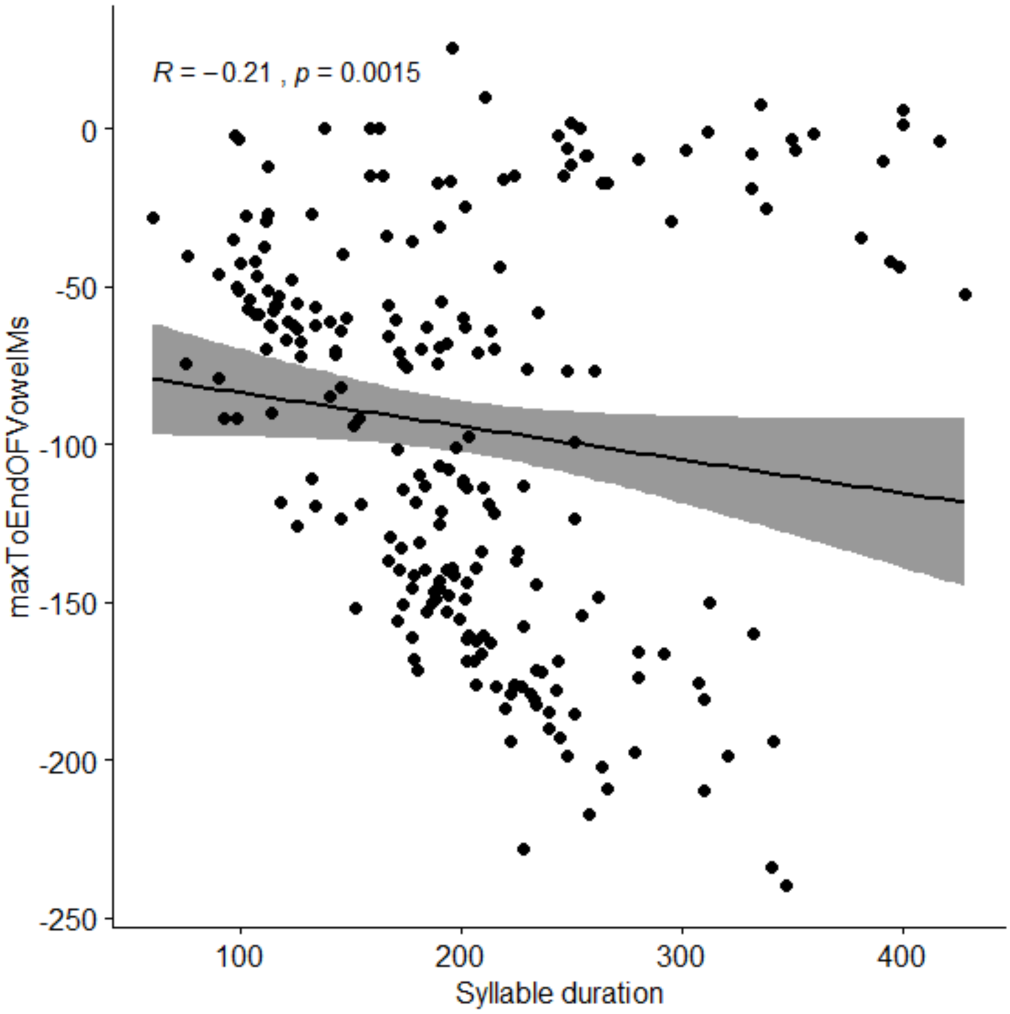
The variables named 'durationSyllableMS' and 'maxToEndOFSyllableMs' are significantly correlated (r=-0.7123112, p < 2.2e-16). The relationship is negative and strong.



2a. Correlation between ‘durationSyllableMS’ and ‘maxToEndOFVowelMs’ combining intonational levels

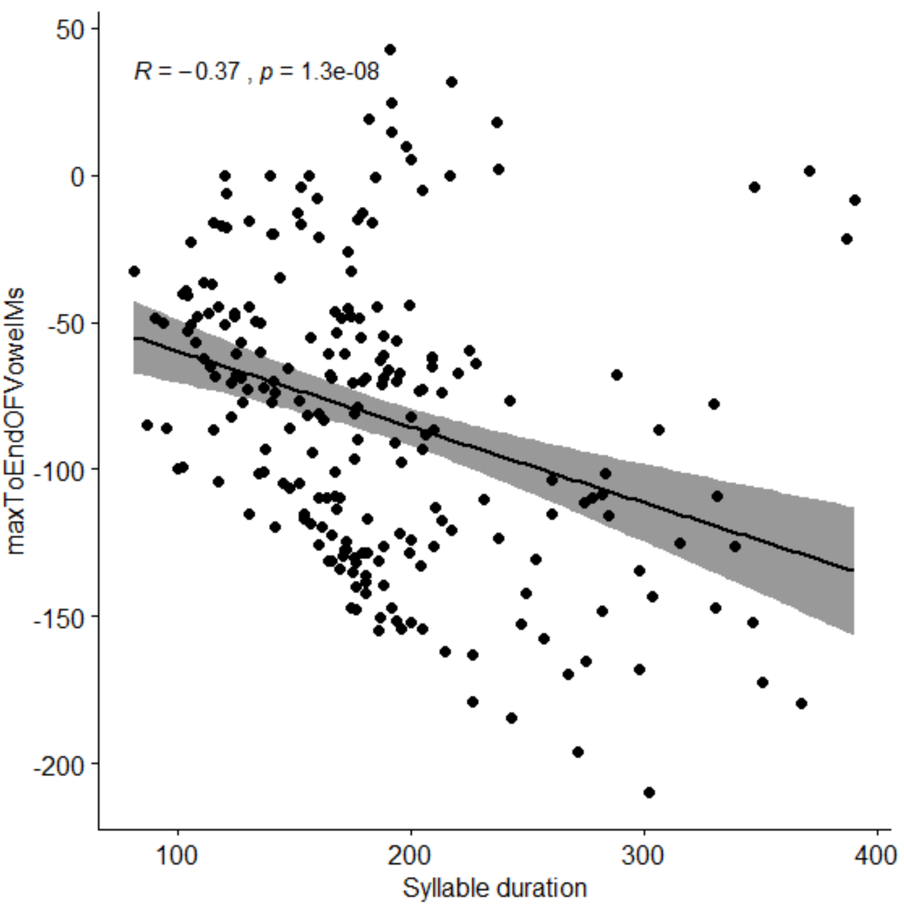
For females:

The variables named ‘durationSyllableMS’ and ‘maxToEndOFVowelMs’ are significantly correlated (r=-0.208994, p=0.001506). The relationship is negative but weak.



For males:

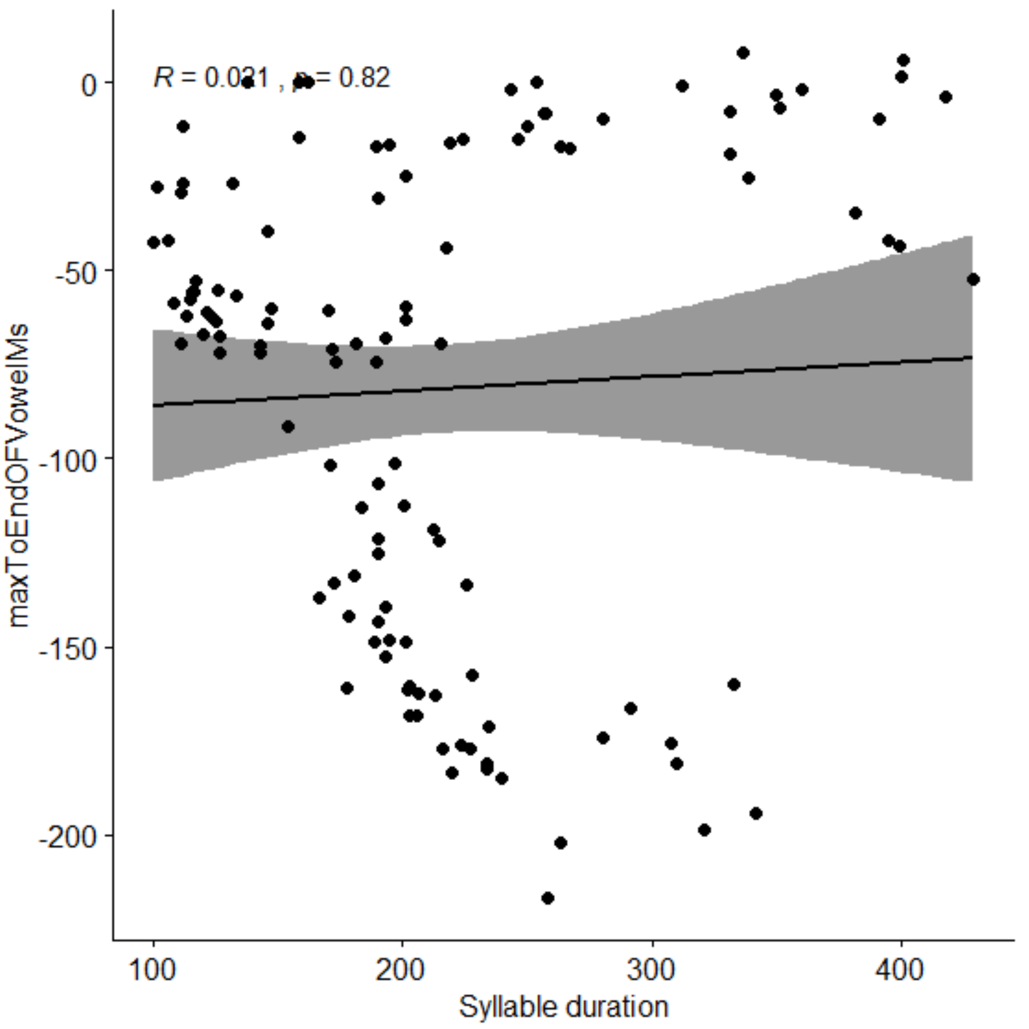
The variables named 'durationSyllableMS' and 'maxToEndOFVowelMs' are significantly correlated (r=-0,366747, p=1.331e-08). The relationship is negative but weak.



2b. Correlation between ‘durationSyllableMS’ and ‘maxToEndOFVowelMs’ for intermediate phrases

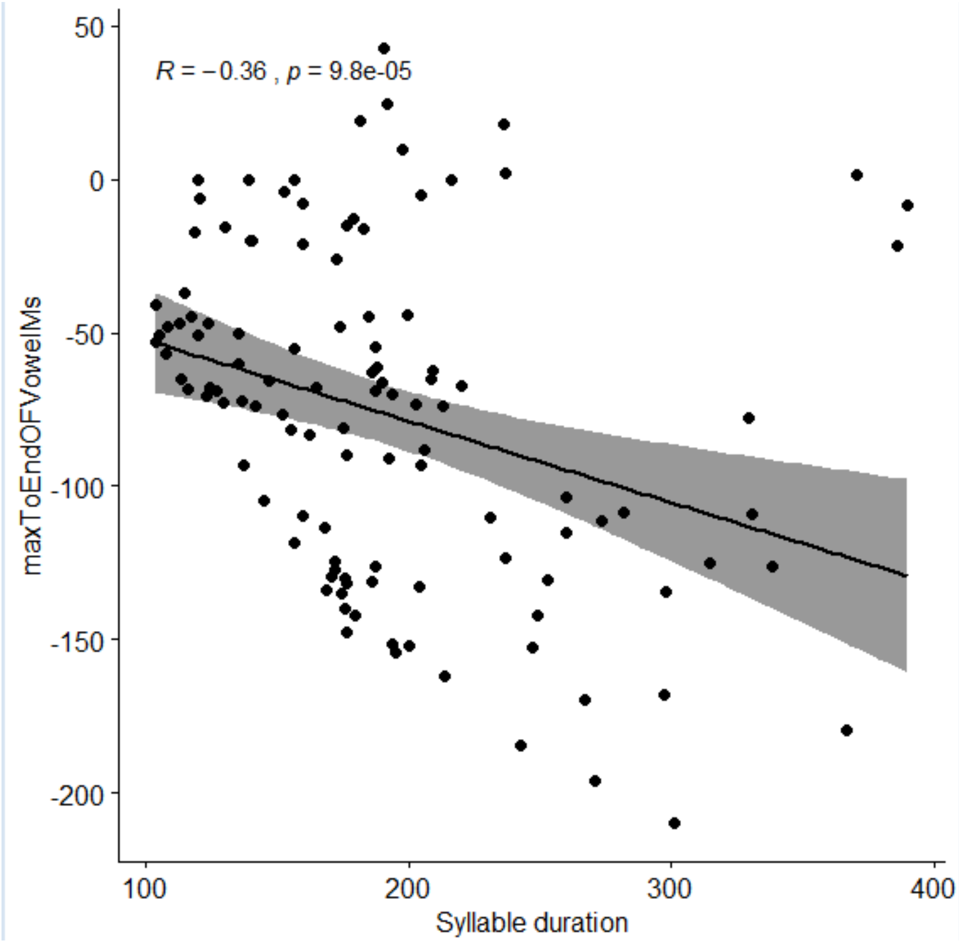
For females:

The variables named ‘durationSyllableMS’ and ‘maxToEndOFVowelMs’ are not significantly correlated (r=0.02143072, p=0.8178).



For males:

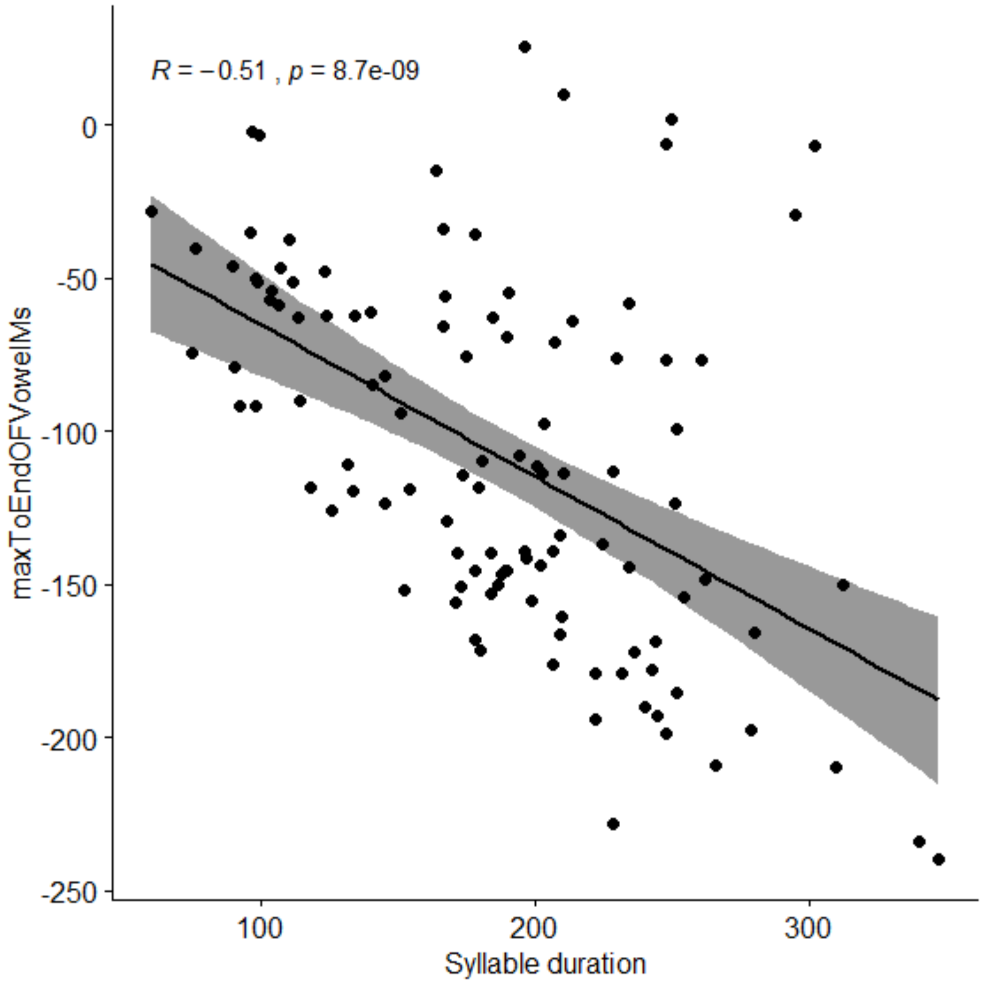
The variables named 'durationSyllableMS' and 'maxToEndOFVowelMs' are significantly correlated (r=-0.3582587, p=9.774e-05). The relationship is negative but weak.



2c. Correlation between 'durationSyllableMS' and 'maxToEndOFVowelMs' for intonational levels

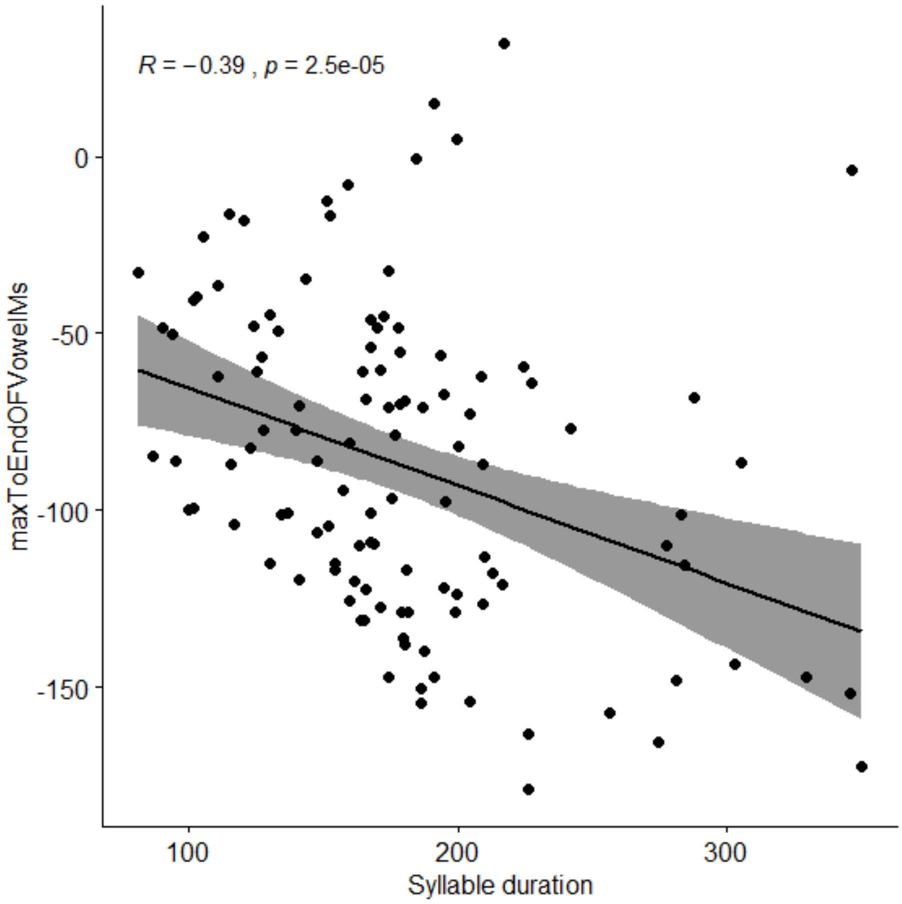
For females:

The variables named 'durationSyllableMS' and 'maxToEndOFVowelMs' are significantly correlated (r=-0.5148526, p=8.697e-09). The relationship is negative and medium level.



<u>For males:</u>

The variables named ‘durationSyllableMS’ and ‘maxToEndOFVowelMs’ are significantly correlated (r=-0.385699, p=2.455e-05). The relationship is negative but weak.



Appendix D: List of target sentences

كان لازم توضح انو للحفلات اصلاً

kaan laazim tiwaDDiH 2innu lilHaflaat 2aSlan[17]

٣١x٣١ +٣٣x٣٣=٢٠٥٠

waaHidu talaatiin fii/Darb waaHidu talaatiin zaa2id talaata w talaatiin fii/Darb talaata w talaatiin yisaawi 2alfeenu xamsiin

كان لازم تقول عسكري، بس قالت ملازم

kaan laazim tiguul 3askari, bas gaalat mulaazim

متى رحلتكم للندن؟

mita riHlatkum lilandan?

حلوة سجاجيد المدينة سجاجيد مكة حلوة بس حقت المدينة احلى

Hilwa sajaajiid almadiina, sajaajiid makka Hilwa bas Haggat almadiina 2aHla

٥٠+٥٠=١٠٠

xamsiin zaa2id xamsiin yisaawi miyya

كان لازم توضح انو لَهَفْو رِزقو

kaan laazim tiwaDDiH 2innu lahafu rizgu

كان لازم توضح انو لَهَفْو مكاني

kaan laazim tiwaDDiH 2innu lahafu makaani

كلمني بطريقة احسن من كدا!

Kallemni bTariiga 2aHsan min kida!

شفت الفيلم الجديد؟

shuft alfilm aljadiid?

ضَرَبَتو؟

Darabatu?

كان لازم تقول بالحفلات، بس قالت للحفلات

kaan laazim tiguul bilHaflaat, bas gaalat lilHaflaat

كان لازم توضح انو لَهَفو طَلَبات

kaan laazim tiwaDDiH 2innu lahafu Talabaat

كان لازم توضح انو رحلتهم حلوة

kaan laazim tiwaDDiH 2innu riHlathum Hilwa

كان لازم تقول لَفَحْو، بس قالت لَهَفْو

kaan laazim tiguul lafaHu, bas gaalat lahafu

اتصل عليا دحين!

2attaSil 3alayya daHiin!

سمعت انو اخوها اخد بنت الجيران مو بنت عمو

simi3t 2innu 2axuuha 2axad bint aljiiraan muu bint 3ammu

ابويا ارسلنا عشان نجيب انابيب وكراتين ولصق

2abuuya 2arsalna 3ashaan nijiib 2anaabiib wu karaatiin w laSag

كان لازم تقول للحفلات بس ماقِدْرَت

kaan laazim tiguul lilHaflaat, bas maa gidrat

كان لازم توضح انو للحفلات في القصور

kaan laazim tiwaDDiH 2innu lilHaflaat filguSuur

كان لازم تقول رحلتكم، بس قالت رحلتهم

kaan laazim tiguul riHlatkum, bas gaalat riHlathum

كان لازم تقول لَهَفْو بس ماقِدْرَت

kaan laazim tiguul lahafu bas maa gidrat

---

[17] The highlighted sentences are the tonal alignment experiment material. 4 underlined target words, each in 5 structural positions.

كان لازم توضح انو ملازم تاني
kaan laazim tiwaDDiH 2innu mulaazim taani

ممكن تفتحلي الباب؟
mumkin tiftaHli albaab?

١٠x ١٠-٥
3ashara fii/Darb 3ashara naagiS xamsa

كان لازم توضح انو ملازم تحقيق
kaan laazim tiwaDDiH 2innu mulaazim taHgiig

كان لازم توضح انو للحفلات اماكن
kaan laazim tiwaDDiH 2innu lilHaflaat 2amaakin

منى كتبت معروض وارسلتلو للملك قبل امس
Mona katabat ma3ruuD w 2arsalatu lilmalik

تعرف تتكلم فرنسي؟
ti3rif titkallam faransi?

كان لازم توضح انو رحلتهم قصيرة
kaan laazim tiwaDDiH 2innu riHlathum gaSiira

كان لازم توضح انو رحلتهم للبنان
kaan laazim tiwaDDiH 2innu riHlathum lilibnaan

كان لازم توضح انو ملازم تحقيقات
kaan laazim tiwaDDiH 2innu mulaazim taHgiigaat

كان لازم تقول ملازم بس ما قِدْرت
kaan laazim tiguul mulaazim, bas maa gidrat

الساعة كم دحين؟
2assaa3a kam daHiin?

كان لازم تقول رحلتهم بس ما قِدْرَت
kaan laazim tiguul riHlathum bas maa gidrat

The focus experiment stimuli:

ريتال خرجت مع سجى قبل امس؟ قبل امس ريتال خرجت مع لانا
Narrow Focus: Ritaal xarajat ma3 saja gabil 2ams? gabil 2ams Ritaal xarajat ma3 Laana

ريتال خرجت مع لانا مو سجى.
Filler: Ritaal xarajat ma3 Laana muu Saja

ايش صار قبل امس؟ ريتال خرجت مع لانا قبل امس
Broad Focus: 2eesh Saar gabil 2ams? Ritaal xarajat ma3 Laana gabil 2ams

ايش صار قبل امس؟ قبل امس ريتال خرجت مع لانا
Broad Focus: 2eesh Saar gabil 2ams? gabil 2ams Ritaal xarajat ma3 Laana

سمعت الجيران يقولو سجى خرجت مع ريتال بس قلتلهم لأ لانا خرجت مع ريتال.
Filler: simi3t aljiiraan yiguulu Saja xarajat ma3 Ritaal bas gultallahum la2, Laana xarajat ma3 Ritaal

مين خرج مع ريتال قبل امس؟ لانا خرجت مع ريتال قبل امس
Narrow Focus: miin xaraj ma3 Ritaal gabil 2ams? Laana xarajat ma3 Ritaal gabil 2ams

لانا خرجت مع سجى؟ (لا) لانا ماخرجت مع سجى
Filler: Laana xarajat ma3 Saja? la2 Laana maa xarajat ma3 Saja

قالولي لانا اختبرت مع ريتال بس قلتلهم لأ لانا خرجت مع ريتال.
Filler: gaalOOli Laana 2axtabarat ma3 Ritaal bas gultallahum la2, Laana xarajat ma3 Ritaal

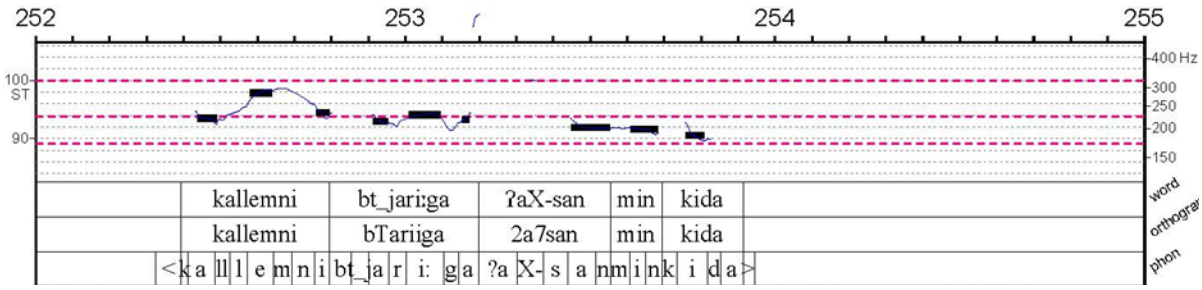
ريتال خرجت مع خالد قبل امس؟ ريتال خرجت مع لانا قبل امس
Narrow Focus: Ritaal xarajat ma3 xaalid gabil 2ams? Ritaal xarajat ma3 Laana gabil 2ams

ايش صار قبل امس؟لانا خرجت مع ريتال قبل امس

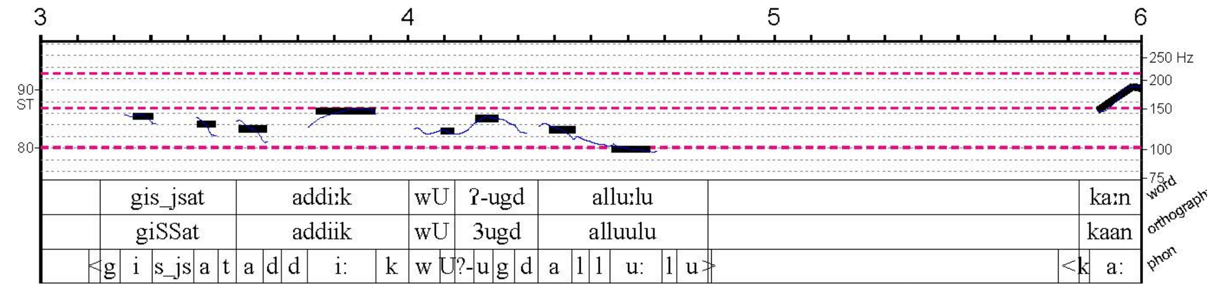
Broad Focus: 2eesh Saar gabil 2ams? Laana xarajat ma3 Ritaal gabil 2ams

Appendix E: Prosograms

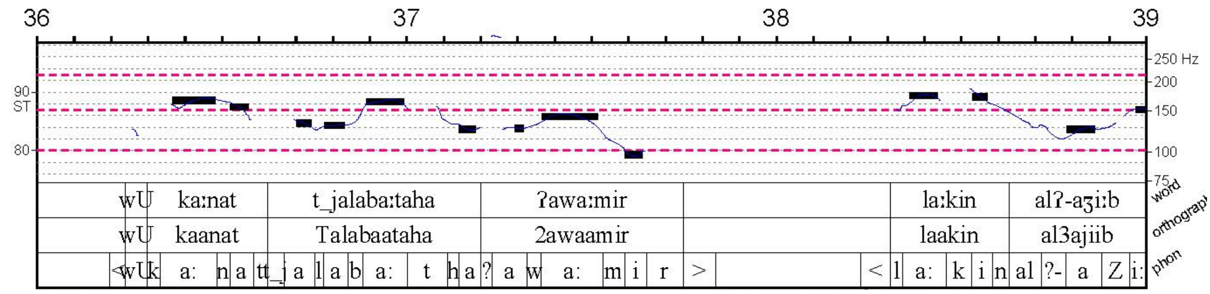
The output Prosogram contains three horizontal lines representing top, median and bottom speaker pitch range. Black stylisation in semitones, and a superimposed F0 curve in blue. The black stylisation occurs on every syllabic nucleus, that is, Prosogram detects each vowel and nucleus centre and represents it as a black line. This black stylisation represents the sonorant nuclei direction and duration detected by the tool. Those do not always indicate pitch accents. Following ToBI notation, accent types are based on the F0 contour. Moreover, the Prosogram contains three tiers representing the segmentation of words and syllables, while the vertical lines correspond to their boundaries.



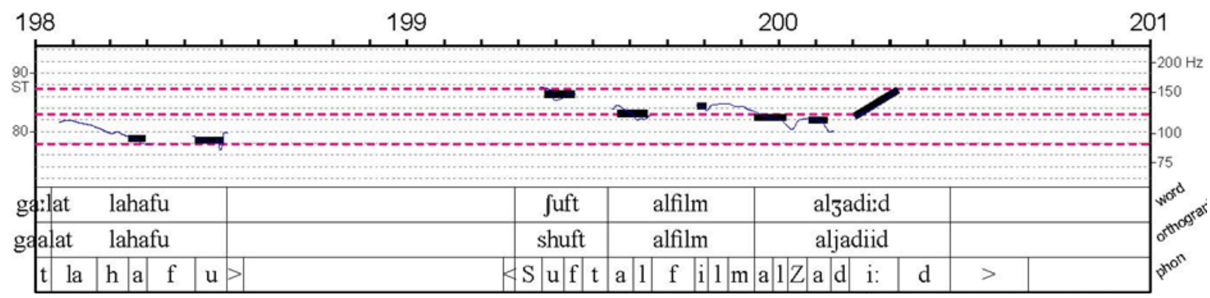
Chapter 4, imperative tune. 03_F



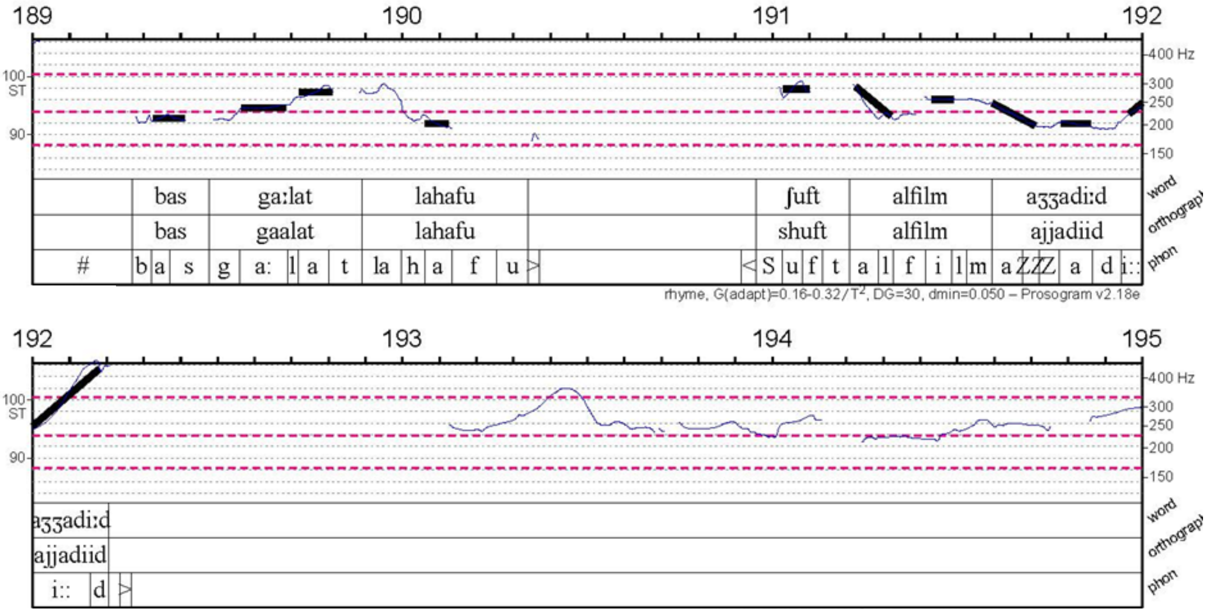
Chapter 4, declarative tune, 01_M



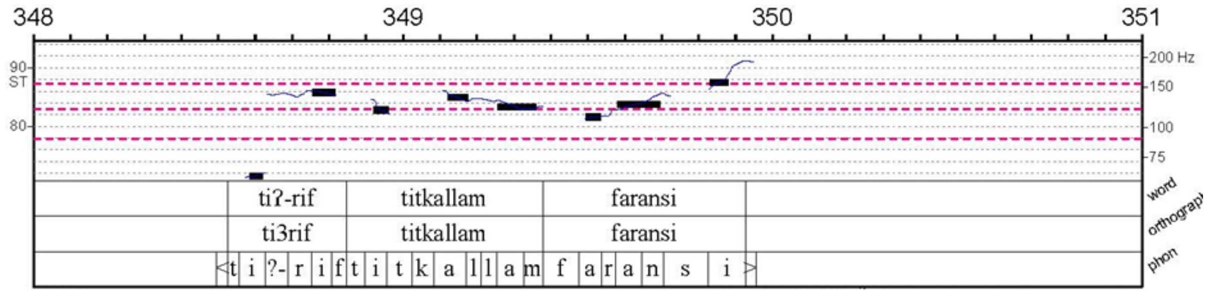
Chapter 4, declarative tune in narration, 03_M



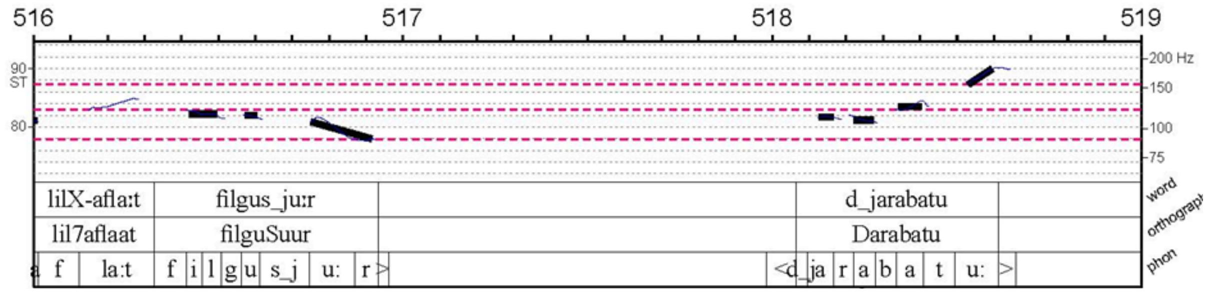
Chapter 4, yes/no question, 05_M



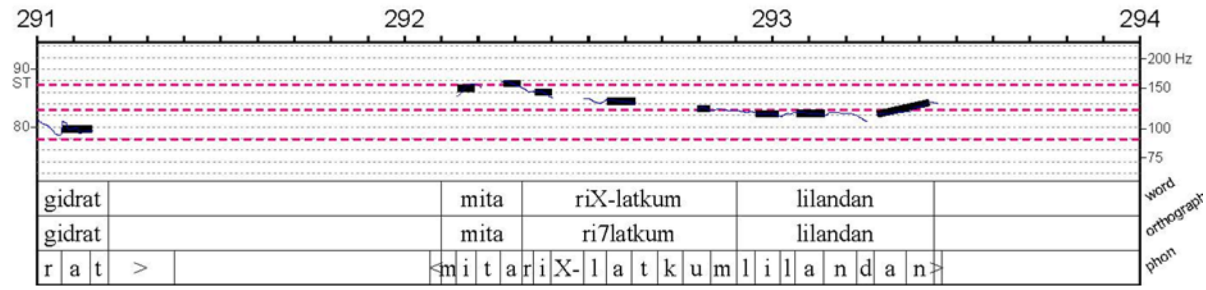
Chapter 4, yes/no question, 07_F



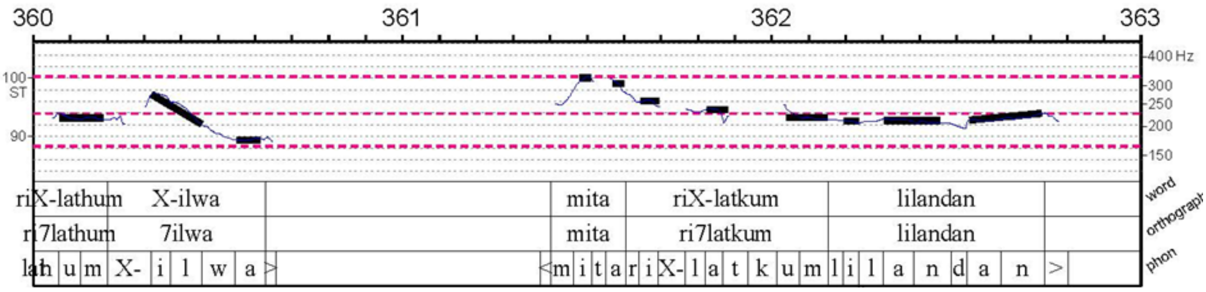
Chapter 4, yes/no question, 07_M



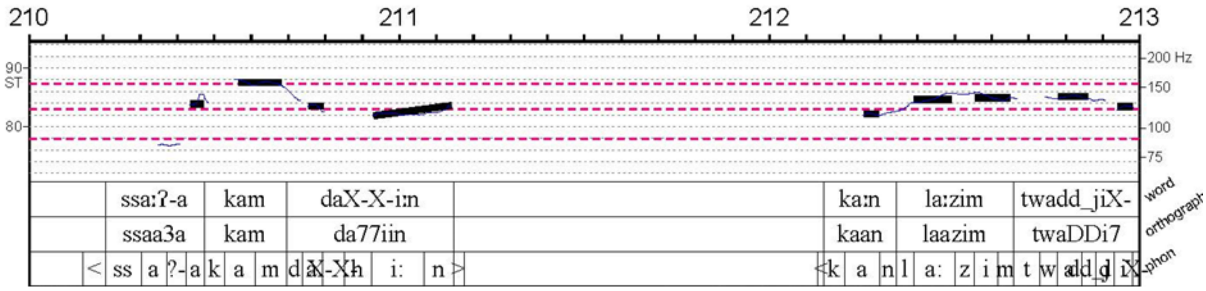
Chapter 4, yes/no question, 10_M



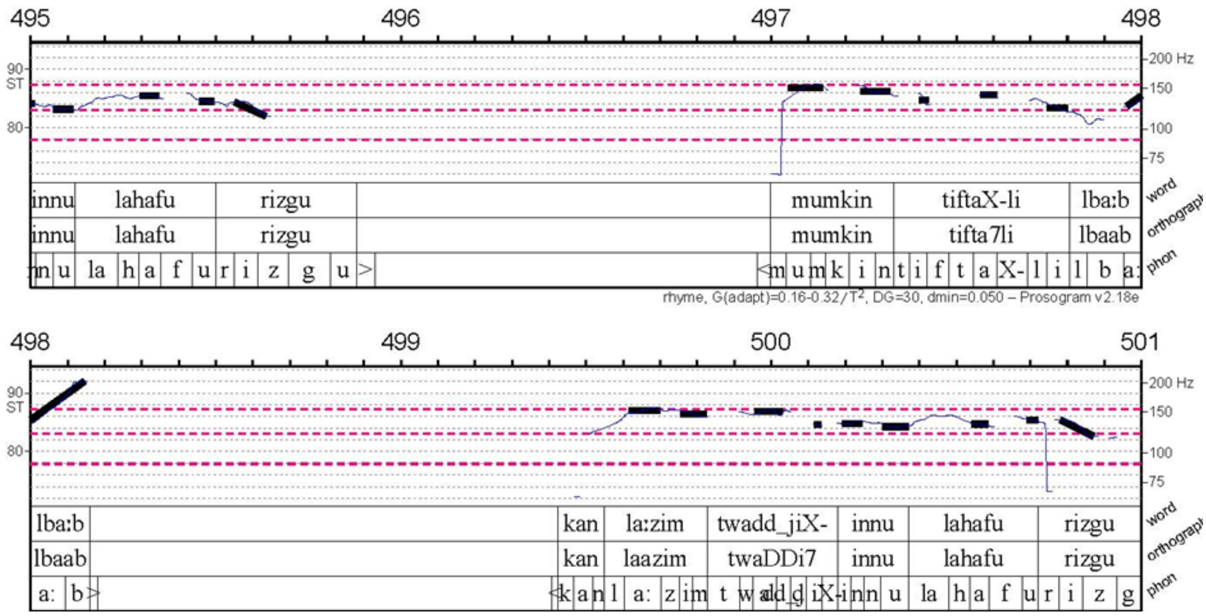
Chapter 4, WH question, 02_M



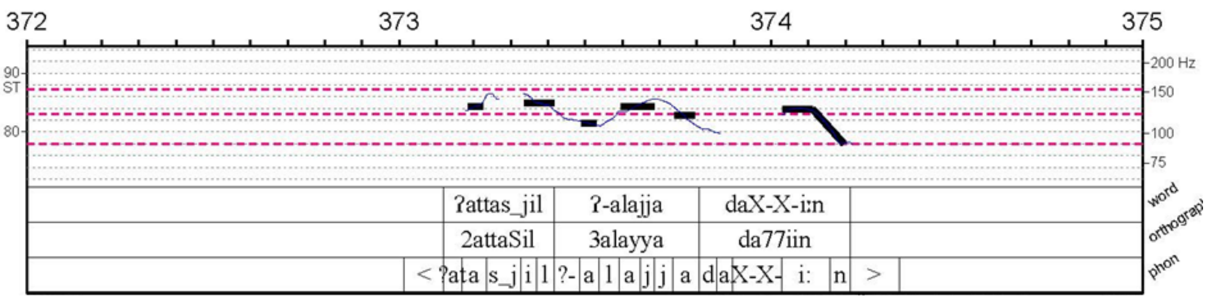
Chapter 4, WH question, 06_F



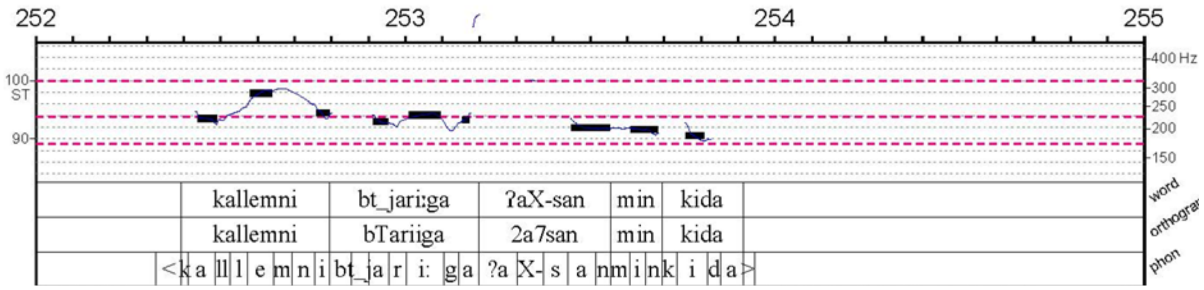
Chapter 4, WH question, 03_M



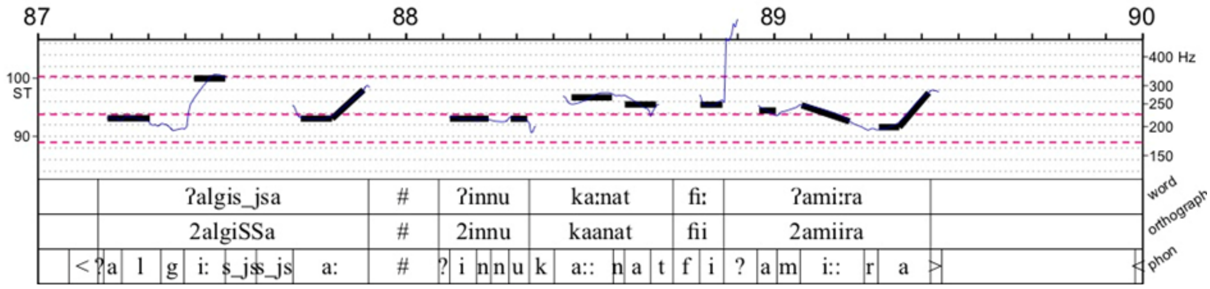
chapter 4, polite request, 10_M



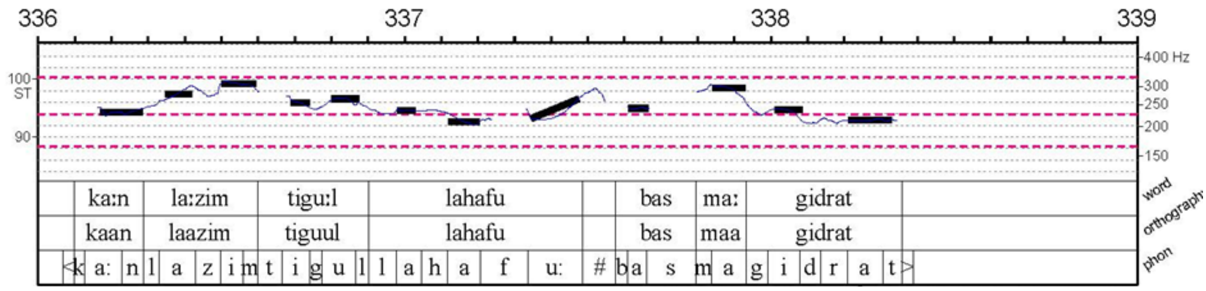
Chapter 4, imperative tune, 04_M



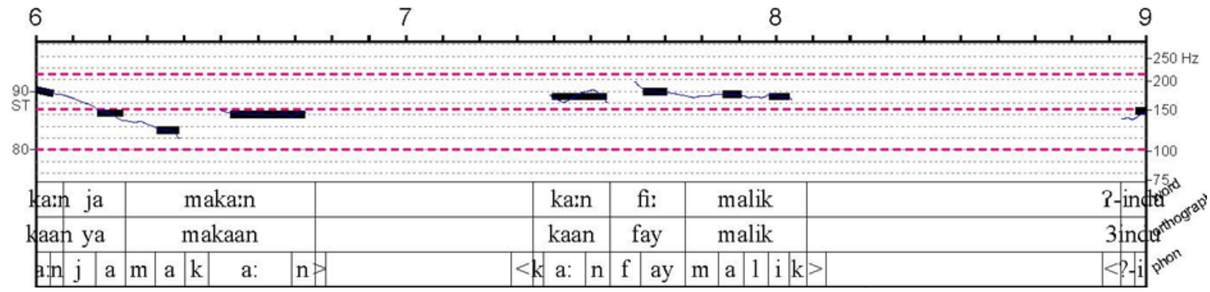
Chapter 4, imperative tune, 03_F



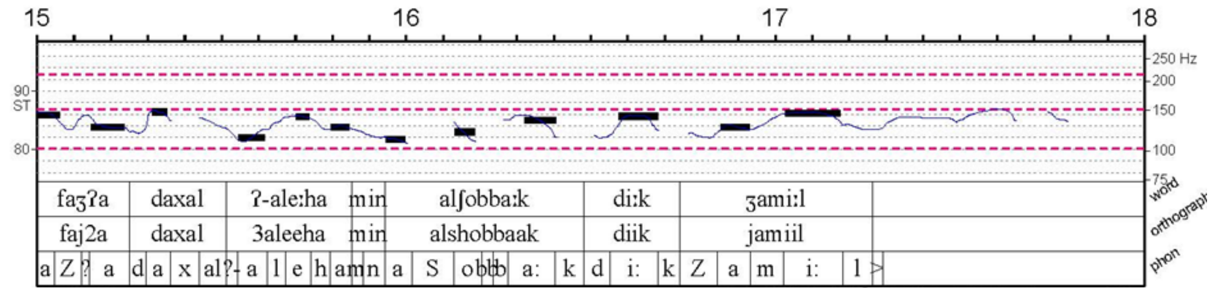
Chapter 4, rising plateau tune, 03_F



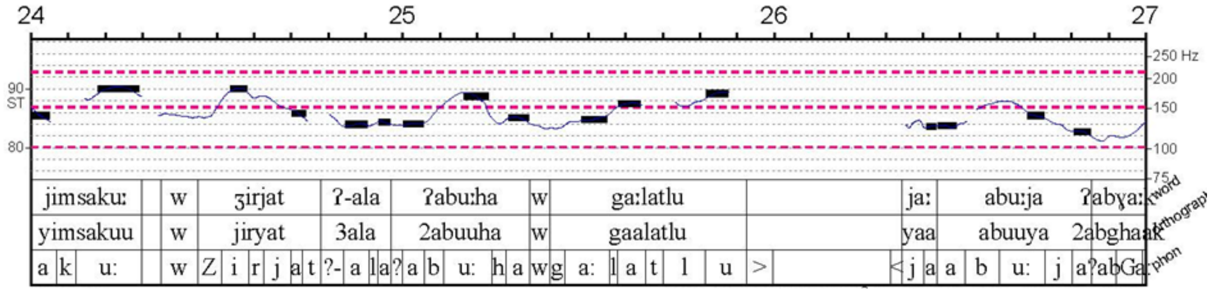
Chapter 4, falling plateau tune, 07_F



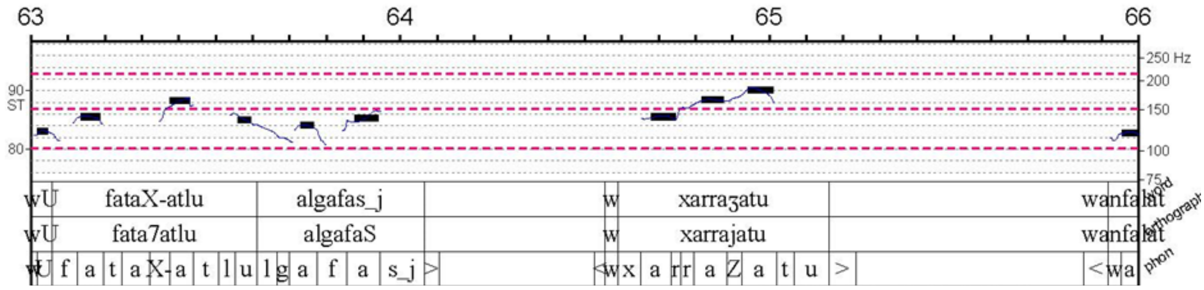
Chapter 4, falling plateau tune in narration, 09_M



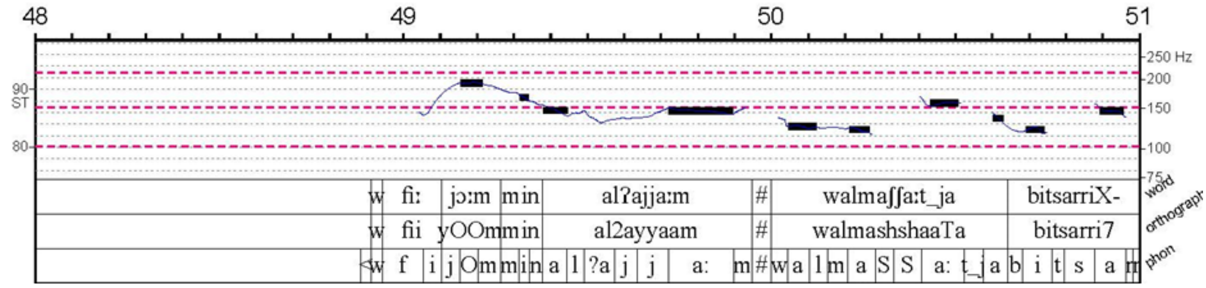
Chapter 4, falling plateau tune in narration, 03_M



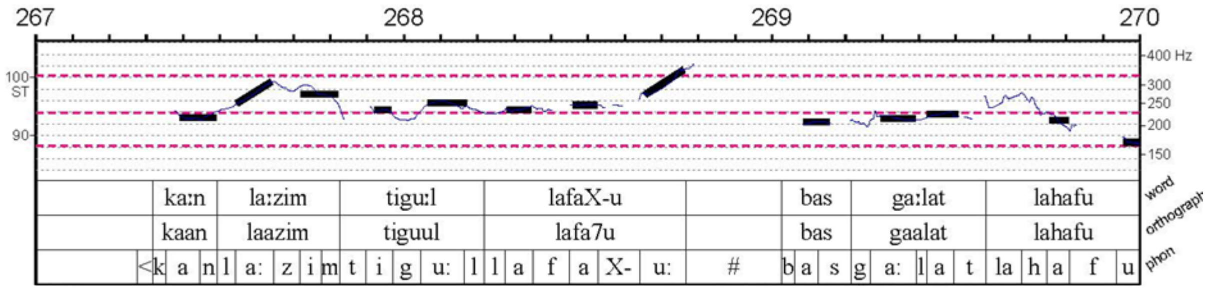
Chapter 4, rising continuation tune in narration, 05_M



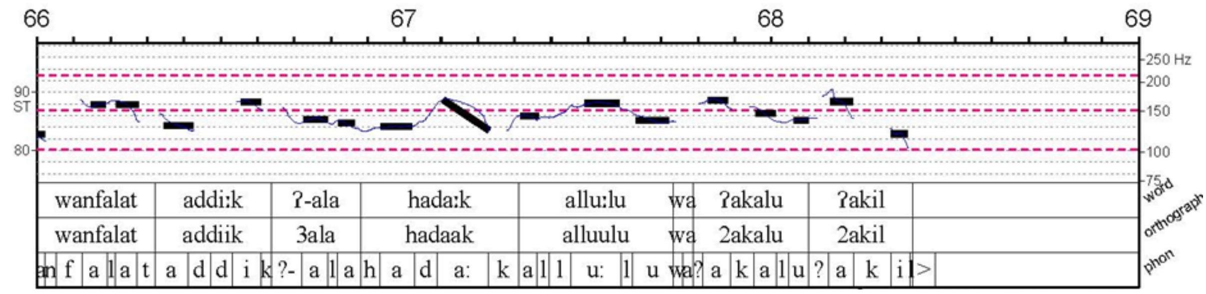
Chapter 4, rising continuation tune in narration, 09_F



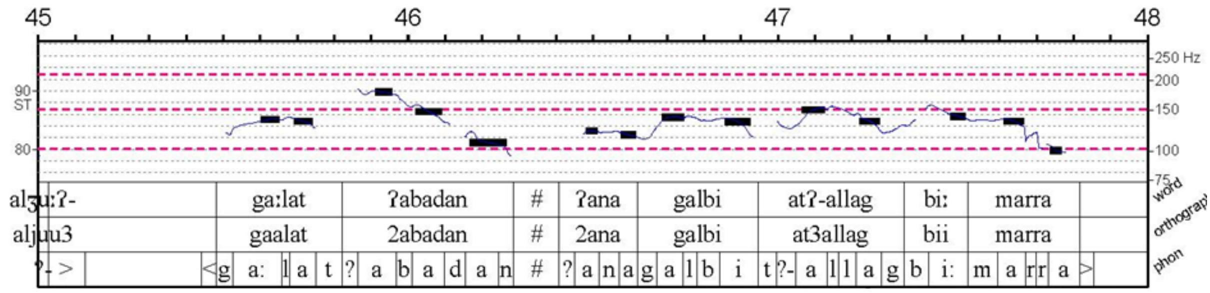
Chapter 4, plain and downstepped H*, !H*, 01_F



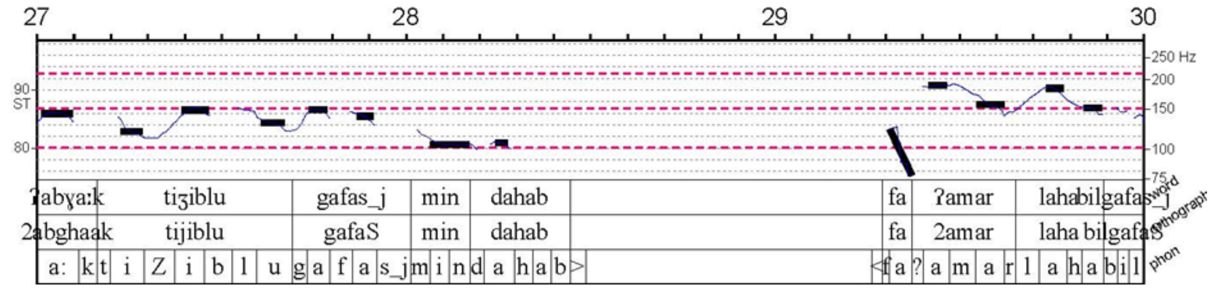
Chapter 4, L+H* accent, 08_F



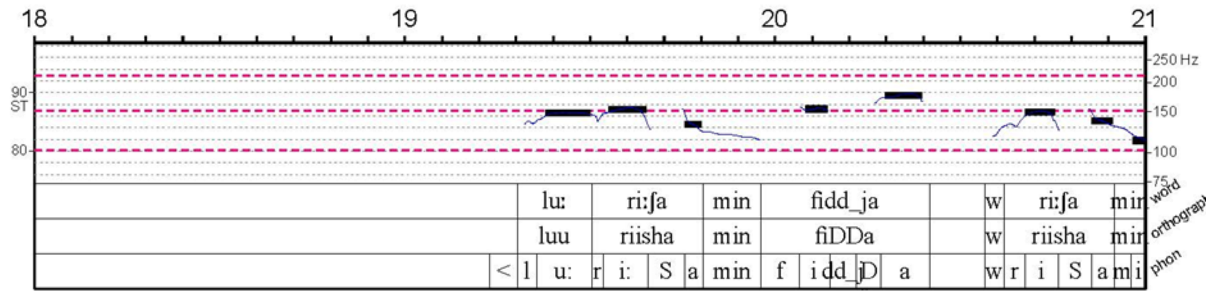
Chapter 4, downstepped L+!H* accent, 01_M



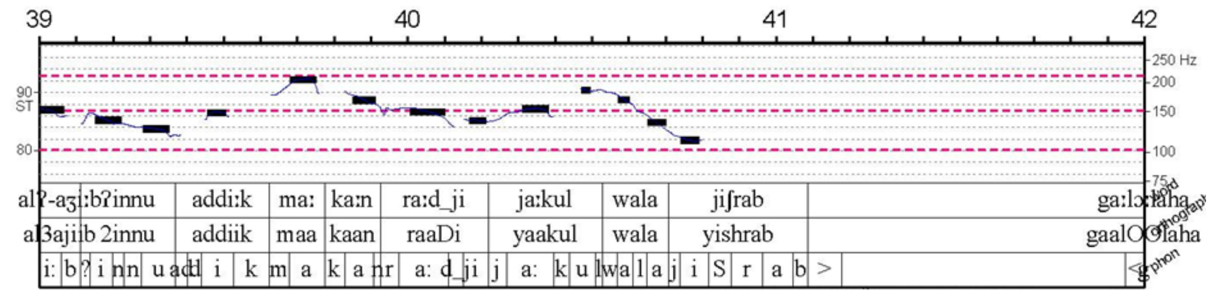
Chapter 4, L* accent, 05_F



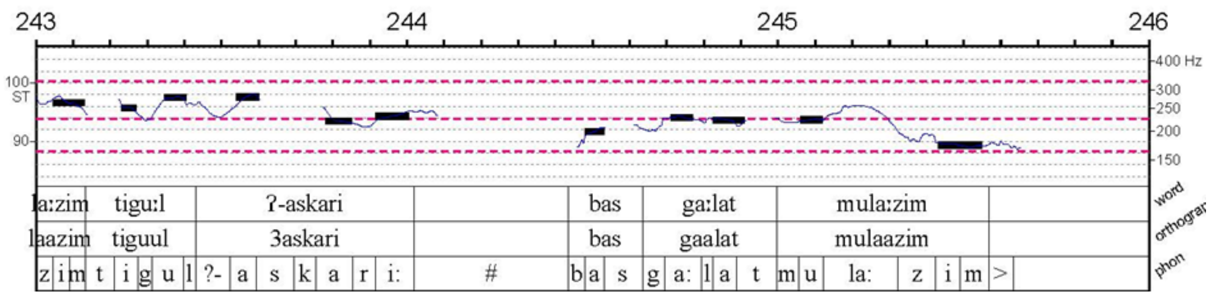
Chapter 4, L* accent, 03_M



Chapter 4, H- phrase tone, 09_M

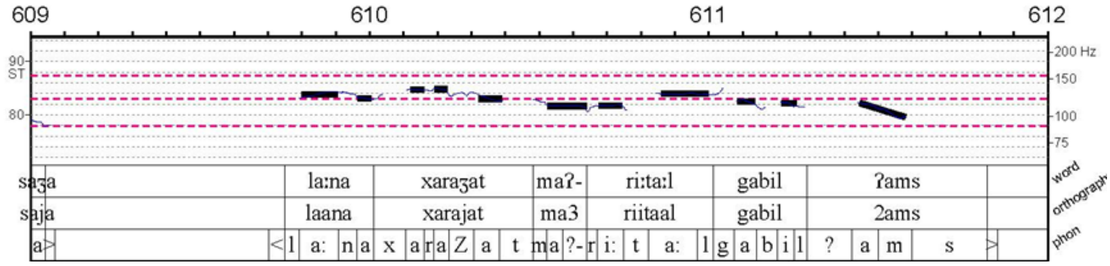


Chapter 4, H- phrase accent, 01_M

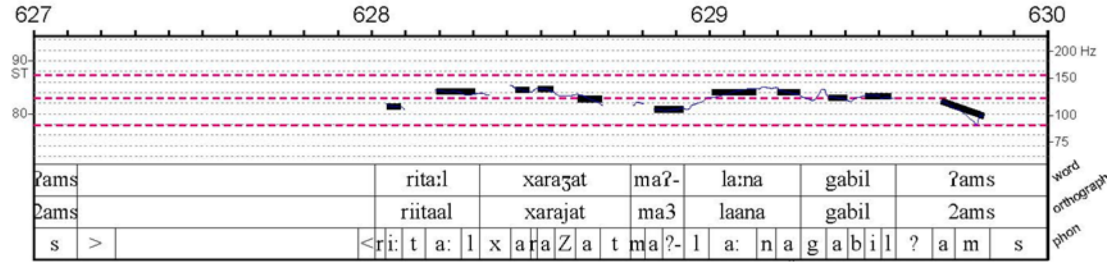


Chapter 4, downstepped phrase accent !H-, 07_F

Focus chapter:

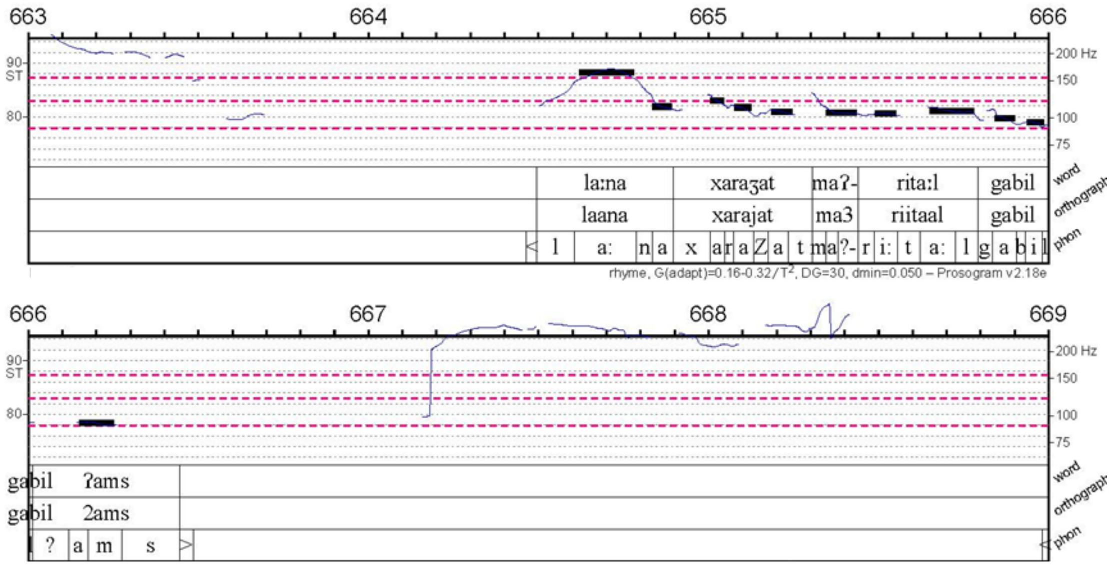


Chapter 5, broad focus initial, 04_M

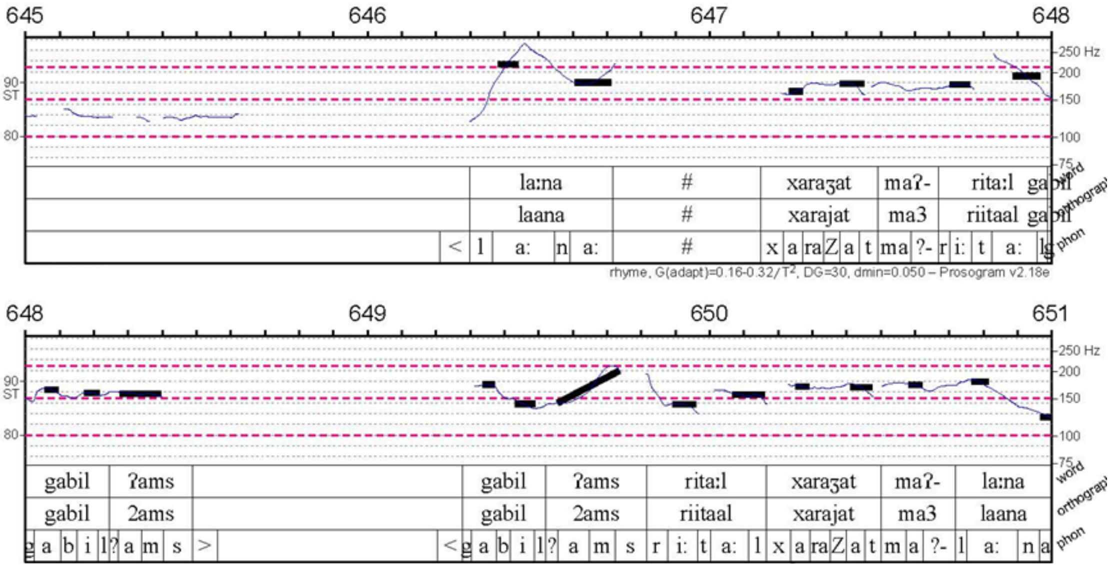


Chapter 5, broad focus medial, 02_M

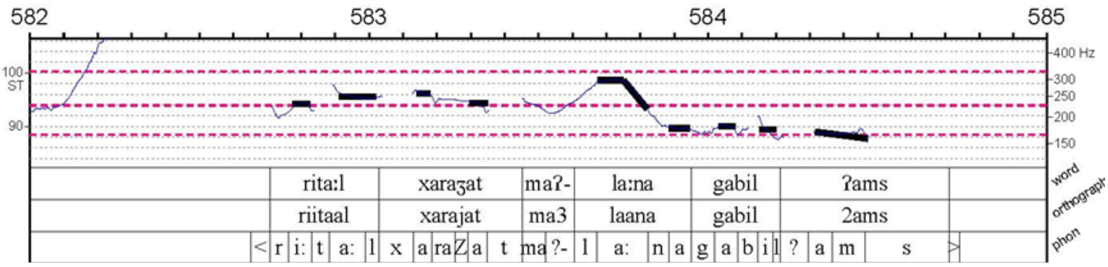
Chapter 5, broad focus final, 01_M



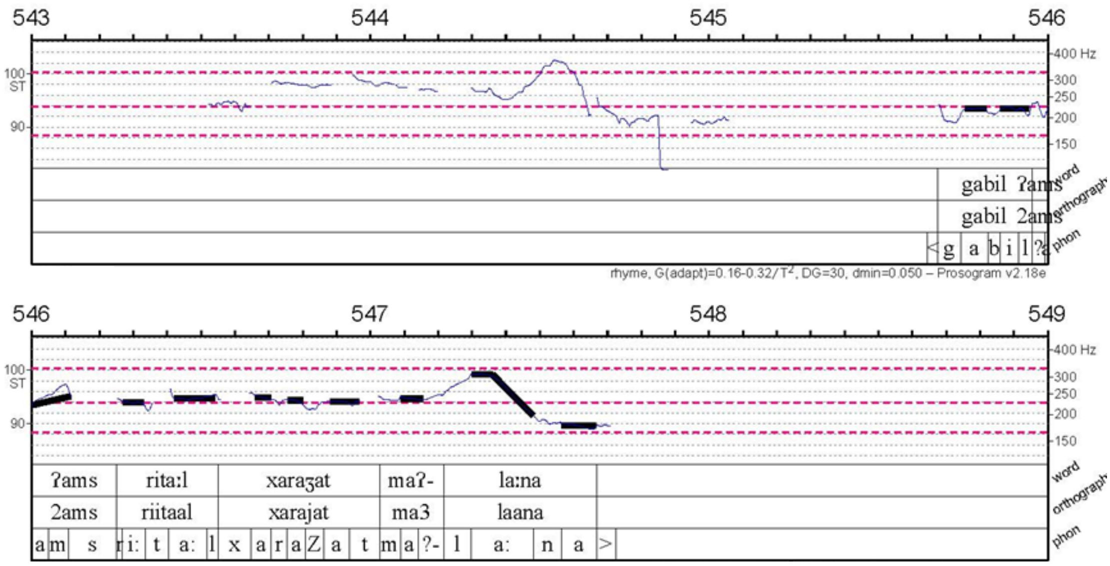
Chapter 5, narrow focus initial in one IP, 10_M



Chapter 5, narrow focus initial as two prosodic phrases, 01_M



Chapter 5, narrow focus medial, 09_F



Chapter 5, narrow focus final, 03_F