Lexical and postlexical prominence in Tashlhiyt Berber and Moroccan Arabic

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Philosophischen Fakultät der Universität zu Köln
im Fach Phonetik

vorgelegt von

Anna Bruggeman
aus Utrecht, Niederlande

Köln, 29. Januar 2018

Erste(r) Referent(in): Prof. Dr. Martine Grice
Zweite(r) Referent(in): Dr. Sam Hellmuth
Datum der mündlichen Prüfung: 16. Mai 2018

Please cite this thesis as:

Bruggeman, Anna. 2020. *Lexical and postlexical prominence in Tashlhiyt Berber and Moroccan Arabic*. University of Cologne PhD dissertation.

This version supersedes the unpublished version from 2018.

Abstract

Tashlhiyt Berber (Afro-Asiatic, Berber) and Moroccan Arabic (Afro-Asiatic, Semitic), two languages spoken in Morocco, have been in contact for over 1200 years. The influence of Berber languages on the lexicon and the segmental-phonological structure of Moroccan Arabic is well-documented, whereas possible similarities in the prosodic-phonological domain have not yet been addressed in detail.

This thesis brings together evidence from production and perception to bear on the question whether Tashlhiyt Berber and Moroccan Arabic also exhibit convergence in the domain of phonological prominence. Experimental results are interpreted as showing that neither language has lexical prominence asymmetries in the form of lexical stress. This lack of stress in Moroccan Arabic is unlike the undisputed presence of lexical stress in most other varieties of Arabic, which in turn suggests that this aspect of the phonology of Moroccan Arabic has resulted from contact with (Tashlhiyt) Berber.

A further, theoretical contribution is made with respect to the possible correspondence between lexical and postlexical prominence structure from a typological point of view. One of the tenets of the Autosegmental Metrical approach to intonation analysis holds that prominence-marking intonational events (pitch accents) associate with lexically stressed syllables. Exactly how prominence marking is achieved in languages that lack lexical stress is little-understood, and this thesis' discussion of postlexical prominence in Tashlhiyt Berber and Moroccan Arabic provides new insights that bear on this topic.

A first set of production experiments investigates, for both languages, if there are acoustic correlates to what some researchers have considered to be lexically stressed syllables. It is shown that neither language exhibits consistent acoustic enhancement of presumed stressed syllables relative to unstressed syllables.

The second set of production experiments reports on the prosodic characteristics of question word interrogatives in both languages. It is shown that question words are the locus of postlexical prominence-marking events that however do not exhibit association to a sub-lexical phonological unit.

A final perception experiment serves the goal of showing how native speakers of Tashlhiyt Berber and Moroccan Arabic deal with the encoding of a postlexical prominence contrast that is parasitic on a lexical prominence contrast. This is achieved by means of a 'stress deafness' experiment, the results of which show that speakers of neither language can reliably encode a lexically-specified prominence difference.

Results from all three types of experiment thus converge in suggesting that lexical prominence asymmetries are not specified in the phonology of either language.

Contents

Acknowledgments

First of all I thank the DAAD for their financial support for the past 3+ years.

There are very many people I am indebted to for helping me shape, start and complete this PhD.

I owe Martine Grice for a great deal more than I can express here. As a supervisor, and as Martine, she supported, questioned and challenged me. She has contributed enormously to how I have grown academically and personally. She identified and created learning opportunities at every turn, opened doors of whose existence I was not aware and made me steer clear from many potential hurdles. I am immensely grateful for everything that she has done for me on all levels.

Sam Hellmuth has been pivotal in getting me back on track when I needed it the most, as a supervisor and in all other possible capacities. I would not have managed if it had not been for her unwavering belief in me, her contagious enthusiasm, and for helping me see clearly through the forest of (metrical) trees. Not least of all, I owe her for entrusting me with her Moroccan Arabic data. This thesis would simply not have had its present form without her.

In Cambridge I owe much to Brechtje Post and Francis Nolan, who were exemplary teachers and motivated me to take the path into phonetics. Their continued interest and open doors during my PhD have been invaluable.

I am also grateful to Carlos Gussenhoven, Rachid Ridouane, Maarten Kossmann and Harry Stroomer for their input, suggestions and interest throughout.

There are a number of other people who have had a profound influence on me academically. Timo Roettger, from and with whom I have learned a great deal, and with whom working together on Tashlhiyt was always stimulating. Francesco Cangemi, for his readiness to share ideas, methodologies and complex thought processes. Stefan Baumann, for always making time when I needed input and for invariably being positive. Bodo Winter, for demystifying statistics before I even started my PhD. András Bárány belongs in this list too, for asking both obvious and obscure questions, and for showing what a great PhD could look like.

Many other people in the last few years have supported me in various ways.

From the I*f*L Phonetik in Cologne I want to thank Anne, Aviad, Bastian, Christian, Christine Ri., Christine Rö., Doris, Francesco, Henrik, Jane, Janina, Jessica, Luke, Leonie, Martina, Martine, Simon R., Simon W., Stefan, Tabea, Theo, Timo, and Phuong for everything from a discussion to a laugh, and from a chat (on request even in German!) to a dance. Martina, Christine and Tabea specifically for joint Muskelkaters, and Henrik, Aviad, the Simons, Luke and Francesco for providing daily dosages of entertainment and motivation.

In York a number of inspiring people helped shape the environment that brought me back on track at the start of my third year, including Sam Hellmuth and Tamar Keren-Portnoy who believed in me right from the start. I was especially lucky to meet Julia Kolkmann and Miriam Aguilar: I owe them both for keeping me sane and motivated at the same time. I also had the pleasure to meet Rana Almbark again, for Arabic discussion and joint elation about sugar.

Elsewhere I owe my (at times) relatively well-kept sanity to a number of good friends. Steffie, Maaike, Nicole and Maarten: Thanks for always being up for anything and providing the nicest kind of changing continuity. Andrea, things have been tumultuous for us in different ways, but I am glad we shared it. Camilla, Claudia B., Claudia N. and Laura: Thank you for providing me with shelter and distraction in various countries whenever I needed it.

A special thanks goes to all my contacts and participants in Morocco, and especially to Abderrahme Charki, Sanae Oubraim and Nabila Louriz. Thank you for welcoming me into your homes and lives. I simply could not have done this PhD without you. Additionally, learning from you (and with you) about your languages and culture has been one of the most fun and rewarding aspects of my PhD.

Aan mijn familie in Nederland: Dankjulliewel voor jullie ondersteuning en rotsvaste vertrouwen in me.

Finally, thanking András here does not do justice to the role he played in the process that led to this thesis – he lit my way.

Abbreviations

1 first person

2 second person

3 third person

ACC accusative

AOR aorist

CL clitic

COMP complementizer

DAT dative

DIST distal

EA état d’annexion (dependent state)

FUT future

F feminine

IMP imperative

IPFV imperfective

M masculine

NEG negative

PFV perfective

PL plural

POSS possessive

Q interrogative particle

REL relative

SG singular

AM Autosegmental Metrical (phonological)

AP Accentual Phrase

GLMM Generalized Linear Mixed Model

IP Intonational Phrase

IPO Instituut voor Perceptie-Onderzoek (‘Institute for Perception Research’)

IVAr Intonational Variation in Arabic

LRT Likelihood Ratio Test

MA Moroccan Arabic

MSA Modern Standard Arabic

PP Phonological Phrase

SRT Sequence Recall Task

ST semitones

TB Tashlhiyt Berber

TBU Tone Bearing Unit

Part I

Introduction

1 General introduction

1.1 Aims and goals of thesis

This thesis investigates and compares aspects of lexical and postlexical prominence structure in two Afro-Asiatic languages of Morocco, Tashlhiyt Berber (TB) and Moroccan Arabic (MA). The main goal is to find out how prominence structure in these languages should be characterised.

In doing so, two secondary goals can be identified. Firstly, this thesis serves as a detailed comparison of aspects of phonological prominence structure in two distantly related languages on different branches within the same language family. These languages have nevertheless been in contact for around 1200 years and are known to exhibit convergence in many aspects of linguistic structure. One of the additional goals therefore involves identifying similarities in these languages in the prosodic-phonological domain. Secondly, the present thesis will provide a contribution to the theoretical discussion about the mapping between lexical and postlexical prominence structure. In brief, prominence is conceived of as a phonological phenomenon with an abstract representation in the grammar, rather than a more surface-oriented understanding of prominence as acoustic-perceptual salience. A detailed discussion and a precise definition of 'prominence' as used in this thesis will be given in the next chapter. At the heart of the phonological definition of prominence is the question what structural linguistic elements are specified as such. This thesis will contribute insights into the specification of prominence at the *lexical* level, and will discuss how the realisation and distribution of *postlexical* prominences relates to these lexical prominence specifications.

Our current understanding of postlexical prominence structure in languages that lack lexical prominence is limited, mainly because there are only few languages that are convincingly argued to lack it. The best-known languages lacking any kind of lexical prominence are French and Korean, which are often considered to lack 'stress' (note that they also lack lexical tone and lexical pitch accent). Terminology is crucial here: Plenty of languages are considered to lack 'stress', such as Tokyo Japanese, but such languages may still have another type of lexical prominence specification (in the case of Japanese this is lexical pitch accent). Again, a detailed discussion of the terminology can be found in Chapter 2.

The two languages discussed in this thesis, TB and MA, form particularly interesting case studies for claims about prominence at both lexical and postlexical levels of phonological structure.

At the time of writing there appears to be consensus on the lack of lexical prominence in Tashlhiyt Berber (Stumme 1899; Dell & Elmedlaoui 2002; Kossmann 2012; Ridouane 2014; Roettger, Bruggeman & Grice 2015), although it has previously also been claimed that lexical stress is present (e.g. Sadiqi 1997; Gordon & Nafi 2012; Laoust 2012). For Moroccan Arabic, the details of lexical prominence are a matter of disagreement and remain as of yet unresolved (e.g Mitchell 1993; Benkirane 1998; Boudlal 2001; Watson 2011).[1]

While it is generally acknowledged that some languages lack *lexical* prominence structure, the question as to whether all languages have phonological prominence specifications at the *postlexical* level has been addressed to a lesser extent. At this point languages do seem to exist in which postlexical structural prominence does not exist or is unspecified, notably Ambonese Malay (Maskikit-Essed & Gussenhoven 2016) and Korean (Jun 1993, 2005a). Most of the literature however focuses on the apparent majority of languages that exhibit structural prominence at both levels. In languages with lexical stress, these levels are characterised by a clear correspondence, where postlexical prominence in the form of pitch accents co-occurs with lexically stressed syllables. In light of this, acknowledging the existence of languages lacking either or both levels of structural prominence is important, not only for purposes of prosodic typology, but also for our understanding of intonational structure in general, as most models have attempted to link lexical prominence structure to postlexical prominence. For present purposes, it should be noted that not much is known about the postlexical prominence structure in TB and MA, beyond general observations that suggest it is rather variable (Mitchell 1993 for MA, Dell & Elmedlaoui 2008 for TB). The investigation of the principles guiding the placement of postlexical prominence in these languages is therefore interesting in its own right already, but even more so by virtue of its being traditionally linked to lexical stress. In this thesis, therefore, the theoretical implications of a potential absence of a correspondence between the two levels of prominence structure will also be addressed.

The general claim that I will make in this thesis, based on the results from five experiments, is that lexical prominence asymmetries (in the form of lexical stress) are absent in both languages. In the context of the aforementioned goals (describing lexical and postlexical structure as well as any possible correspondences), the specific contributions of the different experimental chapters can be categorised as follows:

- Language-specific description of prosodic-phonological structure;
  - Lexical–phonological structure: Acoustic correlates of lexical stress (Chapter 3 for TB and Chapter 4 for MA)
  - Postlexical–phonological structure: Prosodic correlates of question word interrogative intonation (wh-questions) (Chapter 6 for TB and Chapter 7 for MA)

[1]As this thesis is officially published in 2020, it should be clarified that Chapter 3 is a more elaborate analysis of the same experiment that has previously been published as Roettger, Bruggeman & Grice (2015) and Roettger (2017). Similarly, the claim that stress is absent in Moroccan Arabic has in the meantime been published in Bruggeman et al. (in press) (based on Chapter 4).

- The influence of native phonology on the perception of prosodic prominence (Chapter 8 for both languages)

The specific questions and theoretical background motivating each of the experiments will be discussed in more detail in the next chapter (Chapter 2: Approaches to lexical and postlexical prominence).

1.2 The linguistic landscape of Morocco

Morocco is a highly multilingual country, with many of its inhabitants (total as of 2017: ca. 36 million) fluent in both Moroccan Arabic (Darija) and some variety of Berber. The number of individual speakers of Moroccan Arabic for 2014 is estimated at around 25 million. The number of speakers of Tashlhiyt Berber speakers was 3.9 million in 2004 (Simons & Fennig 2017). Many native speakers of Berber have at least some competence in Moroccan Arabic and many are fluent bilinguals, especially those who have received at least secondary education. The reverse, with first-language speakers of Moroccan Arabic being fluent in Berber, is rather uncommon.

The institutional languages of Morocco complicate the picture further as these include neither Moroccan Arabic nor any variety of Berber that is actually spoken. The official languages of Morocco are Modern Standard Arabic (MSA) (since 1961) and Standard Berber (since 2011). MSA is the supranational variety of Arabic that is used throughout the Arabic-speaking world but is never acquired as a mother tongue. In Morocco it is taught in schools to children from around the age of 6, and Moroccans have exposure to it on international radio and TV, as well as in all formal written communication. Some 14.5 million Moroccans are reported to be competent in Modern Standard Arabic (Simons & Fennig 2017). Standard Berber, or Standard Tamazight Berber, is a standardised variety of Berber rather close to the variety of Berber spoken in the Central Atlas, but like MSA it is not acquired as a first language. As of 2011 it is being taught in (primary) schools.

While Moroccan Arabic is not an official language, it does function as the "[d]e facto national working language" (Simons & Fennig 2017). Additionally, among the 'foreign' languages, French has had a privileged status ever since the start of French (and Spanish) protectorates in the early 20th century. After Morocco's independence (1956), the use of French has been in decline. Still, as of today it remains a semi-official language used in administration, commerce and the like, and it is the first go-to foreign language (Maas & Procházka 2012a). Spanish as a foreign language does not quite enjoy the same status, although it is still relatively popular in the north.

Despite the recent institutionalisation of Berber, not only in Morocco but also in other parts of North Africa, Berber is still rather marginalised. The socio-historical reasons that have led to Berber being in this position have their roots in the religious expansion of Islam, whose spread from its origin in the Arabian Peninsula from the 7th century onwards entailed the reach of the Arabic language and culture as far west as Morocco. More detailed overviews of the expansion of Arabic can be found in Holes

(2004: Ch. 1) and Versteegh (2014). For specific reference to the Moroccan context see Maas & Procházka (2012b: Sec. 3), Mitchell (1993) and El Aissati (2005).

In the following section, I will review claims about convergence between Moroccan Arabic and Tashlhiyt Berber in light of the wider context of variation in Arabic and Berber.

1.3 Tashlhiyt Berber and Moroccan Arabic

The linguistic classification of Berber has been subject to some debate, but most researchers now agree that it is a branch of the Afro-Asiatic phylum, sister to Semitic and Cushitic, among others (Simons & Fennig 2017). Moroccan varieties within the Berber branch are subclassified as Northern. Arabic on the other hand belongs to the Semitic languages. Within Arabic, Moroccan Arabic is subclassified as Western (Colloquial).

Currently, Berber languages are spoken in North Africa from its easternmost point in Egypt to its westernmost point in Morocco, while Arabic is natively spoken as far east as Oman, to the north in Iraq and to the south in Sudan. Unsurprisingly, both languages are subject to an enormous amount of regional variability, to the extent that regional varieties of neither Arabic nor Berber are typically mutually intelligible with other varieties of Arabic or Berber.

Of the three main Berber languages spoken in Morocco (Tarifit, Tamazight and Tashlhiyt), Tashlhiyt Berber is the most well-researched (e.g. Stumme 1899; Aspinion 1953; Applegate 1958; Sadiqi 1997; Van den Boogert 1997; Laoust 2012).[2] Some smaller Moroccan Berber varieties have also been studied, including Figuig (Kossmann 1996) and Ghomara (Mourigh 2015). The three main varieties spoken in Morocco are characterised by some mutual intelligibility between Tamazight and Tashlhiyt, and very limited comprehensibility between these and Tarifit. Despite considerable regional variation, good mutual comprehensibility exists within a single variety (Dell & Elmedlaoui 2002).[3] Outside of Morocco, several other varieties of Berber have been subject to research, including Tamasheq ((Sub-)Saharan, Heath 2005; Lux 2014), Kabyle (Algeria, Mettouchi 2008), Zwara (Libya, Gussenhoven 2017) and Siwi (Egypt, Vycichl 2005; Naumann 2012).

Among Berber varieties, the segmental–phonological structure of Tashlhiyt especially has been subject to considerable interest from the linguistic community by reason of its crosslinguistically unusual tolerance of long consonantal sequences (Dell & Elmedlaoui 1985; Coleman 1999, 2001; Dell & Elmedlaoui 2002, 2008; Ridouane 2008, 2014).

[2]For the uninitiated reader, it is not always clear which variety of Berber is being discussed in a given source due to confusing terminology. Tashlhiyt is known by various alternative spellings and names, including Tachelhit(e), Shilha, Soussiya, or Tasoussit. The term Tamazight, in turn, has multiple uses, as it may refer to the variety of Berber spoken in the Atlas as well as to the standardised version of Berber taught in schools. Tamazight might also be used synonymous with Berber and denote the language family in general.

[3]Maas & Procházka (2012b: 330) go as far as to suggest that different varieties of Berber within a village are spoken by individual families.

There has been a long-standing debate about how these sequences should be best represented, with notable disagreement about the status of the central vocoids that are observed to break up long clusters, as either having phonological status or being phonetic transitions.

More recently, experimental work on Tashlhiyt intonation has been conducted, focusing on yes–no question intonation and phrase-final focus (Grice, Ridouane & Roettger 2015; Roettger 2017). Some more impressionistic reference to topic and focus structures in Tashlhiyt can be found in Lafkioui (2010). This thesis adds to existing work on Tashlhiyt intonation by discussing intonation in wh-questions, hitherto undiscussed, in Chapter 6.

While there is thus some past work on Berber, existing research on Arabic has vast dimensions (Classical) Arabic has a long research tradition, but this has been supplemented over the last century or two by work following from specific interest in the variability between synchronic, spoken varieties. The following resources, and references therein, form a good starting point into the literature: Bateson (1967), Holes (2004), Owens (2013) and Versteegh (2014). Contemporary general descriptions of Moroccan Arabic can be found in Harrell (1962) and Maas (n.d.) and Maas & Procházka (2012b). Aspects of MA phonology are addressed in Mitchell (1993) and Heath (1997) and Dell & Elmedlaoui (2002). Comparatively recently, some work has been done on intonation, with a general description in Benkirane (1998), and experimental studies by Yeou (2005), Yeou et al. (2007), Yeou, Embarki & Al-Maqtari (2007), Burdin et al. (2015) and Hellmuth et al. (2015). See Chapter 7 for more a detailed discussion of the existing literature on MA intonation.

It has been suggested that there is no common standard Moroccan Arabic yet, at least not in terms of phonology, although if anything there is a trend towards a local standard variety in the metropolitan areas of Casablanca and Rabat (Dell & Elmedlaoui 2002: 239). Nevertheless, is well known that in general, the segmental phonology of MA has been heavily influenced by contact with Berber (Mitchell 1993; Heath 1997; Maas & Procházka 2012b; Maas n.d., see also Zellou 2010). In fact, MA and (Tashlhiyt) Berber have been said to exhibit similar "surface phonologies" (Dell & Elmedlaoui 2002: 227). For example, the vowel inventory of most contemporary Arabic varieties typically consists of the phonological vowels /i, a, u/ with contrastive use of length (Watson 2011), whereas MA lacks a phonological vowel length distinction (see also Chapter 4). Similarly, MA displays many complex consonantal clusters that are not observed in other varieties of Arabic. Maas & Procházka (2012b) discuss several more phonological properties that are shared between MA and Berber, including prosodic aspects such as what they call "accent" (lexical stress).

Especially this reference to similarities in lexical prominence structure is of interest to this thesis. As previously mentioned, the existence of lexical prominence specifications in the form of stress has been denied for Tashlhiyt, and is subject to debate for Moroccan Arabic. Additionally, the location of postlexical prominence in both languages is highly elusive, along the lines of the following statement about Moroccan Arabic: "Prominence among syllables in MA words is at present imponderable and seemingly lacks any close relationship with prominence in sentences" (Mitchell 1993: 199).

Although in recent years, several studies on intonation and postlexical prominence structure have been carried out for both languages, no direct comparison has so far been attempted of TB and MA in terms of prominence at either the lexical or postlexical levels. This thesis will fill this gap by means of five experiments that test similarities between the languages in terms of (prosodic–)phonological prominence at both the lexical level (stress) and the postlexical level (intonation).

1.4 Overview of thesis

The next chapter, Chapter 2 (part of Part I: Introduction), lays out the theoretical preliminaries for the investigation of lexical and postlexical prominence. It also lays out the research questions of this thesis. The subsequent chapters are dedicated to five experiments that serve to analyse the prominence structure of both languages.

The first experimental part, Part II: Lexical prominence production, is concerned with the acoustic correlates of (presumed) lexical prominence in the form of lexical stress. Individual experiments were conducted for Tashlhiyt Berber (Chapter 3) and Moroccan Arabic (Chapter 4). It will be concluded that the experimental results are compatible with the interpretation that stress is lacking in both languages.

The second experimental part, Part III: Postlexical prominence production, starts off with an introductory chapter that serves to motivate the choice to look at question word interrogatives in order to research postlexical prominence (Chapter 5). The experiments in this part serve to identify the prosodic realisation of postlexical prominence as observed in question word questions (Chapter 6 for Tashlhiyt Berber and Chapter 7 for Moroccan Arabic). Again, both chapters show that postlexical prominence in the form of a prominent pitch event exhibits a type and degree of variability in realisation that is compatible with the absence of lexical anchors in the form of stress.

Part IV: Prominence perception, the final experimental section, contains the last experiment (Chapter 8). This chapter investigates the perception of prosodic prominence by speakers of both languages by means of a short term memory task often referred to as testing 'stress deafness'. The results of this experiment highlight that speakers of both languages show insensitivity to prominence asymmetries at the postlexical level. This is interpreted again as following from a lack of prominence asymmetries specified at the lexical level, and is taken to further support the idea that lexical prominence specifications are absent in both Tashlhiyt Berber and Moroccan Arabic.

Finally, Part V: Conclusion, contains the final chapter of this thesis (Chapter 9) which consists of a summary and general discussion.

2 Approaches to lexical and postlexical prominence

2.1 Introduction

This chapter serves to lay out the theoretical groundwork relating to prominence, lexical and postlexical phonological structure, and the assumptions that underlie the questions asked in this thesis.

Section 2.2 will begin with a few basic definitions, namely those of lexical and postlexical prominence, as used throughout the rest of this thesis. In subsequent sections, the literature on each of these will be discussed in further detail. Section 2.3, entitled 'Lexical prominence: Word stress', will look at phonological prominence at the word level and will review uses of the term 'stress'. The next section, 2.4, entitled 'Postlexical prominence and intonation' continues the discussion of prominence asymmetries at a higher level of linguistic structure (i.e. above the word). This section will involve an overview of the place of postlexical prominence in various models used for the analysis of intonation. It includes an overview of what is currently the most standard theory used in the modelling of intonational structure; AM phonology.

In Section 2.5, I will bring together the discussions of lexical and postlexical prominence and highlight severak topics related to their interaction that are especially relevant to this thesis.

Section 2.6 serves to link the discussion of the current linguistic landscape of Morocco in the previous chapter with the theoretical issues highlighted in the present one. This section will motivate the overarching research questions of this thesis as well as the individual research questions for each experimental chapter. I will also briefly sketch how hypothesised findings may fit in with various typological approaches. This topic is taken up again in Chapter 9.

2.2 Defining prosody/intonation, lexical and postlexical prominence

Prosody will be defined here as all suprasegmental aspects of speech, including the acoustic properties of intensity (perceptual correlate being relative loudness), F0 (pitch), duration (length) and various other acoustic–phonetic properties of speech that can vary, to a large extent, independently of the segmental material with which they co-occur (e.g. Lehiste 1970; Cruttenden 1986). Prosody simply refers to acoustically measurable suprasegmental properties of speech, and the use of the term therefore does not presuppose anything about their contribution to categorical or linguistic structure. The approach taken in this thesis, however, is that prosodic aspects of speech can (and should) be considered in any categorical description of intonation.

The term **intonation** is closely related to prosody but often contains reference to categorical analysis. I am adopting one of the most common definitions here: "The use of *suprasegmental* phonetic features to convey 'postlexical' or *sentence-level* pragmatic meanings in a *linguistically structured* way" (Ladd 2008: 4). Sometimes a distinction is made between intonation in a narrow sense, referring to pitch/F0 only, and intonation in a broad sense, referring to all thinkable prosodic–phonetic dimensions, as in the above definition. While I accept the broader definition, I will nevertheless be concerned mainly with pitch/F0. This is motivated by the primary role that F0 plays in signalling postlexical prominence, rather than any preconceived notions about what does or does not constitute intonation.

The term prominence is used differently in various domains of linguistics. On the one hand, speech elements as small as a single segment may be considered prominent by virtue of acoustic enhancement. On the other end of the scale, linguistic units consisting of several words may be considered prominent within the global discourse context (notably, focused constituents). In the context of this thesis, **prominence** will be used to designate elements that are *phonologically* prominent either at the level of the lexical word, i.e. **lexical prominence**, or at the next higher structural level of description, i.e. phrasal or **postlexical prominence**. The most crucial aspect of its definition here is that it is used mainly to refer to a categorical, paradigmatically or syntagmatically contrastive property. For the notion of 'standing out' without theoretical connotations, different terms will be used, as will be detailed in the following sections.

2.3 Lexical prominence: Word stress

2.3.1 Defining stress

There are various kinds of lexical phonological structure that have been studied, including segmental, suprasegmental (prosodic), and prominence structure. With respect to prominence structure, a common distinction is made between those languages that have lexical stress, those that have lexical tone and those that have lexical pitch accent (Beckman 1986). These properties may moreover occur in various combinations in single languages, or may be absent altogether (Jun 2005b). In this thesis I will be concerned with lexical stress and the absence thereof in TB and MA, as previous descriptions have made clear that neither language has lexical tone and/or lexical pitch accent. The question of lexical stress in these languages on the other hand is rather more controversial (see Chapter 1).

I define **stress** as lexical stress or word stress, and more specifically as the abstract phonological property that makes one syllable (or mora) different from the other ones in that word: It is 'culminative', and as such syntagmatically contrastive. Stress is moreover obligatory: It applies to each polysyllabic word in a given language that has lexical stress (or to none of the words if the language in question lacks stress). Finally, the position of stress in a given word is fixed and invariant, something which would be reflected in the dictionary entry of a word (cf. Abercrombie 1976 in van der Hulst 2014a).

So far, the present definition of stress would cover both lexical stress and lexical pitch accent systems.[1] In this context, a relevant distinction is often made between 'stress–accent' and 'non-stress–accent' (cf. Hyman 1977; Beckman 1986). This distinction refers to the observation that lexical stress asymmetries result from acoustic enhancement in terms of multiple phonetic parameters ('stress–accent'), whereas lexical pitch accent asymmetries result from pitch only.[2] Based on the available evidence about correlates of stress, however, there seem to be languages that have stress but nevertheless exhibit little or no acoustic enhancement of stressed syllables (see Section 2.3.3). It is desirable therefore to define lexical stress without reference to the acoustic properties that *may* mark it, and instead define it to the exclusion of lexical prominence asymmetries that solely involve lexical pitch accent. This way, languages like Tokyo Japanese can be considered to have lexical prominence in the form of pitch accent but not stress, and languages like Swedish and Norwegian have lexical prominence asymmetries in the form of stress as well as in the form of pitch accent. In essence, then, lexical stress is taken to refer to the property of lexically specified prominence asymmetries that does not exclusively involve the lexical marking of pitch.

The present view of stress is similar to the definitions found in Ladd (2008: 50f.): "an abstract phonological property of a syllable within a prosodic structure") and in Hyman (2014: 56): "the phonological marking of one most prominent position in a word" (note that Hyman uses the term 'accent' to denote this property).[3] This definition is also compatible with the one given on WALS (Goedemans & van der Hulst 2013).

The present definition however contrasts with two other commonly used definitions. The first one is the understanding of stress in the literature on Metrical Stress Theory, as in Hayes (1995: 8): "stress is the linguistic manifestation of rhythmic structure". This definition of stress overlaps to a large extent with the present one in the sense that it usually identifies the same syllable as stressed, and uses some of the same diagnostics (see Section 2.3.2). The present definition crucially differs from it in renouncing the idea

[1] In fact, evidence from stress deafness experiments (e.g. Rahmani, Rietveld & Gussenhoven 2015) suggests that it is indeed appropriate to group these systems together for the purpose of lexical–phonological representation.

[2] Even many of those who use the term 'stress–accent' would not go as far as to suggest that the only way of identifying stress is through its acoustic enhancement. However, since acoustic enhancement is so often misanalysed, and there are plenty of alternative ways to identify stress, it seems safer not to use it as part of a definition. If one does want to define a class of languages based on the presence of 'stress–accent', as in Jun (2005b: 440): "A language is categorized to have a 'stress–accent' feature if a certain syllable in a word is more prominent than other syllables by duration and/or amplitude", several languages that are normally thought to have stress but do not exhibit such enhancement would have to be reconsidered.

[3] The main reason why I use the term 'lexical stress' rather than 'lexical accent' is to avoid confusion with the term 'accent' as used in intonational research, where it denotes concrete, measurable, pitch protrusions.

that stress can be identified conclusively by appealing to non-native speaker linguists' perception of prominence. Why this particular way of identifying stress (by means of a non-native ear) is highly problematic will be discussed in more detail below. Moreover, 'stress' in Metrical Phonology is sometimes used to refer to levels of (perceived) prominence above the word, with the term being used for whichever culminating node at a given level of the metrical tree or grid (Liberman & Prince 1977; Hayes 1995). This results in the use of the word stress in collocations such as 'phrasal stress', which are ruled out under the definition adopted in this thesis. A concrete problem with this kind of terminology is seen in the case of French, which is sometimes described as involving 'final stress'. While this might be correct depending on one's definition of stress, such claims are likely to cause confusion as they miss to specify the crucial point that instead of the final syllable of the word, it is the final syllable of a small phrase (an Accentual Phrase, or AP, or a Phonological Phrase, or PP) which is phonologically specified as prominent (Post 2000; Jun & Fougeron 2002). French has inherent prominence asymmetries at the *postlexical*, not the lexical level, and therefore lacks (lexical) stress. The present definition reserves the term stress for lexical–structural prominence only and would classify French as a language lacking (lexical) stress.

A second alternative definition of stress is one that takes it to mean 'perceptually prominent' or 'acoustically prominent'. This understanding was mentioned previously in the discussion of 'stress–accent' and 'non-stress– accent'. This particular use of the term is rather common, especially in earlier work (e.g. Abercrombie 1976 in van der Hulst 2014a).[4] One often finds it used in the collocation 'sentence stress' where it refers to what is perceptually the most prominent word or syllable in a phrase. For the purposes of this thesis, wherever the distinction between acoustic-perceptual prominence and lexical stress as a more abstract notion is important (i.e. repeatedly) I will refer to the former as 'pitch prominent', 'durationally enhanced' or similar, as appropriate.

2.3.2 Identifying stress

The identification of stressed syllables is a complex undertaking, with as of yet no agreed standard on what counts as a reliable diagnostic of stress. In many works, no clear definition is given, making it unclear whether 'stress' refers to phonological lexical stress, (postlexical) pitch accent, or general acoustic enhancement. Criteria for the identification of stress in a given source are also rarely given, especially in work within the Metrical Phonology tradition.[5]

Nevertheless, a number of diagnostics have made repeated appearances in the literature on stress description and typology:

[4] It is an independent question if these authors believe in the existence of a more abstract level of lexical prominence; Abercrombie for example did, and called it 'accent'.

[5] It can often be deduced that stress position is determined based on perceived prominence (by non-native speaker linguists). This seems to be the case even in Hayes (1995), which despite featuring a list of possible diagnostics to stress, cites many sources for which the process involved in identifying stress is unclear.

1. The attraction of postlexical intonational events (Bolinger 1958; Lehiste 1970; Bolinger 1986; Hayes 1995; Hyman 2014). This diagnostic identifies a syllable of a word as stressed when it attracts postlexical prominence in the form of a (nuclear) pitch accent when it is spoken in isolation (see also Section 2.4). As a result, the pitch accented syllable is auditorily the most prominent in the word. This diagnostic is in all likelihood responsible for most claims about stress in languages investigated by non-native speaker linguists, at least prior to the widespread availability of recording equipment.[6]

2. Segmental-phonological properties and phonological processes (or 'rules') (Hayes 1995; Hyman 2014). This diagnostic holds that stressed syllables are special by virtue of their interaction with other aspects of phonology. For example, in many languages only specific syllables may exhibit the full vowel set contrast or the full set of allophones for a given phonemic category. Similarly, some phonological processes apply differently to different syllables, contrasting syllables in terms their phonological activation and/or resistance to changes in form. A well known example would be the intervocalic flapping of stops in American English, which applies at the onset of unstressed syllables only. This differential behaviour often allows for the identification of one syllable that is 'strong' and subsequently classified as stressed.

3. Relative acoustic enhancement of stressed syllables. This well known diagnostic of stress has resulted in many acoustic investigations of word stress (unfortunately not always of equally sound methodology). Some confusion still persists surrounding the interpretation of F0 as an exponent of stress rather than of postlexical pitch accent, with the latter, rather than the former, often turning out to be the actual cause of enhancement (notably: Fry 1955, 1958, see also Gordon & Roettger 2017). There are further issues with the identification of stress based on acoustic enhancement that have to do both with terminology (when 'stress' is originally intended to refer to surface enhancement but is reinterpreted as inherent prominence specifications) and with methodology (when words in isolation are investigated acoustically, and phrase-level prominence is misconstrued as a correlate of word-level prominence). Nevertheless, some languages do display acoustic correlates of stress in the absence of confounds, and this will be discussed in more detail in Section 2.3.3.

4. A further diagnostic for stress is one that considers native speaker judgments. This might involve metalinguistic questions about what is considered the most prominent position in the word, or non-linguistic exercises involving tapping on the 'beat' in parallel to a stretch of speech. Interestingly, this diagnostic is sometimes

[6] In fact, some recent studies still base the identification of stress solely on perceived prominence by non-native speaker linguists, including Zuraw, Yu & Orfitelli (2014). While it is true that the coincidence of pitch accents with stressed syllables is robust in some languages, and therefore that perceived prominence on a specific syllable *might* be a reflection of stress in that position, it is by no means clear that this is a reliable diagnostic crosslinguistically.

considered problematic or its results are simply ignored on the grounds that native speakers disagree on the position of prominence or do not know what is meant by it (see e.g. the discussion of native speaker judgments in Moroccan Arabic in Chapter 4, and discussions of stress in Indonesian, cf. Goedemans & Van Zanten 2007; Maskikit-Essed & Gussenhoven 2016). In fact, it can be argued that the very existence of non-converging or difficult-to-elicit native speaker judgments is highly informative, since such findings suggest either that lexically-determined prominence asymmetries do not exist in the language, or if they do, that they represent a rather different phenomenon from what are comparatively straightforward types of stress characterised by more consistent judgments (as in Germanic).

5. Finally, I would like to suggest that there is a diagnostic which may serve to clarify whether a language has lexically specified prominence asymmetries in the first place. It is based on the results of two decades' worth of 'stress deafness' experiments, culminating in findings by Rahmani, Rietveld & Gussenhoven (2015) and Hellmuth, Muradás-Taylor & Karrinton (to appear). These stress deafness experiments have yielded the insight that a specific kind of memory task (a so-called Sequence Recall Task, or SRT) reliably distinguishes between participants who are native speakers of a language with lexical prominence (showing good performance) and participants who are native speakers of languages that lack lexical prominence (exhibiting relatively poor performance).

In the above, I showed that there are at least four diagnostics (1. to 4.) that may serve to identify stressed syllables (or moras), and one additional, general diagnostic (5.) for the very existence of lexical prominence asymmetries in a language. At this point it is not clear to what extent any of these diagnostics are necessary or sufficient criteria on their own. It is clear, however, that the more diagnostics converge, the stronger any claims about stress are, keeping in mind that not all languages exhibit all of the above diagnostics. There are languages that are considered to have lexical stress but postlexical pitch accent apparently does not dock on these syllables (including Kuot, Wolof and Chickasaw, see Section 2.5 for more detail). There are languages that exhibit little if any acoustic enhancement of stressed syllables, like Hungarian (Varga 2002; Szalontai et al. 2016, and see also Figure 2.1 in Section 2.3.3). Such languages might still be characterised by consistent native speaker intuitions about stress and the co-occurrence of pitch accents with stressed syllables. There are also languages that exhibit most or all of these five diagnostics, such as Germanic languages. In these languages, stressed syllables can receive postlexical pitch accents, but even in the absence of pitch accents, stressed syllables are acoustically enhanced in terms of duration. Stressed syllables are also those syllables that display the full set of contrastive vowels in the language, as opposed to unstressed syllables. Finally, speakers of such languages agree on what stressed syllables are, and they do not exhibit stress deafness on SRTs (Rahmani, Rietveld & Gussenhoven 2015).

The observation that stress might be identified by possibly any combination of the aforementioned diagnostics makes it difficult to convincingly argue that a given language lacks lexical stress. Firstly, negative results for any single one of the diagnostics cannot be taken as conclusive evidence that a given language lacks lexical prominence asymmetries. Secondly, the discussion of negative results brings me to the inherent problem of trying to prove a null hypothesis; If a diagnostic or test is negative, it is not logically possible to conclude that the answer to the question asked with that diagnostic or test is indeed negative. Not being to able answer with 'yes' might simply mean one has looked wrong; Negative results can always be due to flawed design. The solution for the purposes of this thesis (where I will argue in favour of the absence of lexical stress in languages), is the following: If multiple diagnostics converge, in the sense that not one is able to provide evidence *supporting* the existence of lexical stress, it should be possible to conclude the opposite: That stress does not exist. Or: "If […] a language makes it so hard to find the stress, one naturally has to ask whether stress is phonologically activated at all" (Hyman 2014: 78).

2.3.3 Acoustic correlates of stress

Here I will briefly review one of the most widely used diagnostics of stress, namely that stressed syllables are the ones that are enhanced acoustically relative to other syllables in the same word.

Among the possible acoustic correlates of stress are differences in duration, vowel quality, intensity or spectral balance or tilt, and F0. The differences between stressed and unstressed syllables are typically manifested as enhancement on the part of the stressed syllable. Stressed syllables have longer duration, more peripheral vowel quality, more dispersed spectral quality, and enhanced intensity and F0.[7] It is well known that stress in different languages may be signalled by means of different combinations of the above correlates, if they signal stress through enhancement in the first place.

Among the possible correlates, F0 and intensity appear to be the trickiest to interpret. Doubt can be raised about the general reliability of intensity and F0 as single correlates. Intensity is problematic by virtue of its high correlations with pitch prominence and duration, and by virtue of its doubtful perceptual status as a retrievable single correlate (Sluijter, van Heuven & Pacilly 1997). F0 on the other hand is well known to be a treacherous correlate of stress due to its employment in signalling postlexical prominence. Several studies that mention increased F0 as a correlate of stressed syllables in fact reported on F0 differences that were later interpreted as correlates of postlexical prominence rather than of lexical stress proper (notably Fry 1955, 1958 for English). If, in a given language, postlexical prominence in the form of a pitch accent seeks out a specific syllable, this fact in itself would be an indication that the relevant syllable is stressed (see diagnostics of stress in Section 2.3.2). It would nevertheless be incorrect in those cases to interpret enhanced F0 as a direct correlate of stress. A more important

[7] Here, the correlate of vowel quality is intended to refer to differences in what are phonologically the same vowels as a function of stress, rather than to phonological differences that distinguish *between* sets of potentially stressable and unstressable vowels. Also, it is worth noting that languages seem to exist that reserve peripheral vowel quality realisation for unstressed rather than stressed vowels. For a more elaborate version of this argument see Crosswhite (2001).

consideration in this context is that not all enhanced F0 should be interpreted as pitch accentuation or prominence marking, since enhanced F0 might result from a host of factors, many of which have little to do with lexical stress. The most well known alternative cause of pitch prominence includes phrasing-related pitch movements. This is especially relevant for languages that exhibit no acoustic correlates of stress (or other diagnostics for that matter) but do exhibit F0 movements on most lexical words, such as Korean. Normally considered to lack lexical stress, Korean has been misconstrued as having lexical stress on the basis of F0 movement alone (cf. Jun 2005a). The same will be shown to have happened for both Tashlhiyt Berber in Chapter 3 and Moroccan Arabic in Chapter 4.

Some correlates of stress are generally held to be robust, or reliable. In a review of more than 100 studies investigating correlates of stress in over 75 languages, Gordon & Roettger (2017) show that duration is reported to be a common cue to stress, followed, in order of importance, by F0, intensity and spectral measurements. Extending these findings to all known languages or interpreting them as standardly held assumptions (which to some extent they already are) about the crosslinguistic manifestation of correlates of stress is dangerous. In addition even to the aforementioned inherent issues with F0 and intensity, several other things should be kept in mind. Firstly, the meta study reviews published works, and thus possibly reflects a bias towards publishing significant results (indeed, the majority of studies found at least some effect). Secondly, some further doubt is cast on the reliability of the original results by the authors' discussion of the methodologies employed in data elicitation. For example, around *half* of the studies reviewed were not clear on the exact context in which target words were placed. Thirdly, many studies exhibited experimental flaws, ranging from very small numbers of speakers to employing elicitation contexts in which postlexical prominence as opposed to lexical prominence was investigated. In sum, while the results of the individual studies are highly valuable in their own right, these observations cast doubt not only on the reliability of some of the results on their own (as reflecting lexical stress proper rather than enhancement due to postlexical prominence) but also on the extendability of some correlates of stress as crosslinguistically common or reliable.

At this point, it can only safely be said that stress in a given language *may* but does not have to be signalled by acoustic enhancement of the relevant syllable or mora. In addition, it is by no means clear that when stressed positions are acoustically enhanced, some correlates are generally preferred over others.

In order to illustrate different ways in which stress may or may not be cued acoustically, consider Dutch and Hungarian. Dutch is a typical Germanic language in the sense that it has variable stress location and displays most of the known diagnostics of stress, including rather robust acoustic correlates: Stressed syllables are longer, have more peripheral vowel quality and are spectrally enhanced *in the absence of postlexical prominence* (Sluijter & van Heuven 1996; Rietveld, Kerkhoff & Gussenhoven 2004). Hungarian is also uncontroversially considered to have lexical stress, which is fixed in word-initial position. Native speakers agree that stress is word-initial, and word-initial syllables consistently attract pitch accents in sentence context. Its acoustic correlates in the absence of postlexical prominence however are minimal: Stressed syllables are not longer and they do not have different vowel quality. Szalontai et al. (2016) show that only intensity is a reliable marker of word-initial syllables, which supports earlier observations that this might in fact be the only acoustic correlate of stress in Hungarian (Varga 2002). Based on the previous discussion, it remains an open question whether this is a perceptually retrievable correlate of stress.

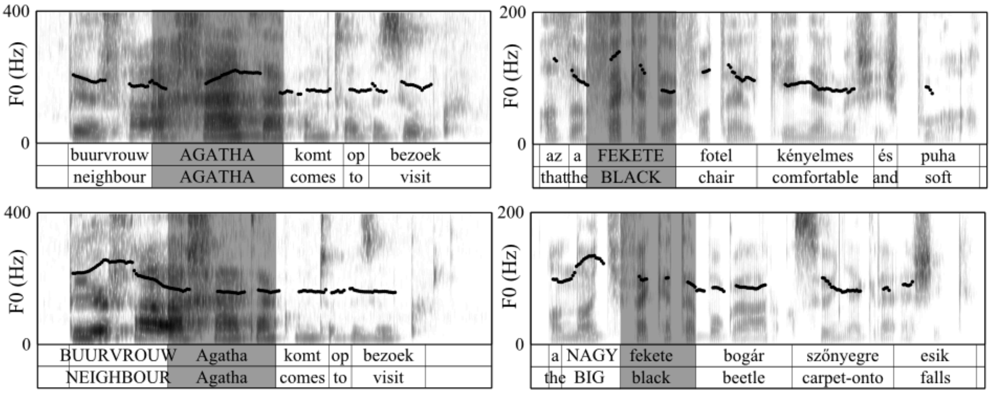

Figure 2.1: Dutch and Hungarian words (*Agatha* and *fekete*, highlighted) embedded in sentence context. Target words carry postlexical prominence in the form of a pitch accent in the top panels. In the bottom panels target words do not carry postlexical prominence. Data recordings are author's own.

Figure 2.1 illustrates the acoustic realisation of a word in each language: [aːˈxaːtaː] '(the proper name) Agatha' for Dutch and [ˈfɛkɛtɛ] 'black' for Hungarian.[8] The top panel gives a sentence in which the target word is focused and receives a postlexical pitch accent that associates with its stressed syllable (in Dutch on the second syllable ˈxaː and in Hungarian on the first syllable ˈfɛ). The bottom panel gives a context in which the target words do not carry postlexical prominence. For Hungarian it is clear that there is no durational asymmetry between the initial stressed syllable in and the other ones in the same word. These are the same phonological vowels and therefore directly comparable. For Dutch, the vowels are phonologically comparable too, but phonetically the non-stressed vowels are much reduced (the first vowel, for example, could be transcribed as [ɐ] in a narrow phonetic transcription). In addition, the stressed vowel is longer even in the absence of pitch prominence. Finally, especially in the case of Dutch with segmentally identical sentences, it can be seen that the highlighted target words in the top panels are longer overall than the target words in the bottom panels. This shows that the simple presence of a pitch accent also results in lengthening (of the entire word, although it will usually target the stressed syllable disproportionately). Pitch-accent related lengthening in part explains the ubiquitous finding in the literature that stress results in lengthening, at least when 'stress' is investigated by means words

[8]It is difficult to find Dutch native words that have full vowels in all syllables, hence the use of a proper name.

that have a pitch accent.

2.3.4 Languages lacking lexical stress

As previously discussed, the existence of multiple diagnostics for stress and the difficulty with null results makes it problematic to argue for the *absence* of lexical stress in a given language. There are nevertheless a number of (non-tonal) languages that are generally thought to lack it. These include French (Post 2000; Jun & Fougeron 2002), Korean (Jun 2005a), Mongolian (Karlsson & Svantesson 2004; Karlsson 2014), West Greenlandic (Jacobsen 2000; Arnhold 2014), and at least some varieties of Indonesian and Malay (Goedemans & Van Zanten 2007; Van Heuven & Faust 2009; Maskikit-Essed & Gussenhoven 2016).

There is an additional class of languages in which the existence of stress is somewhat unclear, with the likely possibility that prominence asymmetries in these languages are in fact assigned at a structural level above the word, as in French. For example, Tamil has been mentioned to lack stress (Nolan & Jeon 2014; Gordon & Roettger 2017) but experimental results leading to this claim are somewhat ambiguous (Keane 2003, 2006b). For Bengali, the terminology used to describe prominence is rather confusing and its intonational system appears dissimilar to languages with clear lexical stress and postlexical pitch accent, suggesting that lexical stress is not a straightforward notion (Hayes & Lahiri 1991; Khan 2014). Other languages that belong to this group are Turkish (Kabak & Vogel 2001; Levi 2005) and Farsi (Mahjani 2003; Sadeghi 2011).

Finally, it should be observed that some lexical tone languages are considered to lack stress, including Vietnamese (Athanasopoulou & Vogel 2016; Brunelle 2017) and many sub-Saharan African languages. Other tone languages including Mandarin are said to have stress (Yip 2002). In most of these cases, stress refers to surface acoustic enhancement, which as argued here does not necessarily constitute evidence for the existence of lexically specified, phonological prominence asymmetries. Very little is known about actual lexical stress in tonal languages, but it is possible that (some) tonal languages have syntagmatic, culminate prominence in addition to paradigmatic lexical tone, see for example Hyman (2014) and Gussenhoven (2004).

For (non-tonal) languages that clearly lack stress, the arguments that are brought forward to support the relevant analysis usually refer to some combination of the aforementioned set of diagnostics. In most of languages that (are thought to) lack stress, native speakers struggle with the concept of a single syllable that is supposed to stand out. On the phonetic level, these languages also exhibit a lack of consistent enhancement of specific syllables, and on the phonological level there are no known phonological processes that distinguish between strong and non-strong positions. Farsi, French and Indonesian speakers, moreover, have been shown to be stress deaf (Rahmani, Rietveld & Gussenhoven 2015).

2.3.5 Summary: Lexical prominence

Stress was defined here as the abstract property of one syllable (or mora) within a word that marks it as different from the others. In reviewing the literature I showed that there are at least four diagnostics that might serve the identification of stressed positions:

- Stressed positions attract postlexical prominence-marking pitch events
- Stressed positions are distinct in terms of segmental phonology
- Stressed positions are subject to acoustic-prosodic enhancement (in the absence of postlexical prominence)
- Stressed positions are easily identified by native speakers

Additionally, native speakers' performance on a memory task (such as an SRT) involving a postlexical prominence contrast might serve the general purpose of determining whether lexical prominence asymmetries exist in a language in the first place.

While any one single or any combination of diagnostics might confirm the location and/or very existence of lexical stress, the second diagnostic, referring to the acoustic enhancement of stressed positions, was shown to be tricky and should ideally not be used on its own: On the one hand stress might exist in the absence of acoustic enhancement, and on the other hand syllables may be acoustically enhanced for reasons that have nothing to do with lexical stress.

2.4 Postlexical prominence and intonation

2.4.1 Defining postlexical prominence

Intonation is often considered to have two main functions: The grouping of speech into units, and the marking of prominence relations within these units. It is this latter, postlexical prominence-marking function of intonation which this section will review. The relationship between postlexical prominence and its correlates is somewhat different from the relationship between lexical stress and its correlates: It was argued in Section 2.3.2 that lexical stress is, or should be, identifiable even in the absence of acoustic correlates. Postlexical prominence is less abstract, since something can only be postlexically prominent by virtue of surface prosodic enhancement. In its prototypical form, postlexical prominence is brought about by a localised pitch event ('pitch accent') which marks out an element of speech as more prominent than others: "Sentence-level prominence amounts to the categorical presence or absence of a pitch accent" (Gussenhoven 2015: 11).[9] While determining the categorical presence versus absence of a pitch accent is typically unproblematic, it is much more difficult to reach agreement on the phonological categorisation of a given pitch accent type, see e.g. Grice et al. (2017).

[9]There is some debate, and by no means consensus, whether postlexical prominence exists that does not take the form of pitch accentuation. It has been suggested that Italian, for example, has postfocal prominence cued by spectral enhancement but not F0 (e.g. Bocci & Avesani 2011).

The question as to how prosody and intonation in general, and postlexical prominence in particular, can be described categorically will be discussed in the following. The current standard model of intonational structure is couched in Autosegmental Metrical (AM) phonology (for AM see Goldsmith 1976; for AM analysis of intonation see Pierrehumbert 1980; Beckman & Pierrehumbert 1986; Pierrehumbert & Beckman 1988; Gussenhoven 2004; Ladd 2008). In Section 2.4.2 I will first briefly review interpretations of prominence in pre- and non-AM phonological approaches to intonational analysis. In Section 2.4.3 I will then discuss the most important aspects of AM-style analyses of intonation and prominence ().

At this point I should clarify that I will only be concerned with a subset of models that deal with the phonological representation of intonation, to the exclusion of models that are primary phonetic, for example those with the aim of speech synthesis, including Paint-E (e.g. Möhler 1998), the Fujisaki model (e.g. Fujisaki & Hirose 1984) and PENTA (Xu 2004a,b), the discussion of which is beyond the scope of this thesis. A more comprehensive overview can be found in, for example, Reichel 2010.

2.4.2 Non-AM phonological approaches to postlexical prominence

There are several pre-AM approaches to intonational analysis that refer to phonological, categorical interpretation. I will start with some observations about intonation that originated in a syntactic perspective on sentence prominence distribution, before moving on to approaches that are more concerned with the study of intonation for its own sake.

One of the first formalisations of the observation that (perceived) prominence asymmetries exist, and that there seems to be a default prominence location for linguistic units larger than the word, at least in English, is the Nuclear Stress Rule in Chomsky & Halle (1968). The 'nuclear' part of the name reflects the syntactic embedding of the element which typically receives the 'stress' (the highest perceived prominence), namely the finalmost word in the phrase. Much generative work in syntax since then has been concerned with formalising the location of 'sentence stress' or 'phrasal stress', thus making a distinction only between the highest level of perceived prominence versus everything else within a given multi-word domain.

This approach of trying to identify default or at least rule-based prominence patterns in a sentence stands in stark contrast to the approach taken by Dwight Bolinger, who famously called one of his papers "Accent is predictable (if you are a mind reader)" (Bolinger 1972). Bolinger's proposal was to appeal instead to information value, in the form of semantic weight, focus, givenness, and speaker intentions to explain patterns of sentence accentuation. Note also that he was not exclusively concerned with main prominence or nuclear pitch accent but with accentuation in general.

Generative approaches to phonology following the appearance of Chomsky & Halle (1968) started to focus more on the details of accent distribution in general, rather than focusing on the nuclear accent. On the one hand, Metrical Phonology (Liberman & Prince 1977) was concerned with how rhythmic structure arises as a function of the distribution of perceived prominence (called 'stress' at all levels), and how levels of prominence structure map onto each other as well as onto syntactic structure. On the other hand, the tradition of Prosodic Phonology (Selkirk 1986; Nespor & Vogel 2007) is more generally concerned with phonological structure and its interaction with other domains of linguistic structure, including syntax and morphology. This approach addresses structural prominence relations as one of several factors involved in determining the grouping of phonological units. As in Metrical Phonology, prominence is conceived of as involving the alternation of (perceived) strong and weak elements.

The conception of prominence in approaches that aim to describe intonational phonology is somewhat different. In contrast to both aforementioned phonological traditions, the Autosegmental-Metrical (AM) approach to intonation is less concerned with the interaction between various domains of linguistics. This approach is also less concerned with degrees of perceived prominence (although perception has always continued to play a role) and instead focuses on the phonetic realisation and phonological-categorical classification of different F0 events.

Before moving on to a more detailed discussion of AM analysis of intonation, two further approaches that aim(ed) to develop analyses of intonation for its own sake should be mentioned. These are the British School and IPO (Instituut voor Perceptie-Onderzoek) approaches.

The British School described phrasal intonation by appealing to the succession of phonological units, ‘tone units’ (equivalent to Intonation Phrases, or IPs, in AM analyses), that are defined based on their forming a coherent unit of linguistic meaning (Crystal 1972; O’Connor & Arnold 1973; Tench 1996). Each of these tone units has a single locus of main prominence (‘nucleus’) which is associated with the stressed syllable of the most perceptually prominent word. The distinctive pitch of the nucleus takes the form of a glide, or occasionally a level. Postlexical prominence in this tradition thus refers to the further enhancement of an already structurally prominent position through a pitch glide or obtrusion.

The IPO or ‘Dutch School’ approach to intonation, like the British School, considers ‘rises and falls’ to be the basic primitives needed for a model of intonation (Cohen & ’t Hart 1968; ’t Hart & Cohen 1973; ’t Hart & Collier 1975; ’t Hart, Collier & Cohen 1990). Its researchers aimed to provide a rule-system that yields all possible Dutch intonation contours, by specifying the combinations in which basic primitives may occur, and by restricting how these combinations (‘configurations’) can combine into phrasal intonation contours. They did not specify exactly how these intonation contours are mapped onto text, although they do note the consistent co-occurrence of specific configurations with specific structural positions such as stressed syllables. The IPO approach distinguishes between pitch and prominence (the latter conceived of as a purely perceptual phenomenon), considering only certain pitch events to have a prominence-lending function.

The idea of glides or configurations as intonational primitives (be it the contours from the British School or the contour primitives from the IPO) has since lost ground to the idea that intonation can be analysed crosslinguistically with a relatively small set of (level) tonal targets as primitives. This is the Autosegmental-Metrical approach to intonation and will be discussed in more detail in the next section.

2.4.3 AM phonological approaches to postlexical prominence

2.4.3.1 Background

Autosegmental-Metrical (AM) phonology derives its name from a combination of research traditions. 'Autosegmental' refers to the analysis of languages with lexical tone in the tradition of autosegmental phonology, in which lexical tone is specified independently of the lexical–segmental material with which it co-occurs (Goldsmith 1976). Transferred onto the analysis of intonation, this meant that intonational tonal targets could be specified as sequences, or tunes, with tune-to-text alignment being a separate analytical step following the identification of the targets themselves. The 'Metrical' part of the name refers to the fact that intonational tonal events are not just realised anywhere: Their distribution can be explained by appealing to positions specified by metrical structure, which may include lexical positions of strength, and edges of phonological constituents.

Research on the phonological structure of intonation, or intonational phonology, as analysed by means of the AM framework has its roots in Bruce (1977), with subsequent seminal works including Pierrehumbert (1980) and Beckman & Pierrehumbert (1986) (see also Ladd 2008 for an overview of work in AM tradition since the 80s). Bruce, swiftly followed by Pierrehumbert and Beckman, was among the very first to analyse the continuously varying prosodic features of speech (in Swedish, in his case) as consisting of a small set of discrete subunits that can be combined in multiple ways. Two of the ideas central to AM phonology, namely that i) intonation can vary independently of the segmental material with which it co-occurs, and ii) intonation is still sensitive to certain structural specifications, are also common assumptions underlying other models of intonation, as discussed in the previous section. The AM phonological approach to intonation, however, differs in terms of the primitives it uses for phonological classification, because it uses level tonal targets: Typically only H(igh) and L(ow) (compare with the glides used in British School analysis or the dynamic intonational categories used in the IPO descriptions).[10] These tonal targets can in turn be allocated to categories such as pitch accents and boundary tones to reflect their function as prominence-marking or edge-marking. This functional distinction is the subject of the next section.

2.4.3.2 Postlexical prominence in AM phonology

As mentioned previously in Section 2.4.1, intonation has two main functions: The grouping of speech into units and the marking of prominence relationships within these units. It is important to note that this is not a distinction that is originally due to

[10] While AM analyses of intonation are primarily concerned with the structure of the F0 contour, it is widely accepted that other prosodic properties are relevant too. Notational systems have been developed in which degrees of perceived prominence and various types of perceived juncture can all be transcribed (cf. Beckman, Hirschberg & Shattuck-Hufnagel 2005).

AM phonology. Several traditions have made the observation that aspects of speech melody can be characterised as either prominence-marking (sometimes explicitly acknowledged to be linked to co-occurrence with stressed syllables) or edge marking:

- The Prague school: 'culminative' or 'distinctive' versus 'delimitative' functions (Trubetzkoy 1969);
- American Structuralism: 'pitch phonemes' versus 'juncture phonemes' (Trager & Smith 1951);
- The IPO: 'prominence-lending' or 'accent-lending' parts of a contour versus 'non-prominence-lending' or 'terminal' parts of a contour ('t Hart, Collier & Cohen 1990, and similar observations in 't Hart & Cohen 1973).

The main point of interest here is how AM phonology deals with the distinction between intonational events that mark prominence and those that do not. In the context of this distinction two further notions need defining: Alignment and association.

Alignment "refers to the temporal implementation of fundamental frequency (F0) movements with respect to the segmental string" (Prieto 2011: 1185). This general definition of alignment is related to the notion of 'segmental anchoring', an observation about alignment patterns in a number of languages (originally termed the 'Segmental Anchoring Hypothesis' Arvaniti, Ladd & Mennen 1998; Ladd, Mennen & Schepman 2000). Specifically, segmental anchoring refers to a situation in which the points marking the start and end point of a pitch movement are both found to temporally occur at locations that can be defined with reference to some segmental landmark (Ladd 2006). Over the years the term has also been used to describe situations in which individual tonal targets (i.e. local high and low turning points) are consistently found in some specific segmental location. Whereas segmental anchoring refers to a type of consistent alignment behaviour, the term alignment alone refers simply to the phonetic-temporal location of an intonational event and does not presuppose consistency.

Association, on the other hand, refers to the phonological interpretation of F0 movements as belonging with specific structural phonological positions, which may be domain edges, or Tone Bearing Units (TBUs) such as syllables or moras. In many languages, positions of metrical strength (stressed syllables) are common TBUs.

Both notions, alignment and association, are crucial to an understanding of the role of prominence within intonational analysis, because whether some intonational primitive (in the form of a tonal target) is considered to contribute to prominence depends in part on its alignment and its interpreted association.

In AM phonology, those pitch movements that occur in the vicinity of stressed syllables are considered prominence-marking and are called *pitch accents*, while movements that are found elsewhere are considered to be edge-marking (*edge* or *boundary tones*): "More or less by definition, a tone that seeks to associate with a lexically stressed syllable is a *pitch accent*" (Ladd 2008: 145). The distinction between pitch accents and boundary tones is in first instance based on differing phonetic alignment, but it is to a large extent backed by perception, since pitch excursions near stressed syllables result in the perceived enhancement of a linguistic element that is already felt to be prominent. Pitch movement in other locations does not typically contribute to the percept of prominence. Whether edge tones should be interpreted as associating with syllables at or near the edge (in addition to associating to the edge) is still an open question and differs from one AM model to the other. Some accounts posit secondary association to explain the behaviour of edge-marking tones that phonetically align on a specific syllable close to the edge (e.g. Pierrehumbert & Beckman 1988; Grice, Ladd & Arvaniti 2000). Moreover, what seems to be a robust dichotomy between prominence-marking and edge-marking has been challenged somewhat by work on a number of well-studied languages including English and German, which has showed the existence of intonational events that have an intermediate character, so-called 'phrase accents'. Unlike pitch accents these are not felt to contribute to prominence, but nevertheless dock on stressed syllables (Grice, Ladd & Arvaniti 2000).

Even if the dichotomy between prominence-marking and edge-marking intonational events is generally accepted, the factors governing prominence distribution are somewhat more elusive. On the one hand, it is well known that some languages do not use postlexical prominence (pitch accentuation) to mark constituents which have semantic-pragmatic prominence, but phrasing instead. Some further reference will be made to these languages in Chapter 5, and in the discussion of whether postlexical prominence is used at all in certain languages, in Section 9.3.2. On the other hand, in languages that do have prominence-marking pitch accents, not all lexically stressed syllables in a sentence typically receive a pitch accent (although there are language-specific differences in this respect, see Section 2.5.1). In the following, I will briefly review two contrasting views on how postlexical prominence in the form of pitch accents is distributed across phrases or sentences in languages that clearly do have pitch accents.

The first view on sentence prominence or accent distribution is what Ladd (2008) calls the 'highlighting' view, according to which the location of pitch accents stands in a direct relationship to high information value. Words that are considered important (or informative, or prominent from the speaker's point of view) are accented, and those words that are accented are important. This view has trouble explaining the entire range of accentuation patterns even in languages like English, in which semantic-pragmatic prominence and accentuation relatively often coincide and can generally (though not always) be predicted, and makes highly inadequate predictions about information value based on accentuation patterns and vice versa for many other languages.

The second view is the 'Focus-to-Accent' view (Gussenhoven 1983; Ladd 2008). A particular version of this view seems capable of capturing several cross-linguistic accentuation patterns. This version, referred to as the 'structural' view in the earlier edition of Ladd's book, holds that accentuation partly reflects which elements of the sentence have relatively high information value. The location of accents is, however, also influenced by specific structural requirements in the grammar of the language in question. In particular, this view incorporates the idea that focus (information value) may project from a lower-level constituent up to a higher-level constituent, as a consequence of which an accent does not end up on the stressed syllable of each smaller constituent. This idea of focus projection forms one of the reasons why there is no direct mapping between semantic-pragmatic prominence and accentuation in most if not all languages.

Thus far, research has shown that in appealing to the combination of information value paired with language-specific accent assignment rules, several types of crosslinguistic patterns of pitch prominence distribution can be accounted for, and I will use the term prominence-*marking* (as opposed to prominence-*lending*) to reflect the relevant property of pitch accents.

Finally, it should be noted that the Focus-to-Accent view might not be able to account for every detail of the distribution of accentual prominence in a given language. It seems possible that in some cases, pitch accents may occur on a constituent for reasons that have nothing to do with information value (projected or not), with notions like 'ornamental accents' (Büring 2007), as well as the phrase accents previously discussed perhaps being of some explanatory value.

2.4.4 Summary: Postlexical prominence

In the preceding sections I discussed different ways in which the term postlexical prominence has been understood, and of the place of postlexical prominence within various approaches to the study of linguistic and phonological structure.

For the purposes of this thesis, postlexical prominence was defined as a phonological-categorical type of prominence, assigned to phonological units larger than the word (although they will still likely occur *on* a specific word). Postlexical prominence is either present or absent, and typically takes the form of an intonational pitch accent. Despite pitch accents being assigned at a postlexical level, based on pragmatic or phonological-structural factors, pitch accents are interpreted to seek association to a position of metrical prominence somewhere within the domain they are assigned to. This is most likely the stressed syllable of a lexical word. This makes the correspondence between postlexical prominence and lexical prominence in languages that have lexical stress a direct one (even if not a bi-directional one; A pitch accented syllable will be stressed, but not all stressed syllables obtain a pitch accent).

Postlexical prominence was also contrasted with (postlexical) edge-marking: Whether an intonational event is interpreted as serving the purpose of prominence-marking or of edge-marking depends on both its phonetic alignment (in the vicinity of a stressed syllable or at the edges of a speech unit) and its perceptual status (contributing to prominence or speech unit delimitation).

2.5 Interaction between lexical and postlexical prominence

2.5.1 General issues in the correspondence between lexical stress and postlexical pitch accent

One of the main aims of this thesis is to form a theoretical contribution to the discussion about how prominence specifications at lexical and postlexical levels of linguistic structure might or might not be interdependent.

For Germanic and other well-studied European languages, the correspondence between lexical prominence in the form of stress and postlexical prominence in the form of pitch accent is well established.[11] Moreover, many definitions of lexical stress can be found that refer directly to this correspondence, including:

- "[o]ne possible kind of phonemic stress is potential for pitch accent" (Bolinger 1958: 149)
- "Word-level stress is the capacity of a syllable within a word to receive sentence stress when the word is realized as part of the sentence" (Lehiste 1970: 237)
- "Lexical stress provides the designated terminal elements for the assignment of intonational tones ('pitch accents')" (Hyman 2014: 58)

While none of the above sources would claim that every stressed syllable will always be pitch accented, the correspondence can account for the 'conceptual merger' (Gussenhoven 2015) as a result of which the F0 correlates of pitch accent are attributed to lexical stress (see also Section 2.3.3).

In general, however, it is well known that the correspondence between lexical stress and postlexical pitch accent is not one-to-one: Languages differ in the percentage of lexically stressed syllables that actually receive a pitch accent, from around 30% in German or English to virtually all lexical stresses (i.e. words) in Egyptian Arabic (Hellmuth 2006). Even varieties within a single language can differ considerably, with 17% of stressed syllables accented in Southern European Portuguese, to 74% in Northern European Portuguese, to almost a 100% in Brazilian Portuguese (Frota & Vigário 2000; Vigário & Frota 2003).

Additionally, there are specific contexts in which the correspondence between lexical stress and postlexical pitch accent is not entirely clear. In the following, I present a number of difficulties relating to this correspondence. The first three points have to do with interpretative difficulties, while the last two points form (possible) exceptions to the overall rule that postlexical prominence maps onto lexical prominence.

A main problem is that some prominence-marking intonational events are analysed as being associated with a stressed syllable, while the phonetic alignment of the relevant turning points in the contour is not synchronised in time with that specific stressed syllable. This observation has caused considerable debate in the case of bitonal pitch accents, in which one of the tonal targets is usually considered to be the more essential

[11] Lindström & Remijsen (2005) even call it a "linguistic universal".

one, indicated by a following star '*' and interpreted in terms of greater strength. If neither of the tonal targets can be defined easily with reference to the phonological unit it is thought to belong with, such as the edges of a stressed syllable, it is not clear which of the tonal targets should be starred (detailed discussions and specific experiments addressing this question can be found in Arvaniti, Ladd & Mennen 2000; Beckman, Hirschberg & Shattuck-Hufnagel 2005; Dilley, Ladd & Schepman 2005; Prieto, D'Imperio & Fivela 2005; D'Imperio 2006; Prieto 2011).

Secondly, despite the existence of segmental anchoring (in some cases), pitch accents differ in their phonetic realisation depending on context, making the uniform analysis of prominence marking pitch events difficult to sustain even within a single language. For instance, the same pitch accent category will be aligned differently in different words, depending on factors such as the segmental structure of the syllable with which it associates, the relative position of that stressed syllable within the word, and the length of the phrase or utterance as a whole. Even when the same pitch accent is associated with the same syllable in a word, say L* + H with the first syllable in the English word *wonderful*, it will align somewhat differently depending on the phrasal position in which the word occurs, with earlier alignment of the turning points when the word is in phrase-final position (as in *That's wonderful!* as opposed to *That's a wonderful drawing!*). Another factor influencing realisational details is articulation rate, resulting in earlier alignment in fast speech. There is therefore no one-to-one mapping between the phonetic alignment parameters of an intonational event and its phonological interpretation as pitch accent type A or B. This holds very similarly for edge-marking intonational events.

A third problem is the discussion surrounding phrase accents, a notion invoked by Grice, Ladd & Arvaniti (2000) (and cf. Ladd 2008: Ch. 4) to describe the variable realisation of the right edge marking of yes–no questions in a number of languages (including Hungarian and Greek). Phrase accents designate the intonational events that occur in a phrasal position after the nuclear accent and before the right-edge boundary of the relevant domain. They are not typical pitch accents because in many cases they do not surface on stressed syllables, but rather occur wherever there is space to be realised, often adjacent to an edge tone and thus forming a complex boundary tone sequence. They do not function like typical edge tones either since, when the length of the metrical-segmental structure allows for it, they do seek out a prominent position such as a primary or secondary stressed syllable. Together, these observations suggests that intonational events exist that are neither solely prominence-lending nor solely boundary marking, and instead may be both at the same time.

The fourth point, which might be one of the exceptions to the general rule of postlexical-lexical prominence correspondence, is so-called 'stress shift'. Among Hayes's (1995) diagnostics to metrical stress is the 'Rhythm Rule', which applies in lexical compounds with adjacent stress positions. At least in English, individual words can exhibit a prominence shift when they occur in a clash context of this kind (compare 'BAMboo CHAIR' with 'bamBOO' with final prominence in isolation). For the purposes of the present definition of stress (stress being an invariant property of a lexical word), this phenomenon should be called 'prominence shift' rather than 'stress shift', since the inherent relative prominence of the final syllable of *bamboo* cannot change by definition. The shift can instead be interpreted as only concerning a particular surface realisation of the word. The phenomenon indeed seems to involve a shift of the postlexical prominence, which implies that the shift involves the perception of a pitch accent occurring on a non-primary (but often secondary) stressed syllable (cf. Shattuck-Hufnagel 1994; Gussenhoven 2004).

The fifth issue concerns another possible exception, and relates to the suggestion that there are languages which, despite having lexical stress, in fact have pitch accents that associate with structural positions other than these stressed syllables. More specifically, this has been argued to be the case in Kuot (Lindström & Remijsen 2005) and several native languages of North America (as reviewed in Gordon 2014).[12] In the first place, these statements cause terminological confusion since the standard definition of a pitch accent presupposes that it is an intonational event that seeks out a stressed syllable.

Setting the definition aside, in none of these cases do there seem to be clear criteria or robust acoustic evidence in favour of claims about prominence, be it lexical or postlexical. While experimental in nature, the study on Kuot is based on only two speakers and stress was identified by a non-native speaker linguist. Similar issues apply to the claims about the Northern Iroquoian language Onondaga (Gordon 2014). Here, the argument appeals to a different position for 'pitch accent' depending on the position of the word in the phrase: If the word is final, the accent would target the penultimate syllable, whereas if the word is phrase-medial, the accent would target the final syllable, suggesting that pitch accents do not go to predetermined stressed syllables. Note however that the term pitch accent in this context is somewhat underdefined, with the only indication that it involves 'raised F0' (Gordon 2014: 89). Such evidence hardly makes a convincing case for an interpretation along the lines of a prominence-marking pitch accent, as it is well known that not all F0 protrusions serve structural prominence marking. Moreover, the issues with interpreting stress location based on perceived pitch prominence are well known (see Section 2.3) and it should be noted that the original descriptions of (pitch) accent in the relevant North American languages (e.g. Chafe 1970; Foster 1982) date from before the widespread use of experimental methods with which claims about perceived prominence can be cross-checked. Finally, the exact same prominence pattern, involving varying final or penultimate lexical prominence depending on position of the word in the phrase, is found in at least two other cases in the literature: Halim (1974) as cited in Maskikit-Essed & Gussenhoven (2016) on Indonesian, and Boudlal (2001) describing lexical stress assignment in Moroccan Arabic. For the former, there is now a consensus that (varieties of) Indonesian lack lexical stress, suggesting that it was not lexical stress that the original claim referred to. With respect to the latter, as I will discuss in Chapter 4, this particular interpretation of stress is

[12] Both these sources cite Rialland & Robert (2001) as another study which makes similar claims, but the main argument of that study is that the language has stress but no postlexical prominence, which is a different claim from the one that stress and postlexical prominence both exist, but do not necessarily coincide.

mistakenly based on pitch excursions that represent edge-marking, not prominence.

In conclusion, while postlexical prominence in the form of a pitch accent often corresponds in position with that of lexical prominence, i.e. a stressed position, the two levels of prominence stand in a more complicated relationship to each other than simply 'pitch accent-goes-to-stress'. It is also clear that even in languages in which the existence and position of stressed syllables is uncontroversial, the translation of surface alignment patterns of individual turning points, on the one hand, into phonological representations of intonational categories on the other, is as of yet not a matter of complete understanding or agreement (Ladd 2006; Prieto 2011). Nevertheless, for languages that have clear lexical stress and postlexical prominence, the general correspondence seems a rather reliable one overall.

The next section will approach this discussion from another angle by reviewing insights from postlexical prominence distribution in languages that do *not* have lexical stress.

2.5.2 Lexical and postlexical prominence in languages lacking stress

Section 2.3.4 listed a small number of languages which are generally accepted to lack stress: French, Korean, West Greenlandic, Mongolian, as well as varieties of Indonesian and Malay. Work on intonation in these languages has highlighted that there might be varying strategies according to which *postlexical* prominence is assigned in the absence of lexical anchors. As stressed syllables are considered pivotal in determining the location of postlexical pitch accents in most languages with stress, the mechanisms governing prominence placement in its absence are of some interest. A binary distinction between stressless languages has already been made in terms of whether they mark prominence at the postlexical level at all (Jun 2005b, 2014a).[13]

On the one hand there is French, which is considered to have a fixed position for postlexical prominence, namely final at the phrasal level of the AP or PP (Post 2000; Jun & Fougeron 2002). These phrases are analysed as consisting of a string of tones (LHiLH* in Jun & Fougeron 2002) which map in a prespecified way onto the syllables in that phrase. The initial rise (/LHi../) and the final rise (/..LH*/) behave somewhat differently, as reflected in the star notation for the latter but not the former. Star notation typically indicates the relative strength and/or alignment of the tone that is marked with it, as it does in the present case of French: It reflects both the tone's interpreted association to the final syllable in the phrase, which is inherently prominent, and its phonetic alignment with this syllable. While Jun & Fougeron (2002) interpret the tone as an edge tone with an additional association to a prominent position, Welby (2006) calls the pitch event in question a H* pitch accent that 'does not convey prominence' (Welby 2006: 368). This use of the term pitch accent seems rather misleading, at least if one defines a pitch accent as a prominence-marking tonal event that associates with a lexically stressed syllable (although it is in keeping with the idea that the intonational

[13] West Greenlandic, as one of the aforementioned stressless languages, is left out of this discussion because little reference can be found to its prominence structure, e.g. Arnhold (2014).

event occurring in this position has some link to structural prominence). A potentially better characterisation would be that of a pitch accent with phrasal-metrical association, or 'phrasal pitch accent'/'postlexical pitch accent', which would contrast with a 'standard' pitch accent which seeks lexical-metrical association. A similar analysis might be appropriate for Mongolian. The existence of prominence asymmetries at the postlexical level in Mongolian is less clear than in French, but there is some evidence that the first syllable of an AP is privileged in terms of being the only position that exhibits a paradigmatic vowel contrast (Karlsson & Svantesson 2004). While the tonal movement marking the same position (a rising gesture at the left edge of each AP) is not by definition considered to contribute to prominence, focus may be achieved by enhancing the initial rising gesture, suggesting that the AP-initial position is indeed a position of metrical prominence (cf. Karlsson 2014).

On the other hand there are languages like Korean and Ambonese Malay. For these languages there is no evidence to support the existence even of postlexical prominence. Apparently, in these languages, pitch is used neither in the sense of associating to phrasal culminative positions (as in French), nor in the sense that edge-aligned pitch movement can be interpreted as having a prominence-marking function.[14]. These languages certainly have intonation, but it can be analysed exclusively in terms of edge-marking tones without reference to functional prominence-marking.

As an aside, while it seems possible for languages without stress to also lack postlexical prominence, it is unclear at this point whether languages with stress can lack postlexical prominence. To my knowledge, this has only been suggested to be the case for Wolof (Rialland & Robert 2001).

The case for Korean as one such language is made by Sun-Ah Jun. Jun (2005a) reviews a number of studies concerned with the perception of prominence in Korean, but all these studies involved non-native speakers who seemed to equate high F0 with prominence. While surface differences in F0 clearly exist, these have nothing to do with prominence. Firstly, there does not seem to be a role for structural prominence asymmetries between syllables within an AP, as in French (Jun 2014a). Secondly, information structure is signalled by means of phrasing, not accentuation of (some part of) a specific constituent (Jun & Oh 1996; Jun 2014a, although a different, impressionistic, view can be found in Choe 1995).

In addition to Korean, Ambonese Malay (Maskikit-Essed & Gussenhoven 2016) has also been argued to be a language of the type that lacks prosodic postlexical prominence altogether.[15] Based on elicited sentences with phrase-final words occurring in different information structural contexts, the authors conclude that the language simply does not have postlexical prominence marking: Contrastively focused words in this position were not realised differently from non-focused words.

[14] A third, but novel, logical option, namely that prominence marking of a constituent may occur somewhat more freely on the constituent, i.e. without the constraints of metrical anchoring, will be argued to be the case for certain intonational patterns in TB and MA, see Chapters 6 and 7)

[15] A difference between the languages concerns the interpretation of edge tones. In Korean, edge tones seek association to a specific TBU, whereas Ambonese Malay edge tones do not. This difference has little bearing on the interpretation of postlexical prominence.

If we consider that lexical prominence may be present or absent in a language, and postlexical prominence may be present or absent as well, the following typology can be suggested:

	Lexical stress	**No lexical stress**
Postlexical prominence	many, including: most European languages, most varieties of Arabic	French, Mongolian
No postlexical prominence	Wolof?	Korean, Ambonese Malay, West Greenlandic

Table 2.1: Proposed typology of languages as a function of lexical and postlexical prominence structure.

The idea that languages can be classified based on their prominence structural properties at lexical and postlexical levels is taken up again in the discussion in Chapter 9, with added insights from findings with respect to Tashlhiyt Berber and Moroccan Arabic.

2.6 Research questions

The overarching research question asked in this thesis is whether Tashlhiyt Berber and Moroccan Arabic can convincingly be shown to lack lexical stress.

Specifically, I am asking the following questions, for both languages:

- **Part II: Lexical prominence production**

  Can prominence asymmetries be observed in the form of relative acoustic enhancement of presumed stressed syllables? That is, are there acoustic correlates of (presumed) stress? This will be tested by means of production experiments that allow for the isolation of correlates of lexical prominence to the exclusion of postlexical prominence (Chapter 3 and 4).

- **Part III: Postlexical prominence production**

  What is the behaviour of postlexical prominence-marking events? What phonological unit(s) do these associate with? Specifically: What are the properties of the intonational movements associated with question words? This question will be answered based on production experiments also, involving question word questions produced in a scripted dialogue context (Chapter 6 and 7).

- **Part IV: Prominence perception**

  To what extent do native speakers exhibit sensitivity to postlexical prominence asymmetries that are parasitic on lexical prominence specifications? This question refers to participants' performance on a lexical Sequence Recall Task. Tasks of this kind are considered to reveal 'stress deafness', and have been shown to provide a reliable insight into the presence of lexical prominence asymmetries in the lexical phonology of the native language (Chapter 8).

In addition to an answer to the main research question and the subquestions, a further contribution is made with respect to language contact. The present experiments allow for the near-direct comparison of results from Tashlhiyt Berber and Moroccan Arabic in the case of the production experiments, and for a direct comparison in the case of the perception experiment. These comparisons have the potential to shed light on the degree of convergence between the two languages in terms of prosodic-phonological aspects of linguistic structure.

2.7 Summary

This chapter has given the theoretical background for the five experiments that will be reported on in the next chapters. The following key definitions were given:

- **Lexical prominence** is the phonological property of one syllable (or mora) within a word that marks it as prominent in relation to the others. **Stress** is a type of lexical prominence asymmetry that does not (exclusively) involve lexical pitch accent.

- **Postlexical prominence** refers to phonological prominence which is assigned above the structural level of the word. In languages that have lexical stress it takes the form of a **pitch accent**.

This chapter's discussion of lexical and postlexical prominence has moreover highlighted the complex interdependence between lexical and postlexical prominence specifications. On the one hand, many languages whose intonation has been analysed within the AM phonological framework are characterised by a highly reliable correspondence in the location of postlexical and lexical prominence. In contrast, in languages in which the question of lexical–postlexical prominence correspondence does not arise, due to the non-existence of lexical stress, the mechanisms accounting for the distribution of postlexical prominence are as of yet poorly understood.

Part II

Lexical prominence production

3 Acoustic correlates of word-level stress in Tashlhiyt Berber

3.1 Introduction

3.1.1 Prior work on the lexical phonology of Tashlhiyt Berber

3.1.1.1 Stress

Prior to 2015, the question whether lexical stress is present in Tashlhiyt Berber had been addressed in passing in a few grammatical descriptions of the language, and it had also been addressed in one experiment.[1]

At the turn of the 20th century, Stumme (1899) observed that lexical prominence ('Wortaccent') is highly variable in sentence context, and with this observation appears to have been the first to propose that Tashlhiyt lacks lexical stress in the sense of lexical prominence specified for each individual word. Later, similar observations were made by Applegate (1958), following a failure to provide a uniform characterisation of word-level and phrase-level prominence patterns, and by Dell & Elmedlaoui (2002), who considered prominence asymmetries to exist only at a phrasal level rather than at the lexical level.

In contrast to the above, the one experimental study that addresses questions about the existence of stress in TB claims that it is in fact present (Gordon & Nafi 2012). The design of the materials in this study however leaves some room for (re)interpretation of the results. Half of the speech material analysed consisted of individual words produced in isolation, which means these words would have been subject to phrase-level prosody. Firstly, this phrase-level prosody could have manifested itself in terms of the presence of phrase-level prominence (such as the nuclear pitch accent which by default occurs on the final content word in Germanic languages). Secondly, the presence of phrase-level prosody on the target words will most likely have resulted in the presence of phrase-level edge marking in the form of continuation intonation. This is a crosslinguistically common form of intonation used for items produced in lists, and judging from the pitch tracks provided by Gordon & Nafi (2012). In sum, these words were subject to the entire range of intonational marking also found on larger phrases, in the present case however condensed onto a single word. Unsurprisingly then, the authors found considerable enhancement of word-final syllables: These had longer duration, greater intensity and higher pitch than their non-final counterparts. While some of these are common, or

[1]The data reported on in this chapter have previously been published, in slightly different format, in Roettger, Bruggeman & Grice (2015) and in Roettger (2017: Ch. 4).

at least possible, correlates of word-level stress (see Section 2.3.3), it is impossible to know whether the attested enhancement was due to phrasal prosodic enhancement or due whether they reflected acoustic correlates of stress proper. The other half of the materials in Gordon & Nafi (2012) consisted of the same words followed by an adverb, meaning that the words were in prefinal phrasal position. While this design feature removes the potential issue of list intonation on target words, a remaining confound concerns the likelihood that target words were inadvertently contrasted in this position, as they occurred, one after the other, in the same carrier phrase. As a result, target words might have received some sort of prosodic prominence signalling contrastive focus, beyond enhancement due to lexical stress proper.

The experiment reported on in this chapter reexamines the question of correlates of lexical stress in TB experimentally, with materials specifically designed to investigate acoustic properties of individual syllables in the absence of postlexical enhancement.

3.1.1.2 Aspects of segmental and syllable structure

Segmental structure, syllable structure and syllabification in Tashlhiyt are well researched due to the interest surrounding its unusual tolerance of long consonantal sequences (as an aside, it shares this feature with Moroccan Arabic) (e.g. Dell & Elmedlaoui 2002; Ridouane 2008, 2014). Accordingly, the discussion of syllable structure has primarily focused on what some consider (phonologically) vowelless syllables. These may in fact be produced with a vocalic element of short duration, the status of which as either a transitional vocoid or a phonological vowel has sparked considerable debate (Dell & Elmedlaoui 1985; Coleman 1999, 2001; Dell & Elmedlaoui 2002; Ridouane 2008). Setting this vowel aside, Tashlhiyt has three unequivocal phonological vowels: /i, a, u/ with no length contrast. Consonantal length on the other hand is contrastive, as is pharyngealisation, which spreads throughout the relevant syllable or word domain (Ridouane 2014).

For the purpose of the present experiment, words were chosen that consisted of relatively straightforward syllable types (CV and CVC) that moreover have phonological vowels as nuclei, to avoid difficulties in determining syllabification and to enable reliable F0 measurements. Consonants in target words were simpletons (rather than geminates) and lacked pharyngealisation to avoid unintended acoustic changes to other segments due to spreading.

3.1.2 Research question

The present experiment was designed to see if Gordon & Nafi's (2012) results could be replicated, to either confirm or refute their claim that Tashlhiyt has final stress. The question asked for the purpose of this experiment is thus: 'Do syllables in Tashlhiyt exhibit distinctly different patterns of acoustic prominence depending on position in the word?'. If the answer to this question is affirmative, this would suggest that there are inherent prominence asymmetries between syllables depending on position. This in turn could be interpreted in terms of lexical stress targeting a specific position in the word. If, on the other hand, results provide a negative answer to this question, this finding would be in line with the assumption that there is no lexical stress in Tashlhiyt (or at least no stress-by-position).

3.1.3 Data

The present chapter reports on an experiment that was designed to test acoustic correlates of stress in Tashlhiyt Berber, and specifically to compare the results with findings by Gordon & Nafi (2012). Recordings were made by the author in November 2014 at the Université Ibn Zohr in Agadir.

3.1.4 Overview of chapter

The next section will present the methodology of the experiment. Section 3.3 presents the results and reports on acoustic correlates of stress as a function of position (initial/final in the word). All results will be considered from both a syllable- and a speaker-based point of view. Any conclusions that are then drawn are thus based on the presence or absence of a systematic pattern, that holds across individual words and across individual speakers. The results section is followed by two separate sections for the discussion (3.4) and summary and conclusion (3.5).

3.2 Methodology

3.2.1 Participants

The speakers in this experiment were 10 native speakers of Tashlhiyt, all students in the Département des études amazighes at the Université Ibn Zohr in Agadir. All participants were multilingual and spoke fluent Moroccan Arabic as a second language (see also Section 1.2). Most additionally spoke some French. One speaker's recording was excluded from further analysis due to poor recording quality. In the results section, individual speakers are referred to by the numbers 1–9 followed by the letter "M" for male speakers, and "F" for female speakers.[2]

3.2.2 Procedure

Participants were given oral instructions about the task by a native speaker. They were then seated in front of a laptop screen in slide presentation mode from which they read out scripted mock dialogues that included the target sentence stimuli (see details below). Dialogues with target items were interspersed with dialogues with filler items. The experiment was self-paced was part of a larger session which lasted 40–45 minutes. Participants did the present experiment reading out dialogues twice, with the full set of dialogues occurring in two random orders, and performed other tasks in the

[2] These codes match the ones used in Roettger (2017: Ch. 4).

interim. The experiment was self-paced and the two repetitions took 15–20 minutes to complete. Recordings were made in a univeristy office using a PreSonus Audiobox solid-state recorder at a sampling rate of 44.1 kHz, and a head-mounted AKG C420 III microphone.

3.2.3 Speech materials

The target words in the experiment were designed to yield CV and CVC syllables that occur as an initial syllable in one target word and as a final syllable in another, such as /di/ in *dima* 'always' and *sidi* 'sir'. In some cases, the target syllable occurred twice in a single word consisting of a reduplicated syllable, such as /ba/ in *baba*. All target syllables and the words in which they occurred are given in Table 3.1.

Syllable	**Initial position**	*Gloss*	**Final position**	*Gloss*
da	/dari/	'with me'	/juda/	'enough'
di	/dima/	'always'	/sidi/	'sir'
ni	/nizi/	'of the fly'	/ʁuni/	'sing'
ta	/tama/	'next to'	/ʃita/	'brush'
ti	/tili/	'sheep, ewe'	/hati/	'here (is)'
	Initial & Final			
ba	/baba/	'father'		
bi	/bibi/	'turkey'		
kaw	/kawkaw/	'peanuts'		
kif	/kifkif/	'it's the same'		
tam	/tamtam/	'drum'		
jan	/janjan/	'one by one'		

Table 3.1: Tashlhiyt target syllable pairs in target words

Target words were placed in a carrier sentence which was in turn embedded within a scripted mock dialogue consisting of three sentences. Sentences were presented on a laptop screen in speech bubbles coming from smiley faces to represent the two imaginary discourse participants.[3]

The context of the dialogue served to ensure that the target word in its carrier sentence (1) was i) given and not marked by intonational prominence, because it was explicitly mentioned in the previous sentence, ii) not marked by any other intonational events, as it occurred in a yes-no question in which the edge-marking intonational event occurs on the phrase-final two syllables (Grice, Ridouane & Roettger 2015), and iii) non-final in the IP so as to avoid phrase-final lengthening effects.

Sentences were presented in Latin script, which is the most common way to write Tashlhiyt. One mock dialogue with a phonological transcription is given in 1 with the target word /dari/ 'with me':

[3]I am grateful to Carlos Gussenhoven for sharing this methodology.

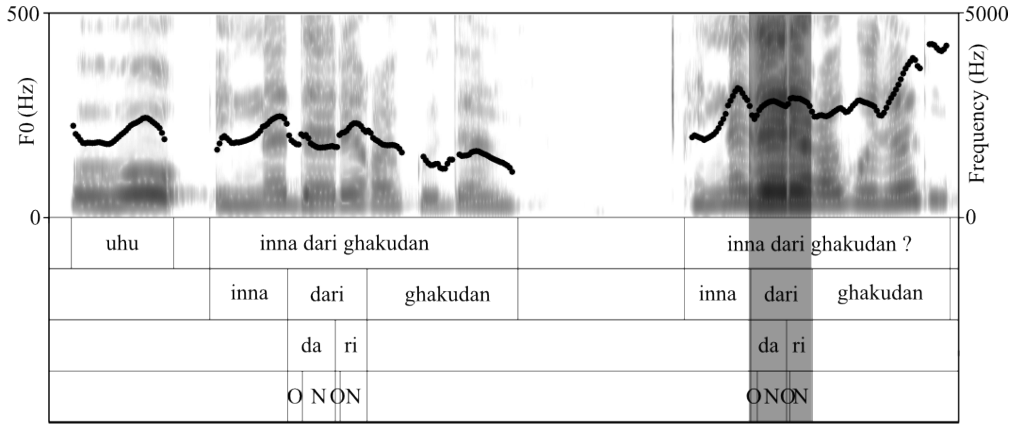

Figure 3.1: Example spectrogram and F0 contour (smoothed) for context sentence 2 and target sentence, spoken by speaker 7f), with target word *dari* highlighted.

Context sentence 1	**Context sentence 2**
is in:a [baba] ʁakudan	uhu, in:a [dari] ʁakudan
Q say.AOR [baba] then	no say.AOR [dari] then
'Did he say [baba] then ?'	'no, he said [dari] then'

Target sentence

in:a [dari] ʁakudan
say.AOR [dari] then
'He said [dari] then ?'

An example spectrogram and waveform are given in Figure 3.1, showing context sentence 2 and the target sentence. As can be seen, the target word *dari* in the target sentence occurs prior to the main pitch event in the sentence (the rise-fall at the right edge of the phrase), and does not appear to be the locus of some phrasal or edge-marking pitch event itself.This means that the target words occurred in the intended context to examine correlates of lexical stress proper.

3.2.4 Analysis

The acoustic parameters investigated were duration, intensity and F0, matching the set of correlates reported in Gordon & Nafi (2012). Annotation was performed manually. Target sentences were segmented into words, and target words were further annotated for the syllable boundary and onset and nucleus of both syllables. This was unproblematic since all words had either the structure /CV.CV/ or /CVC.CVC/.

The theoretical number of target items was N = 396 (9 speakers * 2 syllable positions * 11 target syllables * 2 repetitions). Of these, N = 288 were targetlike in the sense of the correct target word being produced in a fluent sentence. Out of these, N = 279 were produced without any pausing, and the analysis will use this pause-less set of tokens.

F0 measurements are based on a handcorrected version of the fundamental frequency contour provided by the standard pitch-tracking algorithm in Praat (Boersma & Weenink 2015). Manual correction was limited to the correction of pitch-tracking errors, such as octave jumps and the tracking of pitch in cases of phonetically voiceless segments. There were two F0 measurements: Peak in and mean throughout the vowel.

For duration, measurements of target syllables and vowels were taken. Vowel duration was determined as the period of time following the initial consonant with strong periodic energy across the second and third formant. Intervocalic /r/s were mostly realised as either trill or tap, therefore the onset of the /r/ was determined as the start of the (first) closure. Intensity was measured in terms of mean energy throughout the vowel.

To control for variation between speakers, and to facilitate visual comparison, all measurements were z-scored. These data are shown in the speaker-specific graphs. The statistical models use the raw measurements, and account for speaker differences with random intercepts and slopes.

Statistical analysis was performed with linear mixed-effects regression models with the package *lme4* (Bates et al. 2015) in R (R Core Team 2016). Separate models were run for each of the acoustic parameters under investigation, with the relevant parameter as a fixed effect. To allow for potentially varying interactions with the fixed effect, random intercepts for speakers and random slopes for items (syllables) and speakers were included. Statistical significance was calculated by means of likelihood ratio tests LRT)s comparing hypothesised models with the corresponding null model that lacked the relevant fixed effect or interaction term. The R syntax of the models that are used is given in the footnotes.

3.3 Results

3.3.1 Duration

Figure 3.2 shows the raw durational measurements for the vowels in each syllable. There are differences between syllables in whether the final or the penultimate position results in longer duration: For e.g. *ba* and *kaw* the final vowel is longer, whereas for e.g. *kif* and *tam* the penultimate vowel is somewhat longer. This indicates that there is no systematic positional enhancement of syllables: Either the final or the penultimate vowels would have to be consistently longer for this to be true.

Statistically, none of the durational differences are significant. Vowel duration was not different as a function of whether it occurred in the final or penultimate syllable (LRT: $\chi^2(1) = 2.45$, $p = 0.12$). The same holds for syllable duration (LRT: $\chi^2(1) = 0.04$, $p = 0.85$).[4]

[4] syllable/vowel duration ~ syllable position + (0 + syllable position|speaker) + (0 + syllable position|syllable) + (1|syllable) + (1|speaker)

Individual speakers also do not seem to make a position-dependent difference, as visualised in Figure 3.3. If speakers did make a consistent difference, all connecting lines should be going in the same direction.

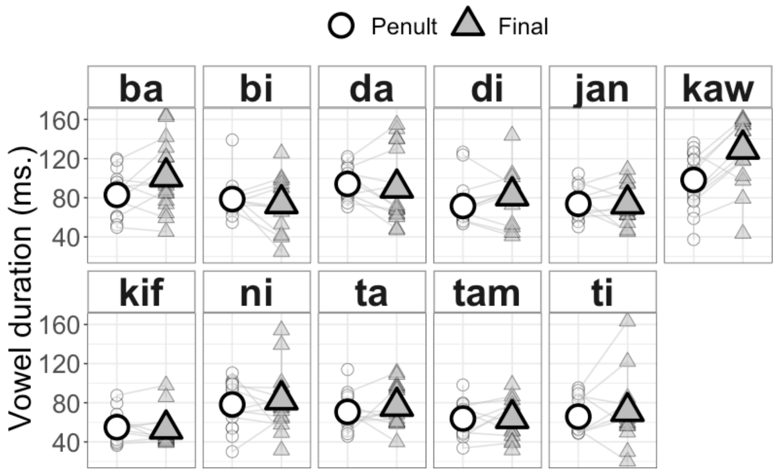

Figure 3.2: Duration (ms.) of vowels as a function of syllable position (penultimate/final in the word). Lines link productions by the same speaker, large dots represent mean.

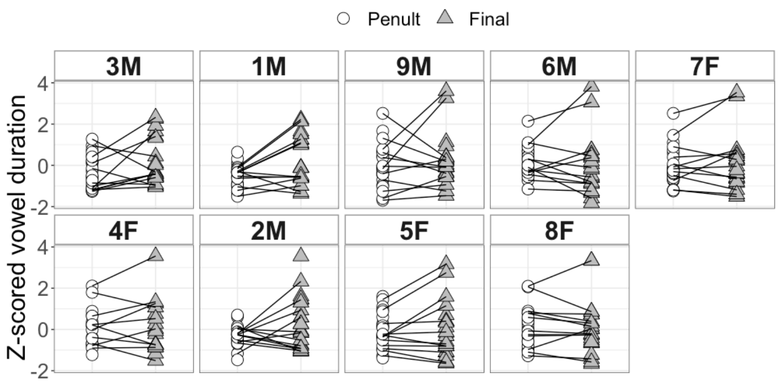

Figure 3.3: Z-scored duration of vowels as a function of position (penultimate/final in the word), per speaker. Lines link productions of the same syllable.

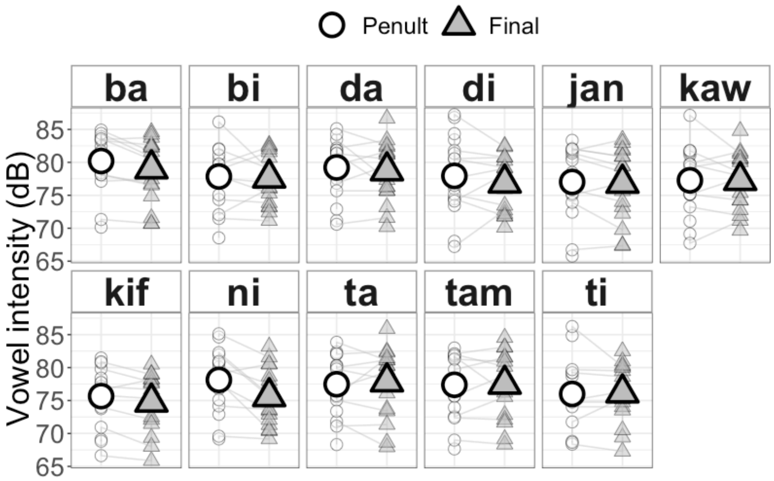

Figure 3.4: Intensity (dB) throughout the vowel for each syllable as a function of position (penultimate/final in the word). Lines link productions by the same speaker, large dots represent mean.

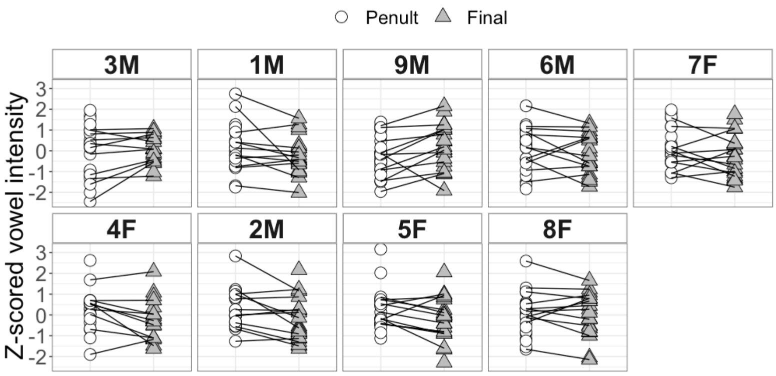

Figure 3.5: Z-scored vowel intensity as a function of position (penultimate/final in the word), per speaker. Lines link productions of the same syllable.

3.3.2 Intensity

Figure 3.4 shows the raw intensity measurements for the individual syllables. There are few apparent differences between syllables in whether the vowel is relatively enhanced in the final or the penultimate position. There is however a small statistical difference in the sense that final syllables have somewhat lower vowel intensity overall than penultimate syllables (LRT: $\chi^2(1) = 4.4$, $p = 0.04$).[5] The estimated difference is 0.56 dB, which arguably constitutes a marginal difference that lies well below the threshold for

[5] vowel intensity ~ syllable position + (0 + syllable position|speaker) + (0 + syllable position|syllable) + (1|syllable) + (1|speaker)

just noticeable differences in intensity of around 1 dB (cf. Lehiste 1970; Beckman 1986 and see also discussion of the relevance of intensity as a cue to stress in Section 2.3.3).

Individual speakers also do not make a consistent distinction between syllables in penultimate and final position in terms of intensity. As shown in Figure 3.5, there is no systematic trend here: It is not the case that all final syllables exhibit lower intensity.

3.3.3 F0

As previously mentioned, F0 mean and peak measurements throughout the vowel were taken. Figure 3.6 shows the mean F0, in semitones relative to 100 Hertz, for penultimate and final syllable nuclei. There are no large differences between mean F0 as a function of syllable position, but the overall data nevertheless suggests a consistent small difference between final and penultimate syllables. This is corroborated by a statistical difference for both the mean F0 measurement (LRT: $\chi^2(1) = 10.0$, $p < 0.01$), and the peak F0 measurement (LRT: $\chi^2(1) = 6.49$, $p < 0.05$).[6] The predicted mean difference was 0.83 ST, and the peak difference 0.55 ST, with word-final nuclei having the lower F0 values. I will return to this finding below.

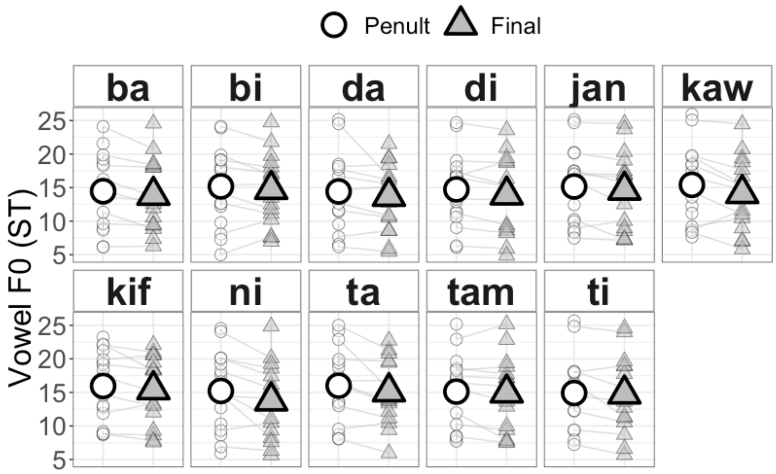

Figure 3.6: F0 (ST) mean throughout vowel for each syllable as a function of position (penultimate/final in the word). Lines link productions by the same speaker, large dots represent mean.

Turning to speaker-specific behaviour, Figure 3.7 shows behaviour that is mostly in line with these findings. Most individual speakers produce lower F0 in syllables that are word-final, with the possible exception of speakers 3M, 4F and 9M who do not make a very large or systematic difference.

The aforementioned predicted differences of 0.83 and 0.55 ST are nevertheless small in absolute size, and may in this respect be irrelevant for the purposes of serving as a

[6]nucleus mean/peak F0 ~ syllable position + (0 + syllable position|speaker) + (0 + syllable position|syllable) + (1|syllable) + (1|speaker)

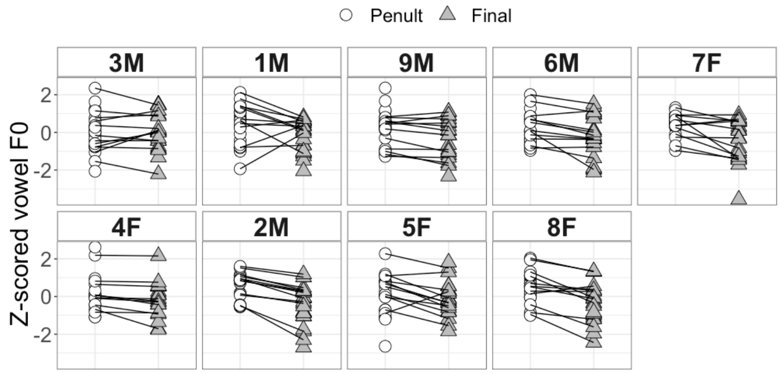

Figure 3.7: Z-scored F0 measurements of syllable nuclei as a function of position (penultimate/final in the word), per speaker. Lines link productions of the same syllable.

reliable phonetic exponent of lexical stress. Firstly, in terms of perception, the more liberal estimates for just noticeable pitch movements in dynamic speech stimuli start at 1 ST (for example, exceptionally good listeners in 't Hart 1981). Most of the literature however has suggested that larger pitch movements are required in order for listeners to reliably identify differences ('t Hart 1976; d'Alessandro & Mertens 1995). Secondly, stress is typically realised as acoustic enhancement. There is no definitive reason why low pitch (rather than high pitch) should not be considered enhancement, but given the known status of high pitch as the prototypical form of enhancement, it would at the very least be unusual (or perhaps it should be taken to point to the interpretation that the initial syllables are the enhanced ones). A third observation, which provides the most likely explanation of the observed differences, concerns the possibility that lower F0 in final syllables results from declination throughout the phrase: The lower F0 values are found on the word-final syllable which coincides with a later position in the phrase. Figure 3.8 shows all F0 contours for the target sentence that contains the word *yanyan* 'one by one' in which both syllables can be compared directly. These contours are characterised by a gradually declining slope throughout the utterance, at least until the starting point of the final rise(-fall) that is typical of yes-no question modality in TB (Grice, Ridouane & Roettger 2015).

These observations together suggest that the small F0 difference does not reflect enhancement of the penultimate syllable as a function of fixed stress in that position. Rather, it forms a predictable phonetic correlate of postlexical (phrasal) structure, in this case reflecting general facts of speech production (declination throughout a phrase is often thought to result from a gradual reduction of pulmonic effort throughout an utterance, cf. Ladefoged 1972; Ladd 2008).

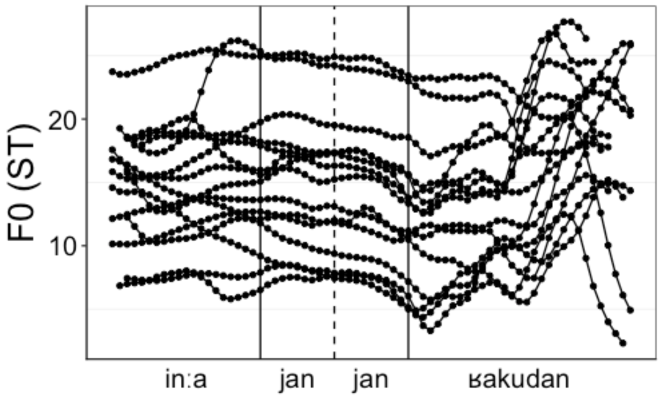

Figure 3.8: Time-normalised contours for all speakers' target sentence *in:a janjan ʁakudan* 'he said [one by one] then'.

3.4 Discussion

Among the acoustic correlates discussed in this chapter (duration, intensity and F0), there were no reliable differences between presumed stressed (final) and unstressed (initial) syllables. The differences that did exist, i.e. slightly enhanced F0 and intensity for *initial* syllables, are not compatible with the results and interpretation of final stress given by Gordon & Nafi (2012). The fact that the enhancement involved both F0 and intensity, but not duration, as well as the fact that the differences were very slight, form indications that the relevant differences are unlikely to form exponents of lexical stress. Firstly, it is well known that both intensity and F0 tend to gradually decline over the course of an utterance. The fact that higher values of each are seen in word-initial syllables, which occur in an earlier phrasal position, could therefore be interpreted as following from this tendency. Secondly, global intensity measures like the present one (mean across vowel) are highly correlated with F0, meaning that the effect of intensity might simply follow from the higher F0 values in the same position. This would then require an explanation for the F0 difference only, as for example above.

A few more observations can be made to discount 'stress' as causing a reliable positional distinction, zooming in on syllables and speakers separately. Not one of the 11 syllables showed consistent differentiation in terms of a combined effect of the three possible correlates (the closest to a possible exception is the syllable *kaw*, which is durationally differentiated — but the other parameters still do not conspire). For all syllables, lines in Figures 3.2, 3.4 and 3.6 typically cross each other or stay level: For no individual syllable was there a consistent, directional effect. Similarly, not one of the 9 speakers exhibit consistent differentiation between 'stressed' and 'unstressed' (final and initial) syllables, for any of the syllable types. Rephrased: Not one speaker exhibited behaviour that caused all lines (in the speaker-specific figures) to go in the same direction. In sum, the degree of overlap in the present data distribution prevents the conclusion that stress, here in the form of a specific lexical position, exhibits reliable acoustic enhancement.[7]

Finally, the possibility should be raised that perhaps the present study failed to find correlates of stress because stress in Tashlhiyt is not assigned by position, but rather is assigned by weight, or is variable in position (rule-based and/or lexically specified). While these scenarios of different stress assignment cannot be dismissed, simply because they were never explicitly tested, there is currently no reason to believe they are very likely. On the one hand, specific stress predictions other than the one of fixed final stress tested in Gordon & Nafi (2012) have not been suggested in the literature (and Gordon & Nafi (2012) also did not set out with an explicit hypothesis about the position of word stress). On the other hand, the possibility of the absence of lexical stress in TB, as suggested in some literature (see 3.1.1.1), is compatible with evidence from intonation suggesting that there is no fixed lexical position to which prominence-marking intonational events associate (Grice, Ridouane & Roettger 2015). It seems more fruitful therefore to try to find further support for the absence of stress from other angles (see also Section 2.3.2). This is the purpose of some of the following chapters: Chapter 4 (which tests for stress in MA), Chapter 6 (which looks at the behaviour of another type of intonational event in TB) and Chapter 8 (investigating perceptual clues to the existence of lexical stress).

3.5 Summary and conclusion

This chapter investigated correlates of lexical stress in Tashlhiyt Berber by contrasting syllables in word-final position with syllables in word-initial position, following up on the earlier claim by Gordon & Nafi (2012)) that syllables in word-final position are stressed. The results from the present experiment could not provide evidence in favour of final stress on the basis of acoustic enhancement: Syllables in final position were not consistently enhanced compared to their word-initial counterparts in terms of duration, intensity or F0. The present data did in fact exhibit slight enhancement of vowels in *initial* position, in terms of F0 and intensity only. These differences were argued not to be correlates of stress, as they were marginal in terms of size and could straightforwardly be explained as reflecting phrase-level prosody. Similarly, the differences Gordon & Nafi (2012) found, involving the enhancement of syllables in word-final position, could be argued to follow from phrase-level intonation rather than lexical stress proper.

The present study therefore provides no evidence to support a stress-by-position interpretation of the lexical phonological structure of TB. These results are compatible with earlier suggestions that lexical stress is absent in this language, and that prominence asymmetries may only be specified postlexically.

[7] Concluding the opposite, namely that stress is absent *because* there is no consistent enhancement, would be unwarranted. The impossibility to prove the null hypothesis is also addressed in Section 2.3.2.

4 Acoustic correlates of word-level stress in Moroccan Arabic

4.1 Introduction

4.1.1 Prior work on the lexical phonology of Moroccan Arabic

4.1.1.1 Stress

The question whether Moroccan Arabic has lexical stress is subject to a long-standing debate and is not currently resolved.[1] Maas (2013), for example, reviews more than 10 sources published between 1894 and 2008 that all differ to some extent in their views on the existence of word stress in MA. Complicating matters more, the different authors he cites use varying terminology, including 'Wortakzent', 'Akzent' and 'Accent' (by German authors; Stumme & Socin 1894; Brockelmann 1908; Fischer 1917), 'accent' and 'accent de mot' (in French works; Cantineau 1960; Benhallam 1989), 'accento tonico' (in Italian; Durand 1994) and finally, 'stress' (in English; Aguadé 2008). Some of these terms may be in fact be interpreted as referring to postlexical pitch prominence (rather than lexical stress as defined here, see Chapter 2). Others do seem to refer to inherent word-level prominence, despite what the choice of terminology suggests. In sum, it is not entirely clear what is meant by these terms other than that they refer to some sort of word prosody, which could be construed as either lexical or postlexical prominence. In reviewing the evidence in detail, however, Maas (2013) argues that the various positions can be allocated to two main groups: One group posits that MA has word stress, the other that MA lacks word stress. It is this latter viewpoint which is assumed by the majority of the authors he reviews.

The first group, the advocates of word stress, propose a range of stress rules and generalisations (for a more comprehenseive overview see Boudlal 2001). For example, according to Benkirane (1998: 348f.), stress falls on the final syllable if it is heavy (i.e. a closed syllable such as CVC) and on the penult otherwise. A similar view is found in Nejmi (1993) as cited in Boudlal (2001). Others believe in a fixed position for stress, such as simply 'final prominence' Watson (2011: 7), or a number of authors cited by Maas (2013) who predominantly posit penultimate stress. Yet another possibility is rather more variable stress that may target an initial CV syllable in a trisyllabic word (Benhallam 1990, as cited in Boudlal 2001).

Finally, a slightly more complicated picture is sketched by Boudlal (2001: 99) who posits that "the location of stress depends on whether or not the items considered occur

[1] This chapter has been published, with some adaptations, as Bruggeman et al. (in press).

in isolation or in context". Accordingly, stress would be final when words are produced in context, but can be captured by Benkirane's generalisation for words produced in isolation. I will return to some specific issues relating to this idea in the discussion in Section 4.4. Assuming that the position of word stress is a fixed property, Boudlal's (2001) observation about varying stress positions for individual words forms a strong indication that he was talking about *postlexical prominence* (since these *can* be variable under the present definition). This interpretation is supported by some of the pitch contours shown in his work. In any case, if word stress is considered an invariable lexically specified property, the observation that lexical prominence is truly variable would imply the very lack of word stress.

The incongruent analyses of word stress are matched by equally uncongruous judgments on the position of word stress by native speakers. A number of studies have investigated stress in terms of where the perceptual prominence of a word lies. These typically involve a few dozen participants underlining the stressed syllable in written words presented in a list, and/or listening to isolated words (see Boudlal 2001: 101f. for an overview). Most of these, including Boudlal's (2001) own underlining test, find that speakers disagree with each other on the location of stress in the same word. This disagreement might in part be due to the fact that written Arabic will automatically invoke Modern Standard Arabic (even if a task explicitly asks for stress judgments on Moroccan Arabic), which has stress assignment that differs from (the many proposals for) Moroccan Arabic. Nevertheless, disagreement among native speakers of the language adds to the elusive nature of word stress in MA, suggested in the first place by disagreement among (native) scholars.

Returning to the stress generalisations that have been proposed so far, Benkirane's is the most easily testable and and it has also been adopted in recent work, including Yeou, Embarki & Al-Maqtari (2007), Burdin et al. (2015) and Hellmuth et al. (2015). Moreover, in most other varieties of Arabic (in which the existence of stress is uncontroversial) stress assignment is subject to weight and position, with stress typically targeting the final superheavy syllable (e.g. CVCC), or a penultimate heavy syllable (e.g. CV:) in its absence (Watson 2011). If one were to assume that MA has lexical stress, and takes into account prior claims about stress position in this variety as well as facts about stress in other varieties in Arabic, the weight of the final syllable would be expected play a role in determining its location.

In this chapter I will test correlates of stress as envisaged by Benkirane (and several others with him), according to whom stress in MA targets either the penultimate or the final heavy syllable.

4.1.1.2 Syllable structure

It is well known that languages differ as to which syllables are considered 'heavy' and 'light'. Heavy syllables are those that are closed or have a long vowel, or the ones that have a long vowel while light syllables would have a short vowel (Hayes 1995). Most Arabic varieties make the first kind of distinction with the presence of a coda or a long nucleus resulting in heavier weight. Given that Moroccan Arabic does not have a vowel length distinction (see next subsection), it is the number of consonantal slots in the coda that determines the weight of the syllable as light (none), heavy (one), or, under some analyses, superheavy (two).

A lot of work has been done on syllable structure in MA, and in various theoretical traditions, yielding varying claims about its phonological representation (Boudlal 2001; Dell & Elmedlaoui 2002; and Benkirane 1982 as cited in Benkirane 1998). What is clear from all sources is that MA allows for more complex consonant clusters than most other varieties of Arabic, but the representation of these clusters in terms of branching onsets or codas is disputed. For example, Benkirane (1998) provides an inventory of syllable types that includes CV, CCV, CCVC, and CCəCC, while Dell & Elmedlaoui (2002) argue that syllable onsets cannot be branching, and that codas can only be branching if they consist of geminates. In order to account for what appear to be syllable-initial clusters, Dell & Elmedlaoui (2002) instead propose a complicated general syllabification algorithm that posits onsetless syllables and empty nuclei. Crucially, a distinction is made in all works between heavy and light syllables, and sometimes superheavy syllables. The degree of consensus is however limited to CV being considered light and CVC heavy, with the exception of CəC (at least according to Dell & Elmedlaoui 2002).

4.1.1.3 Vowel inventory

Various claims have been made about the vowel system in Moroccan Arabic. By some, it is thought to have a five-vowel system consisting of /iː ə aː ʊ uː/ (Hamdi 1991 as cited in Al-Tamimi 2009). On the other hand, many researchers posit only three or four vowels, namely /i a u/ plus a central vowel. The latter is usually considered non-phonological (e.g. Dell & Elmedlaoui 2002: 230). In addition to the number of vowels, the representation of length is also a matter lacking consensus, as can be judged from the juxtaposition of the phonological categories /iː/ and /i/, /aː/ and /a/, and /uː/ and /u/ by the aforementioned sets of authors. This might be caused by the existence of a surface (phonetic) contrast in length, with CVC syllables having longer vowels than CV syllables (Benkirane 1982; Dell & Elmedlaoui 2002; Yeou 2005). Nevertheless, it is widely acknowledged that MA lacks a phonological vowel length distinction for vowels with the same place of articulation, which sets MA apart from most other varieties of spoken Arabic. This is backed by the absence of minimal pairs of the type /sin/ 'tooth' ~ /siːn/ 'the letter S' (example from Iraqi Arabic, Al-Ani 1970: 75). In any case, it is of little consequence for present purposes whether vowels are represented as phonologically long or short. The experiment reported here involves matched pairs of lexical words in which the target syllable always has the same structure, namely /CV./ in word-initial position. Under no account would such a syllable result in potentially varying vowel length depending on context. From here on, for transcription purposes, I will assume that the phonological vowel inventory of MA can be represented as simply /i a u/, with an additional centroid vowel which may be either phonological /ə/ or phonetic [ə].

4.1.2 Research question

The primary aim of this chapter is to present a first systematic investigation of the cues that mark word stress, as conceived of by Benkirane, in Moroccan Arabic. The main question therefore is 'What are the correlates of presumed lexical stress in Moroccan Arabic?'. In answering this question some further contributions are made with respect to details of the acoustic realisation of syllables and vowels in MA in general.

4.1.3 Data

The present experiment forms part of the IVAr (Intonational Variation in Arabic) corpus of speech data (Hellmuth & Almbark 2017).[2] It includes comparable data from seven varieties of Arabic. The corpus recordings of Moroccan Arabic, including the present experiment which was specifically designed to test correlates of stress, were made in the spring of 2014 by Sam Hellmuth and Nabila Louriz at the Université Hassan II in Casablanca.

4.1.4 Overview of chapter

The next section, 4.2 will present the methodology for the experiment. The following section, 4.3, presents the results and reports on the acoustic characteristics of syllables as a function of presumed stress status (stress/unstressed). All results will be considered from both a syllable- and a speaker-based point of view. Any conclusions that are drawn are thus based on the presence or absence of a systematic pattern across individual words and across individual speakers. Separate sections then follow for the discussion (4.4), and summary and conclusion (4.5).

4.2 Methodology

4.2.1 Participants

Two groups of speakers were recorded for the present experiment. The first group consisted of 12 native speakers of Moroccan Arabic who grew up with Moroccan Arabic only at home (the 'monolingual' group).[3] These speakers were aged 21-34. Ten of them were born in Casablanca and had lived there all their life at the time of the recording, one speaker moved to Casablanca aged two, and one speaker was born in very nearby Kenitra. The second group consisted of 12 speakers of Moroccan Arabic who were also native speakers of Tashlhiyt Berber through one or both parents (the 'bilingual'

[2] I am very grateful to Sam Hellmuth and Rana Almbark for their generosity in letting me use some of their Moroccan data recordings.

[3] I use quotation marks here because these speakers were not in fact monolingual. All were highly fluent in multiple other languages, including Modern Standard Arabic and French, see also Chapter 1. For ease of reference, however, I will use the terms monolingual and bilingual and refer this way to their first, home language status.

group). Their ages ranged between 20 and 32. Nine speakers in this group were born in Casablanca, the other three moved to the city at the respective ages of six, twelve and fourteen. All speakers in both groups are fluent in Modern Standard Arabic and French, and had received a number of years English teaching in school.

4.2.2 Procedure

The present experiment was part of a larger recording session for the IVAr corpus which consisted of a number of tasks, including reading out of a mock dialogue that included the qword interrogatives discussed in Chapter 7. For the present experiment, participants were recorded individually in a quiet university room with a Shure SM-10 headset microphone. Participants were first given oral instructions by a native speaker. They were then given a print-out of the experimental stimuli, which consisted of mini-monologues. They read these mini-monologues out loud (with fillers at the top and bottom of each page), with the third sentence in each dialogue forming a target utterance (see below for details). The experiment was self-paced. There were no practice items in order to minimise the duration of the session as a whole.

4.2.3 Speech materials

The experiment was based on the design first used in Bouchhioua (2008) and subsequently in Almbark, Bouchhioua & Hellmuth (2014). The same experimental paradigm was used in the parallel investigation of stress in seven other varieties of Arabic as part of the IVAr corpus project (Hellmuth & Almbark 2017). The experiment contrasts what are identical initial syllables in minimal stress pairs, such as *mu* in ˈ*muka* 'owl' and *mu*ˈ*kat* 'owls' (stress as predicted by Benkirane). Word pairs were designed so as to have stress on the initial (= penultimate) syllable in one member of the pair and stress on the final syllable in the other member.

The full experiment tested the initial syllables of 12 target word pairs (24 words), chosen based on their comparability in the full range of dialects used in the IVAr database. Among these target word pairs, eight exhibited the MA stress contrast as per Benkirane's (1998) generalisation and are shown in Table 4.1. A rough phonetic transcription is given as well because there were considerable auditory differences between vowels that should theoretically belong to the same phonological category given in Section 4.1.1.3.

Target words were placed in a carrier sentence which was in turn embedded within a short scripted monologue consisting of a total of three sentences, given below with target word *muka* 'owl' as an example in 1):[4]

[4]It is not usual for Moroccan Arabic to be written in Arabic script, which is reserved for Modern Standard Arabic. The stimuli in this experiment however are written so as to mirror Moroccan Arabic lexical items and pronunciation. For example, /ʒuʒ mərːat/ 'twice' as used in MA translates to /kamaːn marːə/ in MSA so there was no confusion which variety was being tapped into. Moreover, speakers were explicitly instructed to speak Moroccan Arabic throughout the experimental session.

Syllable	PU stress	*Gloss*	Final stress	*Gloss*
ba	/ˈbaʃər/ [baːʃər]	‘he preached’	/baˈʃərt/ [baːʃərt]	‘I preached’
ma	/ˈmarːə/ [mɑrːə]	‘one time’	/maˈrːərt/ [mɑrːərt]	‘I passed’
mu(k)	/ˈmuka/ [muka]	‘owl’	/muˈkat/ [mukat]	‘owls’
mu(rː)	/ˈmurːə/ [murːə]	‘bitter’	/muˈrːin/ [murːiːn]	‘passers-by’
sa	/ˈsadə/ [saːdə]	‘plain’	/saˈdat/ [saːdaːt]	‘gentlemen’
si(tː)	/ˈsitːə/ [sːtə]	‘six’	/siˈtːat/ [sːtːaːt]	‘sixes’
si	/ˈsirə/ [siːrə]	‘path’	/siˈnat/ [siːnaːt]	plural of the letter *sin*
su	/ˈsurə/ [suːrə]	‘verse’	/suˈdan/ [suːdaːn]	‘Sudan’

Table 4.1: Target minimal stress pairs Moroccan Arabic

Context sentence 1

موكة كلمة سهلة
[muka] kəlma sahla
[owl] word nice
‘[owl] is a nice word’

Context sentence 2

كتب موكة جوج مرّات
ktəb [muka] ʒuʒ mərːat
write.IMP [owl] two times
‘write [owl] twice’

Target sentence

عاود موكة جوج مرّات
ʕawd [muka] ʒuʒ mərːat
say.IMP [owl] two times
‘SAY [owl] twice’

The contextual embedding served to minimise the possibility of postlexical prominence associating with target words in several ways: Target words were i) pragmatically given (mentioned in previous sentences), and ii) postfocal (occurring immediately following contrastively focused *ʕawd*). Target words were designed to be non-final in the IP to avoid phrase-final lengthening effects.

An example spectrogram and waveform are given for one mini-context in Figure 4.1. The left, middle and right phrases show context sentence 1, context sentence 2 and the target sentence respectively, with the target word *muka*.

The target word *muka* in the target sentence occurs after the main pitch event in the sentence (which occurs on *ʕawd*), and moreover does not seem to attract any postlexical prominence itself: It is not phrased separately (judging from pausing or pitch reset) nor does it receive an apparent pitch target. Pitch simply declines throughout the target word towards what seems to be a low target at the right sentence edge. This exactly matches the intended context for examining correlates of stress proper.

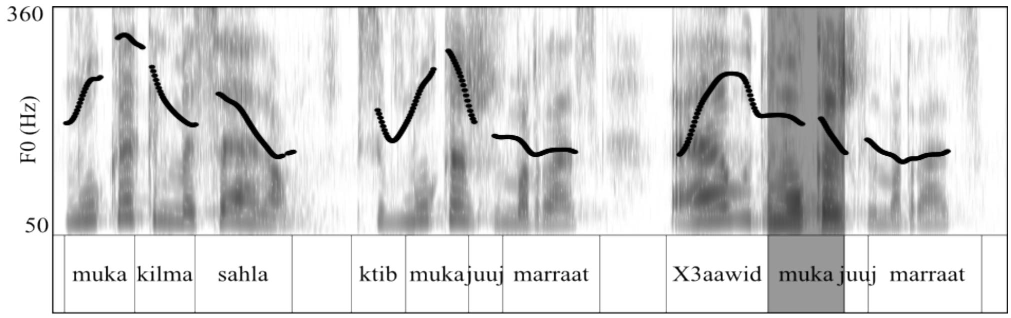

Figure 4.1: Example spectrogram and F0 contour (smoothed) for scripted monologue, with target word *ˈmuka* highlighted in the target phrase. IVAr file moca-slb3-f5.

4.2.4 Analysis

The acoustic parameters investigated for correlates of stress were duration, intensity, vowel quality and F0. Annotation proceeded as follows: Automatic segmentation of utterances into words and segments was performed by means of the Prosodylab Aligner algorithm (Gorman, Howell & Wagner 2011). The segmentation of target words was then manually checked and corrected where needed, and coded for preceding and following pauses. Pauses were defined as periods of silence in the signal and were based on the auditory impression of a pause supported by visual inspection of speech discontinuity in the spectrogram (the auditory impression of a break caused by e.g. pitch reset was thus not a sufficient criterion to annotate a pause).

The theoretical number of target items was N = 384 (2 speaker groups * 12 speakers * 2 stress conditions * 8 target syllables). Of these, N = 360 were targetlike in the sense that the correct target word was produced (MA bilinguals N = 172, MA monolinguals N = 188). Out of these, N = 251 were produced without any pauses, with the most typical location being right after the target word.

F0 measurements are based on a handcorrected version of the fundamental frequency contour provided by the standard pitch-tracking algorithm in Praat (Boersma & Weenink 2015). Manual correction was limited to the correction of pitch-tracking errors, such as octave jumps and the tracking of pitch in cases of phonetically voiceless segments.

For duration, measurements of target words and target vowels were taken. Vowel duration was determined as the period of time following the initial consonant with strong periodic energy across the second and third formant. For the segmentation of intervocalic /r/, the onset of the /r/ was determined as the start of the (first) closure (most /r/s were realised as either trill or tap, although there were some approximants as well). To control for durational differences between speakers, duration was also normalised by z-scoring target vowel duration per speaker.

Vowel quality was measured by F1 and F2 values taken at the midpoint of the vowel. Measurements were extracted by means of the *Burg* method in Praat (Boersma & Weenink 2015), using a 25 ms. Gaussian window, and a 10 ms. step. All values were verified manually, corrected where needed, or excluded where reliable formant values could not be extracted. Results are reported on both the raw F1 and F2 values as well as on Lobanov-normalised values. The latter were calculated with the NORM vowel normalisation suite (Thomas & Kendall 2007) and the R package *vowels* (Kendall & Thomas 2014).

Two measurements for intensity and F0 each were taken i) mean throughout the target vowel, ii) maximum (peak) in the target vowel. As for duration, variation in intensity and F0 between speakers was controlled for by z-scoring per speaker. To allow for a more holistic analysis of F0 movements, phrasal F0 contours were also extracted. This was done by means of taking F0 at 20 extraction points spaced equally per word, with the exception of the target word for which 10 measurements were taken in each syllable in order to be able to compare syllables directly.

Statistical analysis was performed with linear mixed-effects regression models with the package *lme4* (Bates et al. 2015) in R (R Core Team 2016). Models with the same structure were run for each of the acoustic parameters under investigation with presumed stress status as the fixed effect of main interest and an interacting effect of group (bilingual/monolingual). In the case of duration, pausing (present/absent) is taken into account as an additional fixed effect. In some cases, mainly for vowel quality, models are run on subsets of the data after the main model. To allow for potentially varying interactions with the fixed effect, random intercepts for items (target syllables or words) and speakers were included (slopes led to overspecification). Statistical significance was calculated by means of LRTs comparing main models with corresponding null models that lacked the relevant fixed effect or interaction term. The R syntax of the models that are used is given in the footnotes.

4.3 Results

4.3.1 Duration

The total number of targetlike target words was N = 360, as mentioned above. Of this number, 319 tokens actually had an initial vowel. The 41 tokens that did not were all instances of the words *sitta* and *sittat*, which were typically produced as [sːtːə] and [sːtːaːt]. This left only two tokens of stressed /si(tː)/, which both did have a vowel, and no matching instances of unstressed /si(tː)/ that had a vowel. This syllable was therefore also excluded from the statistical analysis (tokens with a vowel are however included for reference in the figures). The remaining set of N = 317 tokens was then submitted to further analysis.

During annotation, differences appeared to exist between stressed and unstressed tokens of the syllables *si* and *su*, in the sense that the stressed versions of these syllables had long, dynamic vowels, whereas the unstressed ones had shorter, steady-state only vowels. An explanation for this observation might be found in the segmental make-up of the word pairs involved. In the target words in which these syllables are stressed (*sira* and *sura*), there is a high vowel followed by [r] or [rˤ]. This rhotic considerably affected the preceding vowel formant structure and resulted in a longer vocalic portion than in the unstressed counterparts *sinat* and *sudan*, which had steady-state only initial vowels (note that the other target syllables followed by /r/, i.e. *mu* and *ma* in *murra* [mərːa] and *marra* [mərˤːa], respectively, are much shorter, but the target vowel in these cases is central rather than high). This is a known effect: Pre-/r/ lengthening of high vowels preceding rhotics is also observed in other languages, including Dutch (Rietveld, Kerkhoff & Gussenhoven 2004). In an attempt to neutralise these segmental effects, vowel duration was re-extracted using only the steady-state portion of the vowel. The analysis below is based on this shorter vowel duration.[5]

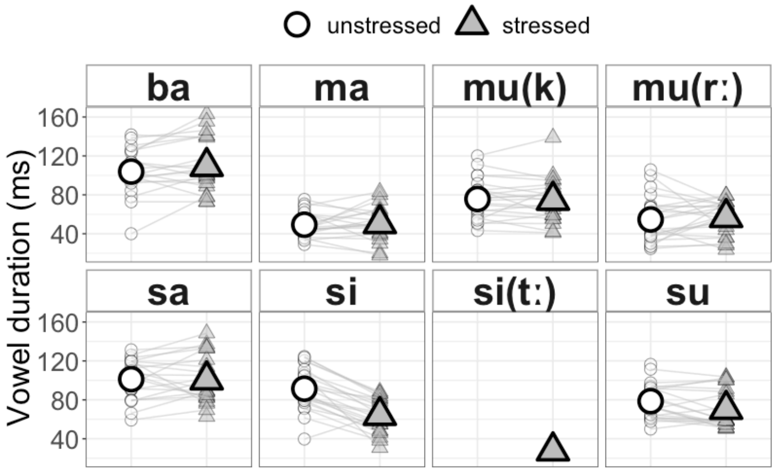

Figure 4.2: Absolute duration (ms) of vowels, per syllable, as a function of status (stressed/unstressed). Lines link productions by the same speaker, large dots represent mean per syllable.

Figure 4.2 shows the distribution of absolute vowel duration, for each syllable separately and pooled across the groups. Firstly, there was no main effect of pause (present or not) on any of the duration measures (for absolute vowel duration, LRT: $\chi^2(1) = 1.73$, $p = 0.18$).[6] The interaction between stress and group was significant only for z-scored syllable duration (LRT: $\chi^2(1) = 4.46$, $p = 0.03$).[7] Since none of the other three duration measures (absolute syllable duration, and absolute/z-scored vowel duration) were significant this effect will not be further considered. There was, however, a main effect of stress on vowel duration (z-scored and absolute), although the predicted difference involves 'stressed' vowels being 3 ms. shorter than 'unstressed' ones (LRT: $\chi^2(1) = 4.70$,

[5] Analysis was performed on the longer vowel duration measurements too, with the result of a significant length difference between stressed and unstressed syllables. This effect disappeared entirely when tokens of *si* and *su* were excluded.

[6] duration (syllable/vowel, absolute/z-scored) ~ pause + stress * group + (1|speaker) + (1|syllable)

[7] Same models as above, with model comparison of stress + group versus stress * group

p = 0.03).[8] This change is not in the expected direction if stress is considered to equal enhancement, but more importantly a change of 3 ms. on an average vowel duration of around 77 ms. does not reflect a meaningful change even if statistically significant.

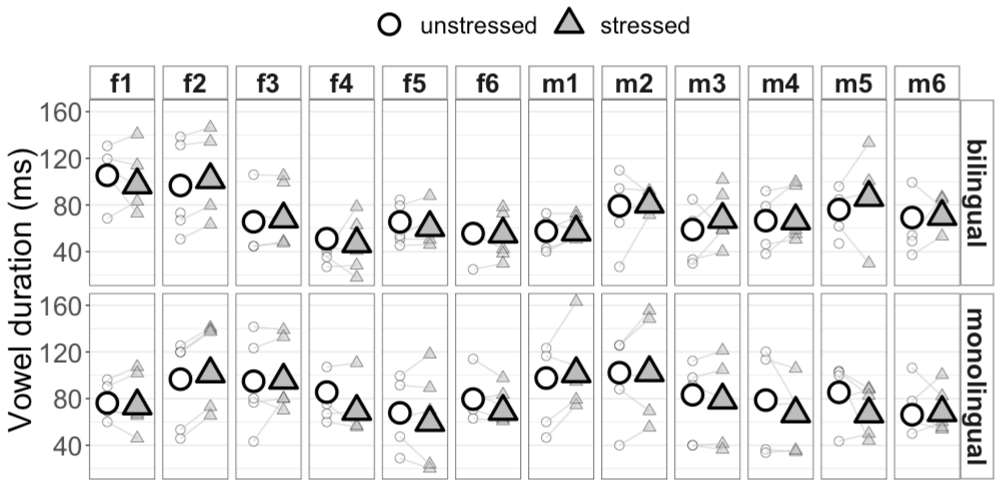

Figure 4.3: Absolute duration (ms) of vowels, per speaker, as a function of status (stressed/unstressed), tokens of *si* and *su* excluded. Lines link productions of the same vowel, large dots represent mean per speaker.

In short, both the statistics and the above observations suggest that there is no evidence to support an interpretation in terms of stress-induced vowel or syllable lengthening in MA across the board. To confirm that there are also no individual speakers who produce consistent durational enhancement of stressed vowels, Figure 4.3 shows speaker-specific behaviour (tokens of *si* and *su* are removed). Unstressed and stressed tokens of the same syllable are connected by lines (e.g. unstressed *mu* in *mukat* and stressed *mu* in *muka*. The varying direction of these lines for almost all speakers, and the large overlap between the presumed stressed and unstressed categories within each speaker, indicates that speakers did not systematically differentiate between stressed and unstressed vowels in terms of duration.

In sum, most syllables do not exhibit durational enhancement under presumed stress, and speaker-specific results confirm that in addition to an overall distinction, individual speakers also do not differentiate consistently between stressed and unstressed vowels in terms of duration.

4.3.2 Vowel quality

Vowel quality was measured by F1 and F2 on the same subset of target words that had an initial vowel as above, but includes the vowels in *sitta* for reference in the graphs (N = 319). As with duration, they are excluded from the statistical models. Of these,

[8] duration (syllable/vowel, absolute/z-scored) ~ pause + stress + group + (1|speaker) + (1|syllable)

formants could not be retrieved for N = 15 items, resulting in N = 304 left for further analysis. Given the aforementioned observations about formant transitions preceding /r/ in target words *sira* and *sura*, formant measurements for these particular words were taken in the middle of the initial, steady-state part of the vowel (rather than at the midpoint of the vowel as a whole, as was the case for the other target words).

There was no interaction of speaker group with stress (LRT: $\chi^2(1) = 1.18$, $p = 0.27$).[9] There was also no main effect of group (LRT: $\chi^2(1) = 2.4$, $p = 0.11$).[10]

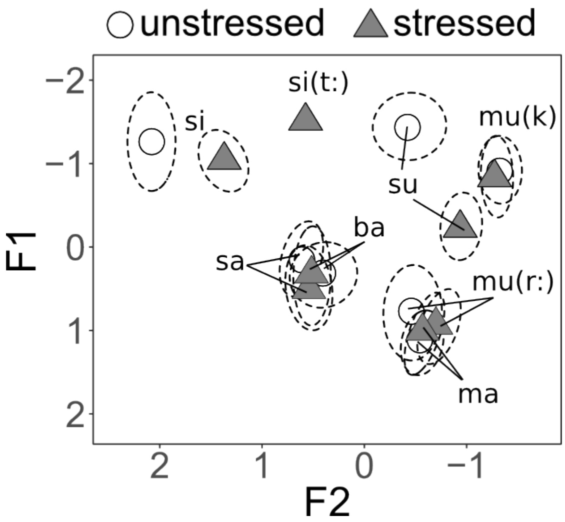

Figure 4.4: Mean formant values (Lobanov-normalised) for N = 304 target vowels, ellipses indicate 1 SD.

Figure 4.4 shows the stressed and unstressed vowels within a Lobanov-normalised vowel space. Firstly, there is a high degree of overlap for several matched syllable pairs: i) *sada* ~ *sadat*, ii) *bashar* ~ *bashart*, iii) *marra* ~ *marrart*, and iv) *muka* ~ *mukat*.

On the other hand, there is an obvious separation between stressed/unstressed vowels in syllables *si* and *su*. This is confirmed statistically with stressed *si* having higher F1 (LRT: $\chi^2(1) = 5.20$, $p = 0.02$) and lower F2 (LRT: $\chi^2(1) = 61.19$, $p < 0.001$), and *su* also having higher F1 and lower F2 when stressed (LRTs: $\chi^2(1) = 51.8$, $p < 0.001$, and $\chi^2(1) = 30.25$, $p < 0.001$, respectively).[11] This effect persisted despite the effort to measure vowel quality in the steady-state portion of the vowel which presumably is less influenced by coarticulation than the exact midpoint. I will return to this finding below.

More surprising, perhaps, is the degree of overlap between the initial vowel in the word pair *murra* ~ *murrin* and with the initial vowel in *marra* ~ *marrart*. This overlap indicates that speakers produce a vowel similar to phonological /a/ (which is phonetically close to [ɑ]) in both cases. On the one hand, it is possible that some speakers indeed produced the same phonological vowel in *marra* and *murra*, since there was no

[9] Lobanov-normalised F1/F2 ~ stress * group + (1|speaker) + (1|syllable)
[10] Lobanov-normalised F1/F2 ~ stress + group + (1|speaker) + (1|syllable)
[11] Lobanov-normalised F1/F2 ~ stress + (1|speaker)

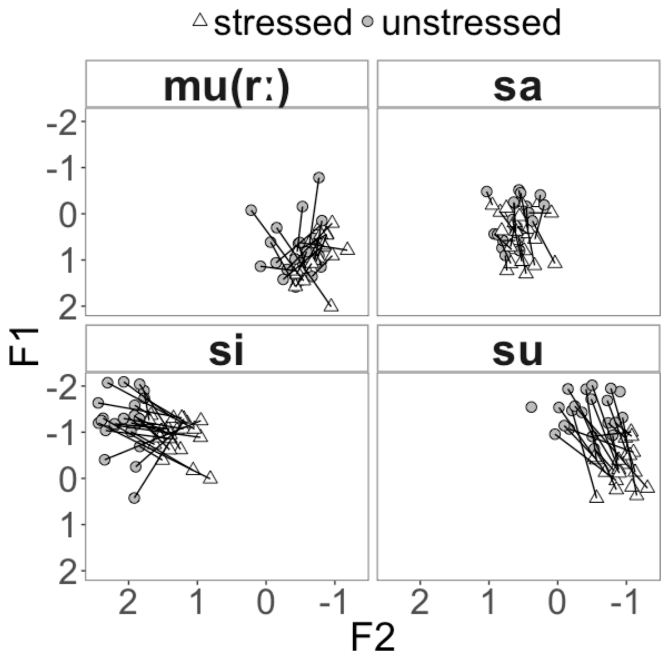

Figure 4.5: All individual tokens of target vowels for the 4 syllables *mu(rː)*, *sa*, *si* and *su*, with lines linking individual speakers' stressed/unstressed renditions of the same vowel. Lobanov-normalised values.

vowel diacritic in the written stimuli, and this is confirmed by auditory impressions. On the the other hand, irrespective of the phonological or phonetic status of the vowel, it is clear that stress status does not result in clear differences *between* members of either of these vowel pairs. For *murra* ~ *murrin*, stress status did have some effect, with lower F2 when the syllable is stressed (LRT: $\chi^2(1) = 12.36$, $p < 0.001$).[12]

Finally, differences were found between stressed and unstressed *sa*, with stressed *sa* having higher F1 (LRT: $\chi^2(1) = 12.87$, $p < 0.001$) and lower F2 (LRT: $\chi^2(1) = 6.0$, $p = 0.01$).[13] For none of the other syllable comparisons were there significant differences between stressed and unstressed vowel F1 and F2.

Figure 4.5 allows for a closer examination of the four vowel pairs that appear to differ as a function of stress. Individual speakers consistently differentiate the stressed from unstressed vowels in *sira* ~ *sinat* and in *sura* ~ *sudan*, which confirms that this is a robust effect that needs explanation.

The differences between the non-high vowels /a/ in *sada* ~ *sadat* and /u/ (realised as [ɑ]) in *murra* ~ *murrin* seem to be of a different kind. In *murra* ~ *murrin*, most of the effect seems to be carried by a small number of speakers. Additionally, the fact that this difference concerns a front/backness distinction might be explained away by appealing to coarticulation with the following vowel. Unstressed /u/ in *murrin* is followed by a

[12]Lobanov-normalised F1/F2 ~ stress + (1|speaker)
[13]Lobanov-normalised F1/F2 ~ stress + (1|speaker)

high front vowel, which might result in its raised F2 values compared to stressed /u/ in *murra*, which is followed by a low or centralised vowel. The great degree of between-group overlap for the target vowels in *sada*∼*sadat* however requires closer scrutiny. While averages suggest that stressed /a/ had 41 Hz higher F1 values and 40 Hz lower F2 values than its unstressed counterpart, the variation among speakers is high, with some speakers producing the reverse pattern from others (i.e. a number of male speakers producing higher F2 in stressed syllables). Whatever differences there are, they are not robust across speakers. I take this to mean that any significant differences between stressed and unstressed *sa* should not be interpreted as meaningful, as all speakers are not comparable in their behaviour (compare with speakers' more uniform behaviour with respect to *si* and *su*).

For these high vowels /i/ and /u/, the pattern of higher F1 and lower F2 for the unstressed member is reminiscent of the effect of pharyngealisation, which typically occurs due to a preceding pharyngealised (emphatic) consonant (Al-Tamimi 2017). The present stimuli did not contain pharyngealised consonants.[14] A potential explanation for the effect on F1 and F2 is anticipatory coarticulation with /r/, which follows both stressed vowels (in *sira* and *sura*) but not the unstressed ones (in *sinat* and *sudan*).

In sum, for three out of seven syllable pairs (excluding the eighth syllable *si(tː)*) there were no F1 and F2 differences between stressed and unstressed vowels. For the other four, I argued that any apparent effects of stress on formant values could be explained by coarticulation effects. These results together do not provide evidence to support the hypothesis that vowel quality reliably distinguishes between 'stressed' and 'unstressed' syllables.

4.3.3 Intensity

Again, there was no interaction between stress and group (among the four models, the one closest to significance was the one for z-scored peak value, LRT: $\chi^2(1) = 1.51$, $p = 0.21$).[15] In the models with absolute values, there is a main effect of group, with monolinguals producing higher intensity (e.g. for peak intensity, LRT: $\chi^2(1) = 5.65$, $p = 0.02$).[16]. This difference is not reproduced in the models used z-scored values (e.g. for peak intensity, LRT: $\chi^2(1) = 0.16$, $p = 0.68$).[17] Even if monolinguals generally produce speech that results in increased loudness (which is by no means to be concluded from these findings alone), the lack of an interaction with stress means that it has little consequence for the present experiment. Group variability (or, at a more essential level, speaker variability) in intensity patterns is to be expected anyway, and

[14]While both pharyngealised [sˤuːrə] and non-pharyngealised [suːrə] are real Arabic words, the stimuli were presented in written Arabic which uses different letters for these sounds so that it was clear that the intended word was [suːrə] not [sˤuːrə]). [sˤiːrə] moreover is not a real word as opposed to [siːrə]. All in all it is unlikely that pharyngealisation is the real explanation of the identical effect in both word pairs.

[15]intensity (mean/peak, absolute/z-scored) ∼ group * stress + (1|speaker) + (1|syllable)

[16]intensity (mean/peak, absolute) ∼ group + stress + (1|speaker) + (1|syllable)

[17]intensity (mean/peak, z-scored) ∼ group + stress + (1|speaker) + (1|syllable)

will simply have to be taken into account in any further models.

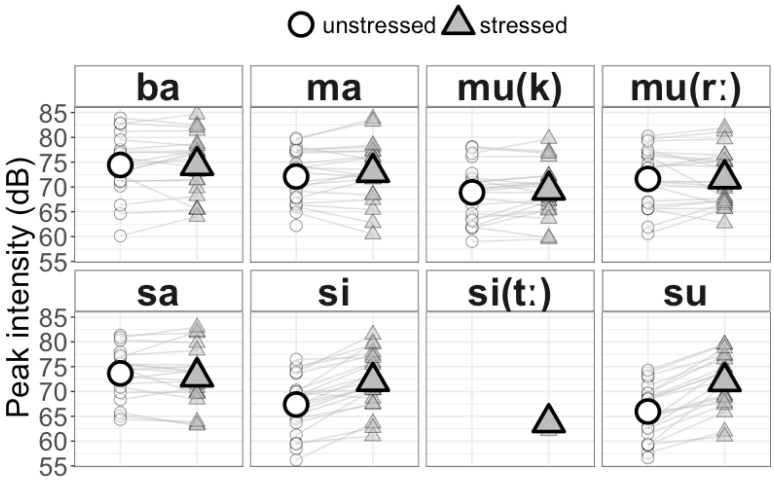

Figure 4.6: Peak intensity (dB) of vowels, per speaker, as a function of status (stressed/unstressed), tokens of *si* and *su* excluded. Lines link productions by the same speaker, large dots represent mean per syllable.

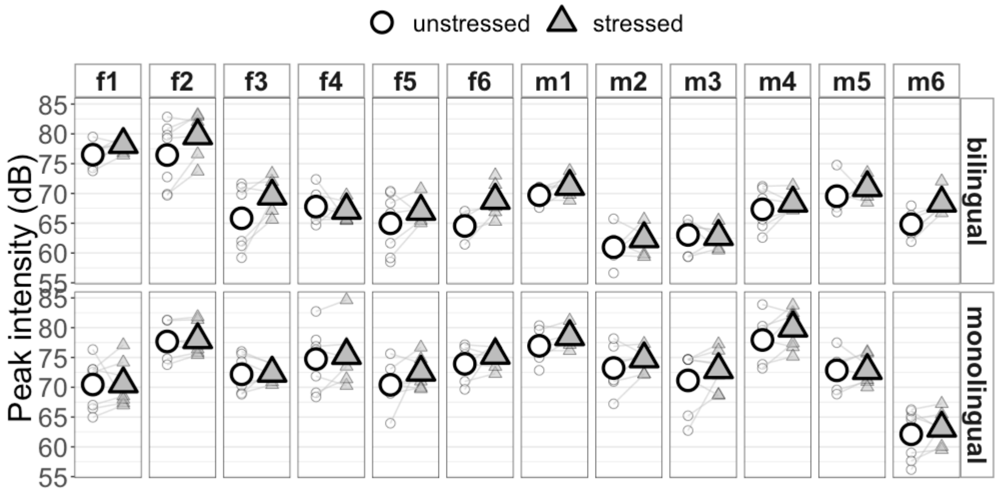

Figure 4.7: Peak intensity (dB) of vowels, per speaker, as a function of status (stressed/unstressed), tokens of *si* and *su* excluded. Lines link productions of the same vowel, large dots represent mean per speaker.

More interestingly, perhaps, there was also a significant main effect of stress on the intensity of a vowel in all of the models (e.g. for z-scored peak intensity, LRT: $\chi^2(1) = 29.00$, $p < 0.001$).[18] Once the stressed/unstressed members of a given vowel

[18] intensity (mean/peak, absolute/z-scored) ~ group + stress + (1|speaker) + (1|syllable)

pair are visualised, however, as in Figure 4.6, it appears that this effect is mainly carried by the syllables *si* and *su*. Without these syllables, the significance disappears (e.g. for z-scored peak intensity, LRT: $\chi^2(1) = 1.69$, $p = 0.19$). This particular difference can easily be interpreted not as a correlate of stress, but rather as a side-effect of the formant transition to /r/ in *sira* and *sura*, which causes an increase in sonority at the end of the vowel relative to the unstressed member.

A depiction of individual speakers' behaviour provides further evidence against the hypothesis that presumed stressed syllables are enhanced. As can be seen in Figure 4.7, most speakers do not clearly differentiate between stressed and unstressed vowels in terms of intensity, reflected by the varying directions of the lines that link unstressed and stressed members of a syllable pair. While there are a few speakers that do seem to mark stressed syllables more consistently with higher intensity (bilinguals f6 and m6), there is no overall trend towards a separation between stressed and unstressed syllables.

One final aspect of the possible link between stress and (enhanced) intensity was considered, by comparing the peak intensity in the target (initial) vowel compared to that of the second vowel in the word. The possibility exists that any intensity asymmetry between the initial and final syllable is augmented in cases where the first syllable is stressed. Accordingly, the intensity differential between the target and second vowel was calculated for the three word pairs *bashar*~*bashart*, *marra*~*marrart* and *muka*~*mukat*. These word pairs each have an initial and final vowel that stays constant across stress conditions (i.e. only the coda changes in the second syllable). Figure 4.8 shows the peak energy differential (first minus second vowel) as a function of the status of the first vowel (stressed versus unstressed).

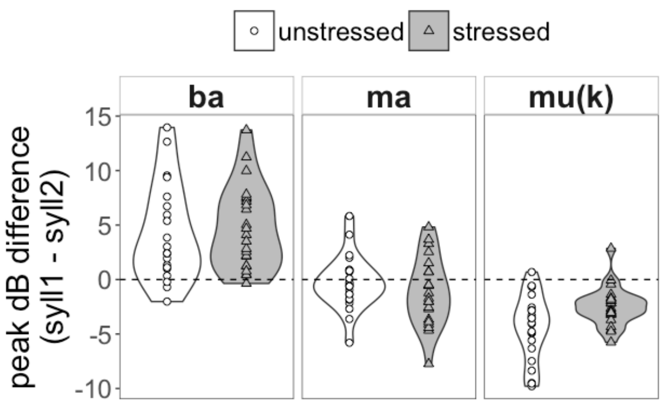

Figure 4.8: Intensity differential between initial and second syllable as a function of initial syllable status (stressed/unstressed). Data subset consisting of the three syllables *ba*, *ma* and *mu(k)*.

Differences can observed here between the word pairs but not as a function of stress, i.e.*within* word pairs. For example, for *ba* in *bashar*~*bashart*, the initial vowel almost always has higher intensity than the second, reflected by positive values for the difference in both the stressed and unstressed conditions. The reverse is true for *mu* in *muka* ~ *mukat*.

The absence of a consistent effect as a function of stress in these cases lends further credibility to the above results that also failed to reveal consistent stress-related differences in intensity between syllables.

4.3.4 F0

In this section I will first report on the global F0 contours characterising the utterances in which target words were embedded. After this I will take a look at static scaling properties of the F0 contours in terms of mean and peak measurements in target vowels.

As described in Section 4.2.4, global contours were normalised for syllable duration. The syllable boundary is taken to be the onset of the second consonant, which occurred in intervocalic position, the exceptions being *sitta* ~ *sittat*, which, as mentioned previously, were typically realised as [sːtːə] ~ [sːtːaːt]. In order to facilitate the comparison, the onset of the geminate /tː/ was taken to be the syllable boundary in these words.

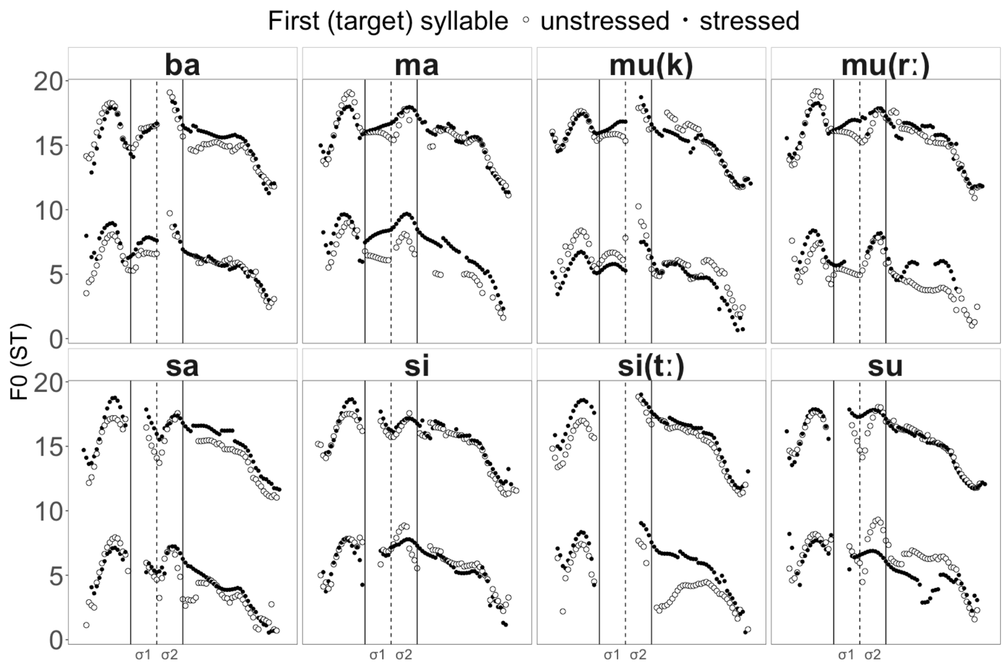

Figure 4.9: Phrasal intonation contours (averaged and normalised for duration) across N = 251 target utterances without pauses, male contours (bottom white and black lines) and female contours (top black and white lines) separated.

Figure 4.9 shows time-normalised mean contours for all utterances in which target words were not followed (or preceded) by a pause (N = 251, as described above). A first observation is that the part of the contour that occurs on the target word is similar in all contexts, suggesting that there is no obvious effect of stress on the F0 characteristics of target words. It is at least expected that the words in this dataset are not subject to postlexical prominence marking, as they occurred in postfocal position, and at first sight this is confirmed here. Nevertheless, there is one unexpected visual difference between female and male speaker behaviour, for at least one syllable pair: The male-speaker contour for unstressed *mu* in *mukat* (white) is realised at a higher level than stressed *mu* in *muka* (black), while the female contour shows the opposite pattern. This observation led to the inclusion of random slopes for the interaction of stress with participant sex in the statistical models discussed below.

It can also be observed that target words exhibit a small rise at their right edge, or somewhere on the second syllable in general, and there are several possible explanations for this. One theoretical possibility, namely that the pitch movement represents postlexical prominence, is unlikely given the focus structure of the phrase as just mentioned. Another is that the higher pitch on word-final syllables is in fact a marker of word-final stress. I will return to this possibility below. A more likely possibility is that the rise represents an edge-marking tonal event, rather than a prominence-marking event, given that they occurrence in a context where the target words are postfocal and given. In this case, the observed patterns would highlight the difficulty in obtaining pitch-neutral stimuli despite careful experimental design.

In the present experiment target syllables are word-initial and thus do not themselves carry this rise. It would have been problematic, however, if some target syllables had been word-final and others occurred in non-final position (as in fact in the stress experiment conducted for TB in Chapter 3), as this would have implied that final syllables are distinct from initial ones by virtue of the rise in the form of a positional effect. This brings me to some of Boudlal's (2001) results. As mentioned in Section 4.1.1.1, Boudlal proposes that stress targets the final syllable in cases where a word is produced in isolation, while stress may be elsewhere in sentence context. Given that his words in isolation were produced in a list, it is not surprising that these words were produced with what looks like a continuation rise, see Figure 4.10, so that all target words were characterised by high pitch in the final syllable. Therefore, what Boudlal (2001) considers an effect of stress rather seems to be a positional difference, and high final pitch in this context should not be interpreted as a correlate of stress.

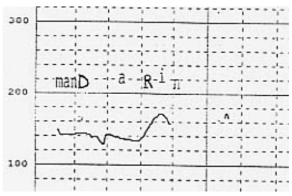

Figure 4.10: Example F0 contour on word in isolation, from Boudlal (2001: 346).

In the present case, it cannot be concluded for certain that the pitch event is edge-marking rather than serving the purpose of (lexical stress) prominence marking. For the purpose of this experiment, the presence of a consistent prosodic event on the final syllable, irrespective of its actual nature, cannot be considered evidence in favour of the view that stress is weight-sensitive and variable between penultimate and final position. Returning to the experimental question about correlates of stress in MA as conceived of by Benkirane: The most important aspect for present purposes is that these target (initial) syllables are comparable in terms of their pitch properties, and they seem to be.

Several static measures of target vowels were taken (N = 315), in the form of peak and mean F0 (in ST), and as before models were run on both absolute and z-scored values. Pausing is included in these models as a fixed effect. Firstly, there was no interaction between pausing and stress (e.g. for absolute peak values, LRT: $\chi^2(1) = 1.54$, $p = 0.21$).[19] There was also no main effect of the presence of a pause on the F0 values of target vowels (LRT: $\chi^2(1) = 0.68$, $p = 0.17$).[20]

There was however an interaction between group and stress status. In bilingual speakers, stressed vowels had F0 values that were around 1 ST higher than unstressed syllables, whereas the effect was less pronounced in monolingual speakers, who exhibited a difference of 0.6 ST (the 'least' significant was z-scored mean F0, LRT: $\chi^2(1) = 3.86$, $p = 0.05$).[21] The general effect therefore is one of slightly increased F0 on the stressed syllables. Figure 4.11 shows the mean F0 values for the N = 315 target vowels for which pitch was correctly retrieved.

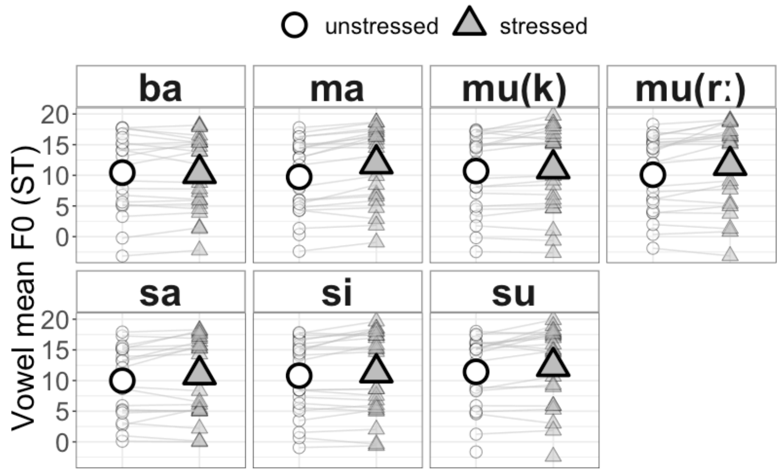

Figure 4.11: Mean F0 (ST) of vowels, per syllable, as a function of status (stressed/unstressed). Lines link productions by the same speaker, large dots represent mean per syllable.

Similarly to the F0 differences in the case of the experiment on Tashlhiyt (Chapter 3), predicted differences of the above-mentioned magnitude might not represent perceptually 'robust' enhancement, as differences of 1 ST to be an absolute minimum in order for listeners to distinguish dynamic pitch movements. Nevertheless, a difference exists,

[19] F0 (mean/peak, absolute/z-scored) ~ pause * stress + stress * group + (1|speaker) + (1|syllable)
[20] F0 (mean/peak, absolute/z-scored) ~ pause + stress * group + (1|speaker) + (1|syllable)
[21] F0 (mean/peak, absolute/z-scored) ~ pause + stress * group + (1|speaker) + (1|syllable)

and it might be enlightening to investigate speaker-specific behaviour with respect to F0 as a possible correlate of stress in MA.

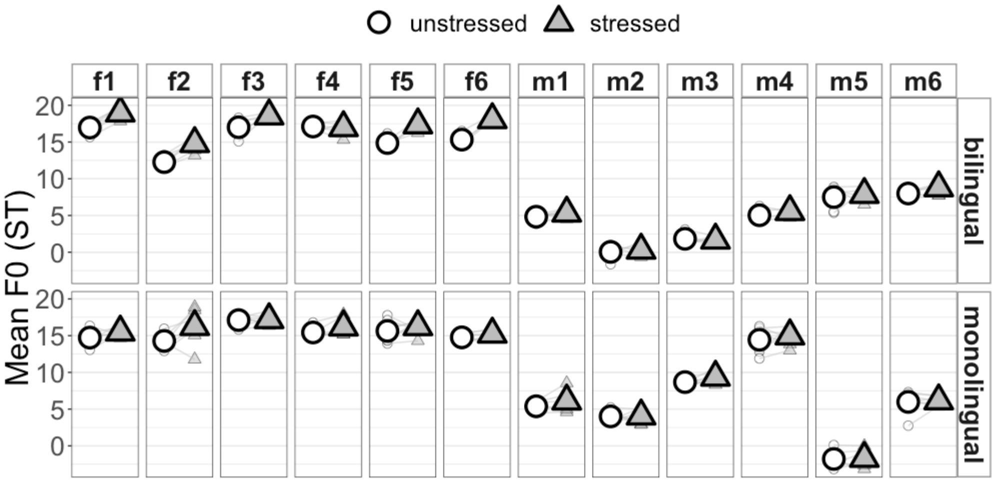

Figure 4.12: Mean F0 (ST) of vowels, per speaker, as a function of status (stressed/unstressed). Lines link productions of the same vowel, large dots represent mean per speaker.

Figure 4.12 shows that about half of the speakers consistently produce higher peaks F0 on 'stressed' syllables, including almost all bilingual female speakers, as well as monolingual f2 and perhaps several others. These bilingual speakers thus appear to carry much of the effect for the bilingual group according to which stressed syllables are produced with increased F0. The remaining speakers display the otherwise familiar pattern of making no consistent distinction between stressed and unstressed vowels.

It follows that, while there is a statistical effect of enhanced F0 on stressed syllables in MA, this effect is not a very robust acoustic correlate, neither in terms of the scale of the difference (involving no more than 1 ST) nor in terms of being observed reliably across the speaker population.

4.4 Discussion

For none of the acoustic correlates discussed in this chapter (duration, vowel quality, intensity and F0) were there convincing differences between presumed stressed and unstressed syllables. In order to conclude that stressed syllables stand out acoustically, there would have to be consistent differences across the board. More concretely: A lexicon-wide effect with acoustic enhancement of stressed positions would require stressed syllable members to stand out from unstressed ones in most if not all words. In the present experiment, the only (marginally) consistent differences were found for the stressed and unstressed counterparts of the syllables *si* and *su*. The differences in duration and intensity however could be attributed to segmental-contextual effects.

Vowel quality in these syllable pairs was also different between stressed and unstressed members, but not in the expected direction, requiring an explanation in terms of coarticulation rather than stress status. In short, any acoustic differences in general are best interpreted as parasitic on this vowel difference.

Across the board, it is crucial to note that no acoustic parameter was used consistently across syllable pairs to mark the distinction between stressed and unstressed syllables, and no syllable was consistently enhanced by multiple acoustic parameters. Additionally, for the differentiation of stressed and unstressed syllables to be robust, individual speakers would be expected to systematically cue this distinction. In this experiment, no speaker made a consistent distinction (either using multiple cues to enhance stressed syllables, or using a single cue consistently), which suggests that speakers do not produce two acoustically distinct phonological categories 'stressed' and 'unstressed'.

While the present findings can thus not be considered to provide evidence *in favour of* the existence of lexical stress, they also cannot be taken to provide evidence *against* it. This is a problem inherent to null results, but it is augmented further by the fact that even if MA lacks lexical stress as envisaged by Benkirane (1998), lexical stress could still, potentially, be captured by appealing to another stress rule. One other testable prediction about stress that has been brought forward is the one that considers stress to be final, irrespective of syllable weight (see Section 4.1.1.1). In order to test this prediction, a different experimental set-up would be needed, where the stimuli should contain identical syllables contrasted in a presumed stressed position and in a presumed unstressed position. This could for example involve stimuli such as the following non-words /daˈba/ ~ /baˈka/, in which the target syllable *ba* could be compared in final (stressed) and in non-final (unstressed) position. Whether this stress rule is correct, however, is somewhat doubtful: To my knowledge there is no concrete evidence that suggests that MA has a stress-by-position system.

Returning to the present null results, in the context of past and present work these appear to be best interpreted as accurately reflecting a situation in which MA lacks lexical stress, as they are in fact compatible with multiple additional observations. Firstly, native speakers have shown varied judgments on stress position in MA and there is a century-long disagreement among scholars on the proper representation of stress. Clearly, the concept of stress in MA is an elusive notion, which is already a good indication that it might not play much of a role in the phonology of the language. The second observation has to do with the intonational phonology of MA and the generally held assumption that stressed syllables serve as docking sites for postlexical pitch accents. In an experiment conducted by Yeou, Embarki & Al-Maqtari (2007), in which specific words in read sentences were contrastively focused, MA displayed somewhat different patterns from Kuwaiti and Yemeni Arabic (varieties that uncontroversially have stress). Specifically, the intonational movement accompanying the relevant words was not as closely tied to the presumed stressed syllable in MA as it was in the other varieties.[22]

[22]The authors do not indicate how they decided on the position of stress, but it seems that they took syllables with long vowels to be stressed. Vowels that were long according to their transcription occurred in penultimate position (/CVː.CV/) or in closed syllables in final position (/CV.CVːC/), so

Typically, intonational movements that have a prominence-lending function (e.g. pitch accents used to signal contrastive focus) tend to be localised movements closely aligned with stressed syllables, as also seems to be the case in Kuwaiti and Yemeni. It is interesting that the same is not true for Moroccan Arabic, as this behaviour is very much compatible with the idea that stress really is absent. Chapter 7 will address the interpretation of intonational movements in MA in more detail.

In sum, the present findings, reflecting an absence of consistent acoustic enhancement of presumed stressed syllables, seem valid null results.

4.5 Summary and conclusion

This chapter investigated correlates of lexical stress in Moroccan Arabic by contrasting presumed stressed syllables with unstressed ones, according to a dominant view on word stress in Moroccan Arabic (Benkirane 1998). According to this view, the penultimate syllable of a word is stressed, unless the final syllable is heavy, in which case stress is final. The results from the experiment reported here could not provide evidence in favour of acoustic enhancement of 'stressed' syllables according to this rule: Excepting a few differences that could be explained by other factors, acoustic properties of stressed syllables did not consistently differ from those found in unstressed counterparts, in terms of either duration, vowel quality, intensity or F0. This lack of consistent enhancement was observed across the board. Specifically, effects were absent for most syllable comparisons (indicating that the effect does not hold across the lexicon), and systematic effects were also absent for individual speakers (indicating that speakers do not have separate realisational categories for stressed/unstressed syllables).

While the present null findings cannot be directly interpreted as providing evidence in favour of the opposite claim, namely that Moroccan Arabic lacks lexical stress, the synthesis of the present results with evidence from other sources points strongly towards an interpretation in which lexical stress does not play a large role, if any, in the phonology of the language. In conclusion, the results in the present chapter are compatible with an absence of lexical stress in Moroccan Arabic.

that their understanding of stress for these words at least matches Benkirane's (1998) predictions.

Part III

Postlexical prominence production

5 Prominence in question word interrogatives

5.1 Introduction

5.1.1 Preliminaries

One of the aims of this thesis is to characterise the nature of prosodic prominence at the postlexical level in Tashlhiyt Berber and Moroccan Arabic. This question is addressed experimentally for the intonational marking of question words (qwords) in the following two chapters. The present chapter serves to motivate why specifically question words (or wh-words) form a good testing ground for the investigation of postlexical prosodic prominence.

The intonation of questions in general has a long research history, but most of this work has been concerned with one aspect of question intonation only, namely the right-edge-marking of yes/no questions. As I will go on to show, qword questions have received comparatively little attention. In particular, a detailed review of the prosodic prominence of qwords is so far noticeably absent, a situation that the second part of this chapter seeks to remedy (Section 5.3).

In stark contrast to the situation for intonation is the amount of work dedicated to the syntactic and semantic properties of qword questions, which have been investigated in great detail both crosslinguistically and for many individual languages. One particularly interesting discussion concerns the structural or inherent prominence of qwords, which is sometimes referred to directly in terms of focus (see Section 5.2.2.2). Syntactically, the prominence of qwords comes about through syntactic movement (at least in approaches that stipulate movement), with qwords often considered to move to (or occur in) a specific position that allows for their intended interpretation. Semantically, the prominence of qwords follows from their contribution to interrogative meaning. Finally, some aspects of qwords' (morpho-)syntactic behaviour also suggest that qwords are similar to focused constituents.

In reviewing the various aspects of qword interrogative structure, I will also discuss one of the more puzzling observations about the correspondence between the *structural* (non-prosodic) and *prosodic* prominence of qwords. Specifically, the observation has been made that the two types of prominence do not coincide in English, as qwords are not typically the words with the highest degree of phrasal prosodic prominence (i.e. they do not tend to get the nuclear accent). In many other languages, however, qwords in interrogatives are consistently the most prominent words in the phrase. Unfortunately, the facts of English have for a long time confused researchers. In the next sections, I will show that this presumed mismatch between structural and prosodic prominence is not very pervasive crosslinguistically.

5.1.2 Overview of chapter

In this chapter I make a distinction between aspects of qword interrogatives that can be allocated to non-phonological linguistic structure (syntax, semantics and pragmatics), and those that are prosodic–phonological. The non-phonological aspects will be discussed in the next section, 5.2, where I give a brief overview of insights from syntactic, semantic and pragmatic literature. The prosodic-phonological aspects of structure will be discussed in the subsequent section, 5.3. This section is divided into two parts matching the division found in past research (which incidentally is also a distinction that lies at the heart of AM analyses of intonation), namely between edge-marking (Section 5.3.1) and prominence-marking (Section 5.3.2). The discussion of the prosodic- and prominence-marking of qwords crosslinguistically (Section 5.3.2) is based on all the sources known to me at the time of writing. The sources can be subdivided into experimental and descriptive studies:

- Experimental studies (Section 5.3.2.1): This set includes the few works that explicitly investigate prosodic properties of qwords with some quantitative aspect beyond mere listing or description of intonation patterns. Claims in these works are based on well-described datasets involving data from multiple speakers.

- Descriptive studies (Section 5.3.2.2): This set contains all other resources, including i) general descriptions of intonational systems with a few single examples of qword contours (most contributions in Jun's *Prosodic Typology* volumes (Jun 2005c, 2014b), and in Hirst & Di Cristo's (1998) *Intonation Systems*), ii) more elaborate descriptions of qword interrogatives in specific languages that do not provide detailed information about the actual data (e.g. Varga 2002) or make little attempt to quantify patterns (e.g. Frota 2002).

Following a discussion of the inventory of prosodic patterns that may be used crosslinguistically in the marking of qwords, I will finish this chapter by raising some general questions about qword prominence that will be answered for Tashlhiyt Berber and Moroccan Arabic in the following chapters.

5.2 Syntax, semantics and pragmatics

Syntactic, semantic and pragmatic aspects of qword interrogatives are closely intertwined. In discussing them as separate aspects of linguistic structure I do not wish to suggest that a clear division of labour between these linguistic domains is always justified, especially when it comes to focus. In the next few sections I will nevertheless separately review a few insights from the different fields before moving on to what links them all together in Section 5.2.2.2 on focus.

5.2.1 Syntax

A large amount of work in syntax has been concerned with characterising the phrasal positions in which qwords may occur in different languages. There seems to be a consensus on the types of patterns that exist crosslinguistically, which can be allocated to roughly two types of structural behaviour (Chomsky 1995: 68; Dik 1997b) (but see discussion below on whether there might in fact be three groups):

Type 1 Languages in which qwords in interrogatives are found at (or close to) the left edge of the phrase;

Type 2 Languages in which qwords are found in the same position as corresponding non-qword constituents ('in-situ' in frameworks that presuppose syntactic movement)

Type 1 covers many languages, some 70% of the world's languages according to Dik (1997b: 283) (although no details are given to substantiate this claim).[1] Depending on the definition, this group of languages may include those that front all qwords in multiple qword questions, as in Slavic, as well as languages like English, in which one qword only is fronted.

Considering only questions with a single qword, the Type 1 group includes most Indo-European languages as well as varieties of Berber and Arabic. An example illustrating the canonical phrase-initial position for qwords in Tashlhiyt is given in (1). A typical declarative, as in (2), has the object(s) corresponding to the qword 'what' in post-verbal position[2]

(1) *ma* *rad* *za* *n-skr* *ɣila* *?*
what FUT then 1PL-do.AOR now ?
'What are we going to do now?'

(2) *rad* *n-skr* *tiz:wa* *aɣjul* *tagant*
FUT 1PL-do.AOR bees donkey forest
'We'll do the bees, the donkey, (and then) the forest.'

Type 2 covers languages such as Mongolian, Japanese, Chinese and Korean, where the qword occurs in the same position as the matching constituent in the corresponding declarative. Compare (3) and (4), based on Cheng 2003: 103):

(3) *Hufei* *mai-le* *shenme*
Hufei buy-PFV what
'What did Hufei buy?'

[1] In fact, the 70% represents those languages that are 'Q-focus' in his definition, a category which also includes discourse-configurational languages that may have a focus position that is not phrase-initial.

[2] Examples are taken from map task interactions available online as part of the CoTaSS corpus (Bruggeman & Roettger 2017).

(4) *Hufei mai-le shu*
Hufei buy-PFV book
'Hufei bought a book.'

Finally, there are languages not subsumed in either one of these categories because they seem to allow qwords in multiple positions, such as Maltese (Vella 2007), and Zulu, Malagasy, and French, which have also been called 'optional in-situ' languages (Sabel & Zeller 2006). Whether these languages should be given their own category is beyond the scope of this discussion. For present purposes it is important to note that most languages of the Type 1 group in fact also allow qwords to occur in non-initial position, and specifically in a position that is typical of qwords in Type 2 languages, so that the English counterpart of (3) is also permissible, as in (5):

(5) Hufei bought what?

This utterance can still be interpreted as a direct interrogative in English, but it has a somewhat different interpretation from a canonical question with an initial qword. In syntactic work such questions are often relegated to the echo question category but perhaps the better characterisation would be that non-canonical qword placement cannot be used for out-of-the-blue questions (as isolated written examples tend to be). Corpus studies or studies on interactional data often reveal that non-initial qwords in (what are assumed to be) Type 1 languages occur not only as echo questions but also in other contexts. For example, Germanic languages such as English are sometimes considered to have a category called reference questions, as exemplified in (6) from Bartels (1997: 4):

(6) A: I just talked to him last night.
B: You talked to WHO last night?

Additionally, languages more closely related to Berber and Arabic, such as Hausa (Afro-Asiatic, Chadic), clearly allow qwords to occur in non-initial default position (Jaggar 2006). In fact, Tashlhiyt Berber and Moroccan Arabic also allow for some variation, based on my own observations of qword questions in daily interaction as well as in the CoTaSS corpus (Bruggeman & Roettger 2017) and the IVAr corpus (Hellmuth & Almbark 2017), respectively. The context in (7) exemplifies the usage of a non-initial non-echo question in a Moroccan Arabic map task:[3]

(7) A: I've passed Bilal mosque, right?
B: You've passed Bilal mosque.
A: You told me to go left. Turn around and what do you see?
dur ʕlih ʁajban lik ʃnu ?
turn around will see what
'Turn around and what do you see?'

[3]IVAr filename: mobi-mp2-m5m6

B: Now, the municipal palace.

In the example above, the target question does not serve to express any of the meanings that are associated with echo questions, such as surprise, disbelief or a request for repetition of earlier-mentioned information. The question moreover elicits a simple response that provides the information requested, which is a further indication that this is a felicitous question at this point in the conversation.

The distinction between questions with non-initial qwords that are echo questions and those that are not echo questions, therefore, is a fine but crucial one. Nevertheless, the occurrence of non-initial qwords in Type 1 languages is rather uncommon, and probably highly dependent on discourse context and pragmatics. It is further possible that the occurrence of non-initial qwords may also depend on a range of other factors including speech register, style, and individual speaker preferences. These are factors that have been invoked to explain patterns of qword placement in French by Baunaz & Patin (2009), one of the aforementioned 'optional in-situ' languages.

There are, in sum, two main points of interest here. Firstly, languages differ in the syntactic slots qwords *typically* occupy. Secondly, within a single language, variation in the location of qwords may be expected. Even languages with a default position for qwords may allow for qwords occurring elsewhere in specific contexts, without necessarily causing an echo interpretation.

5.2.2 Semantics and pragmatics

5.2.2.1 Lexical and phrase-level meaning

The semantics and pragmatics of qwords are intricately linked to discourse and phrase-level context. The lexical meaning of qwords and their contribution to phrase-level meaning will be discussed here, while discourse-level influence on the interpretation of qwords will be discussed in the next section on focus.

Qwords as interrogative pronouns are often considered to be a type of indefinite quantifier with an additional aspect of meaning (i.e. contributing to interrogativity) which they derive as a function of their prosody and/or syntax (Krifka 2011). The quantifier interpretation is supported by the observation that qwords in many languages have homonyms that function as indefinite quantifiers (such as English 'which', which may be either an interrogative pronoun or a relative pronoun, as in this sentence). Other languages use homonyms of interrogative pronouns as conjunctions, such as Italian, where a word like *perché* can be translated as either 'why', or 'because'.

On a phrasal level, qwords may occur in direct interrogatives and in embedded questions. They can also occur in non-interrogative contexts, but on some accounts qwords would not be considered qwords anymore when they do not contribute to interrogative meaning. Typically, a qword that occurs in a direct interrogative helps to performs a speech act with illocutionary force, as in (8). The meaning of direct interrogatives is often construed as the set of answer propositions (possible or true answers) (Hagstrom 2003).

When a qword appears in an indirect or embedded question, as in (9), the utterance does not invite an answer, and its semantic meaning is simply that of the proposition it denotes by itself. The semantic-pragmatic properties of qwords in such contexts are, by consequence, rather different from those in direct interrogatives. Particularly, they are not typically considered focused, in contrast to qwords in direct interrogatives (cf. Krifka 2011; Song 2017).

(8) What languages do they speak in Morocco?

(9) I wonder what languages they speak in Morocco.

5.2.2.2 Discourse-level meaning: Focus

Where syntactic, semantic and pragmatic aspects of qwords come together is with respect to focus. For present purposes, I adopt the definition of focus given by Lambrecht (1994: 208), who considers focus to be "that portion of a proposition which cannot be taken for granted at the time of speech. It is the unpredictable or pragmatically non-recoverable element."[4] Qwords under this definition can be considered foci, as they are typically a new element in the discourse and almost always a new element in the phrase (cf. van der Wal 2016). Clearly, focus has much in common with notions of (non-prosodic) prominence, since it refers to information that stands out relative to information which is known, given, or presupposed. In order to investigate prosodic prominence status, then, it is highly relevant whether qwords are inherently prominent (in the non-prosodic sense), and specifically whether they should be considered focused. I will address exactly this question in the remainder of this section.

There are multiple arguments from pragmatics and (morpho-)syntax that indicate similarities between focused constituents and qwords, with the typical conclusion that qwords in direct interrogatives are in fact focused constituents.

For one, the response to a qword question is widely assumed to have focus on the constituent that provides an answer to the qword constituent (Culicover & Rochemont 1983; Rooth 1992). This correspondence between qwords and focus in the answer is generally interpreted as meaning that qwords themselves are focused too (for an alternative view see Erteschik-Shir 1986).

Perhaps a more convincing argument comes from discourse-configurational languages, in which qwords are often required to occupy what is considered the structural focus position. These are preverbal positions in Turkish (Dik 1997b), Basque (Elordieta & Hualde 2003) and Hungarian (É. Kiss 1995), and other fixed positions in Korean (Choe 1995), Finnish (Vilkuna 1995), and various other languages (É. Kiss 1995). Further

[4] Another commonly used definition of focus, based on Rooth's Alternative Semantics (Rooth 1985, 1992), considers focus to be about alternatives: If a constituent is focused this indicates the relevance of multiple (contrasting) alternatives to the referent of that constituent in that particular context. This definition also lends itself less easily to accounting for the focal status of qwords.

observations that qwords syntactically behave like foci have been made for other types of languages, such as Type 2 languages in which qwords are like foci in terms of their (in-)sensitivity to so-called island constraints (Hagstrom 2003: 192, Krifka 2011).

A third argument has to do with the morphological marking of qwords. Some languages that have morphological markers that attach to focused constituents use this same marker with qwords, such as Wambon (Dik 1997a), Akan (Genzel 2013), Gungbe (Aboh 2007), and Lɛtɛ (Akrofi Ansah 2010).

A fourth similarity between qwords and foci is that qword interrogatives often resemble cleft constructions (clefted constituents being focused by definition). In the generative tradition in particular, clefts and interrogatives with an initial qword are often analysed as involving movement that is motivated by a focus feature. The structural similarity is nevertheless also observed by functionalists (Dik 1997b: Chapter 13). Particularly relevant to this thesis is that the syntactic parallel has been observed to hold in Moroccan Arabic (Ouhalla 1999), and in Berber varieties such as Tarifit and Tamazight (Stoyanova 2004, and Penchoen 1973 as cited in Dik 1997b: 328f). Relevant details will be discussed in Chapters 6 and 7.

5.2.2.3 The qword/focus correspondence

Two main views on the qword/focus correspondence can be found in the literature.

The first holds that qwords (in interrogatives) act as foci across the board, an idea that has been called "generally assumed" (É. Kiss 1995) (see also Erteschik-Shir 1986; Dik 1997b). In a few cases it has even been explicitly stated as a universal (Horvath 1986; Song 2017). The 'qword = focus' idea has however also been challenged, for example by Büring (2012) and Erteschik-Shir (1986, 2007). It seems no accident, however, that both authors' counterarguments are informed primarily by English. English typically marks focused constituents by means of intonational prominence and specifically by means of a nuclear pitch accent. However, qwords in English do not typically receive this nuclear accent. Qwords do however often receive some sort of postlexical prominence, in the form of a prenuclear rather than a nuclear pitch accent. Both authors try to resolve this 'puzzle' by stipulating that qwords therefore are not focused *in English*, but this reasoning risks being circular. While it is true that foci in English can generally be identified by intonational prominence, this is simply not always the case. This invalidates the idea that the absence of the main prosodic prominence on qwords should preclude a focus interpretation. Several examples can be given of contexts in which no consistent one-to-one mapping between intonational prominence and focus is found in English (for a full overview see Zimmermann & Onea 2011). On the hand there are cases of 'focus without accent' in which focused constituents do not receive an accent (or do not receive the main prominence-marking accent), such as:

- 'focus ambiguity' of phrase-final nuclear accents. This kind of accent may indicate all-new broad focus, as well as various kinds of narrow focus including on the single word that carries the accent;

- second occurrence focus, where a focused constituent which has been previously mentioned is not marked by the main pitch prominence.

Additionally, English also has non-prosodic focus strategies, such clefts and pseudo-clefts. This can be considered another kind of focus without accent.[5]

On the other hand there are cases of 'accent without focus' in which pitch accents occur on constituents that are not necessarily focused themselves. Such accent placement may be due to (metrical) phonological considerations or so-called focus projection (see also Section 2.4.3.2 for a discussion of patterns of accent distribution).

The above set of arguments strongly suggests that the absence of intonational prominence on a given constituent cannot alone serve as conclusive evidence against its being focused. An alternative explanation of the typical absence of intonational prominence on English qwords is that qwords function as non-prosodically marked foci. This option is not entertained by either of the aforementioned authors, but there is no obvious reason why this possibility should be discarded. Firstly, qwords are a morphologically distinct group of words, which means that they could be identified as foci even in the absence of intonational prominence (unlike, say, a random noun like 'apple'). Secondly, qwords are all the more distinct due to their occurrence in a specific syntactic slot near the left phrasal edge, which also distinguishes them from homologous relative pronouns.

In conclusion, it seems precipitous to require qwords to be prosodically the most prominent in the phrase in order to qualify as foci, even in languages that usually mark focus by means of intonational prominence.[6] Ideally, there should be pragmatic diagnostics to the focus status of a constituent, which then may or may not be supported by intonational prominence or other grammatical markers of focus (see also Lambrecht 1994; Zimmermann & Onea 2011: 208).

In sum, the absence of prosodic prominence on a qword alone should not constitute evidence against it being focused (in English or crosslinguistically) Taking this further, concluding that qwords are focused by definition, might not be warranted either. The next section will address the issues relating to generalisations about the focal status of qwords in some more detail.

5.2.3 Crosslinguistic generalisations?

The above discussion highlights that a unified treatment of qwords crosslinguistically is difficult at the very least. Even if there are ostensible similarities between qwords and focused constituents in many languages, it is not clear that such patterns should be interpreted as reflecting a universal that attributes focal status to all qwords all languages, or even a generalisation that qwords in interrogatives always function as

[5] Clefted constituents often, but do not necessarily, attract intonational prominence (Bolinger 1986).

[6] It is moreover well-known that the presence of intonational prominence is not a crosslinguistic marker of focus, cf. Erteschik-Shir (2007: 40): "not all languages use stress to mark the focus".

foci in a given language.[7]

An additional difficulty is that there are counterexamples to most of the previously mentioned diagnostics to qword focus. A first one concerns the optionality of the use of focus markers with qwords. While there are languages that consistently use the same morphological markers for qwords and foci, as mentioned above, there are also languages that only optionally use a focus particle with a qword, such as Hausa (Jaggar 2006) and Lele, Amharic and Gungbe (Aboh 2007). Aboh (2007) explicitly proposes that languages may distinguish between focused and non-focused qwords, and attributes variation to discourse context and phrasal position. A second observation that challenges a conception of qwords as always focused is that there are discourse-configurational languages in which the qword does not occur in the designated focus position (at least under some analyses), such as Bulgarian and Catalan (É. Kiss 1995). A third difficulty is that some languages allow for the clefting of qwords, including English:

(10) What is it that they bought?

The clefting structure in (10) indicates that in this particular case, the qword is most definitely focused, but this raises questions about non-clefted qwords in English: Are these also focused, and if so, should this lead us to posit different kinds of qword focus? Does it follow that there is a difference in the focal status of qwords in languages that treat all qwords like clefts, and the status of qwords in languages that may but need not be clefted? Concrete answers to such questions have not been proposed in the literature, but a good explanation would likely appeal to different focus types. Zimmermann & Onea (2011) for example suggest that the difference between in-situ and ex-situ focus in West African languages reflects the difference between contrastive and information focus.

5.2.4 Summary

There are many arguments from (morpho-)syntax, semantics and pragmatics that support an interpretation of qwords as inherently focused in many individual languages. It is one step further to assume that qwords in interrogatives are by definition focused crosslinguistically, although this is in fact a very common view. If a uniform characterisation of qwords in terms of focus is possible at all, semantic and pragmatic arguments are likely to be needed: A mere glance at the possible crosslinguistic patterns for the syntactic and prosodic marking of qwords is enough to abandon hopes for a uniform characterisation from these corners.

At this point, it has to be concluded that qwords in interrogatives cannot be considered focused by definition. Instead, qwords' potential status as foci must be determined on a language-specific basis, and likely even on a discourse-specific basis. Given

[7] This question is further complicated by the variety of focus categories that have been proposed (e.g. information focus and contrastive focus), the discussion of which is beyond the scope of this chapter.

the well-documented crosslinguistic tendency to treat interrogative qwords as foci, it seems nevertheless reasonable to assume that qwords in any language are likely to be focused at least *some* if not *all* of the time.

5.3 Prosodic marking of question word interrogatives

5.3.1 Right-edge marking

I take right-edge marking of interrogatives here to refer to phrasal boundary marking, with the specific aim of reviewing the intonational marking that serves the function of demarcation rather than prominence.[8] Prominence marking will be discussed separately in the next section. For a more in-depth discussion of the distinction between these types of prosodic marking see Section 2.4.3.2.

The abundance of literature on the prosodic marking of interrogatives has led to the insight that yes–no questions in many languages are marked by some sort of high pitch. This can take the form either of a localised high pitch excursion, for example a final rise or final rise-fall near the right phrasal edge, or of more global characteristics, for example in terms of an overall higher register for questions than for statements, or in the form of the suspension of declination (Ultan 1978; Cruttenden 1986).[9] The prosodic marking of yes–no questions as different from statements can in some languages be attributed to the fact that yes–no questions are segmentally identical to statements: In these languages there are no morpho-syntactic cues to questionhood and the burden is on prosody to distinguish the two sentence modalities (e.g. in Italian and Greek). Most languages however (additionally) have morpho-syntactic strategies to signal yes–no questionhood, including the use of question particles and word order changes. Languages that use such strategies often still allow for phrases with declarative word order to function as yes–no interrogatives by virtue of intonation only. These are sometimes called 'declarative questions' to distinguish them from morpho-syntactically cued yes–no questions.

The prosodic marking of qword questions, on the other hand, has not been subject to the same amount of crosslinguistic scrutiny, although some general observations have been made in the literature. For example, qword questions, in contrast to yes–no questions, seem often to be marked by some kind of low pitch at their right phrasal edge (Bolinger 1978; Ultan 1978 as cited in Cruttenden 1986). This observation has been confirmed by several experimental studies: Sosa (2003), Willis (2007) and Henriksen (2014) for in total six different varieties of Spanish; Varga (2002) for Hungarian; Frota (2002) for Portuguese; Hirschberg (2000) for American English; and Grice, Baumann &

[8] Admittedly, this is a somewhat simplified approach: Right edge marking in the literature often refers to 'nuclear contours', which consist of the nuclear pitch accent followed by boundary tones at the right edge of the phrase, in which case at least part of this right-edge marking could be considered prominence marking.

[9] Even if some form of 'high pitch' is commonly attested in yes–no questions crosslinguistically, it is by no means claimed to be the only pattern. For more detailed discussion of right-edge marking in different contexts and in different languages, see e.g. Michalsky (2017: Ch.2.1.1) and Haan (2002).

Benzmülller (2005) for German. Most of these languages also have qword questions in which the right edge is high or rising, although this is less commonly attested than low pitch. High pitch additionally seems more restricted in its occurrence than final low pitch, being pragmatically or socially conditioned by specific interactional contexts or speech style, or even phonologically conditioned by segmental structure (Bartels 1997; Frota 2002).

An explanation for the different amount of attention yes–no and qword questions have received might take into account the structural differences between the two types of questions. In contrast to yes–no questions, qword questions are by definition lexically and/or morphologically differentiated from statements. This follows from the use of an interrogative pronoun, which moreover often occurs in phrase-initial position, which leaves little uncertainty about the modality of the phrase.[10] There is therefore no real need for prosody in qword interrogatives to mark sentence modality the way it does in yes–no interrogatives. Accordingly, it is perhaps the *prominence* marking, rather than the *edge* marking within qword questions, which is more interesting to compare crosslinguistically. So far prominence marking of questions has received little mention in the literature, and I will therefore use the next section to provide an overview of crosslinguistic patterns of qword intonational prominence.

5.3.2 Question word marking

In Chapter 2 it was argued that postlexical prominence could be conceived of as the categorical presence or absence of a pitch accent (which by definition associates with a lexically stressed syllable). In the next few sections, I will describe the crosslinguistic prosodic marking of qwords as neutrally as possible, in terms of surface pitch prominence patterns rather than in terms of analytical concepts such as pitch accents. This also serves to facilitate the comparison between languages, as some of the languages that are discussed (Tamil, Korean, Mongolian) lack lexical stress and their pitch patterns are not analysed in terms of pitch accents in the original sources.

Section 5.3.2.1 deals with experimental studies, whereas Section 5.3.2.2 deals with the more descriptive studies of qword interrogative marking.

5.3.2.1 Experimental studies

The criteria for the inclusion of a study in this section were the following: The work needs to investigate qword prosody by means of an analysis that involves some quantitative element, and should be based on data from at least two speakers. Using these basic criteria, experimental studies investigating the intonation of qword interrogatives are nevertheless surprisingly rare. Collapsing across varieties of a single language, they cover six languages only: Spanish, Maltese, Greek, Dutch, Tamil and Korean. The order in which I will present the results from these studies is regionally informed, with

[10] Korean forms an exception, as interrogative pronouns in this language are homologous to indefinite pronouns, and interrogative pronouns in qword questions occur in the same position as indefinite pronouns in declaratives.

first the discussion of European languages, then Tamil and finally Korean. The studies involve varying numbers of speakers, ranging from 2 to 10. Most studies had a similar set-up, with speakers reading out scripted questions following a prompting context.

In the experiments on European languages, qwords occurred in phrase-initial position, with only Maltese also allowing for qwords in other positions. The qword in all these languages is typically marked with the highest pitch or the most extreme pitch excursion (rising towards a peak) in the phrase: This holds for Manchego Spanish and Peninsular Spanish (Henriksen 2014 and Prieto 2004, respectively); Maltese (Vella 2007; Grice, Vella & Bruggeman 2019); Greek (Arvaniti & Ladd 2009); and Dutch (Haan 2002). Among these, Dutch exhibited the most variable patterns: In 15% of qword interrogatives (28 of 185 cases), the qword did *not* receive the main pitch prominence in the phrase. This means that even for the least consistent language, the qword was in fact the most prominent in the vast majority of cases (85%). For all languages, the intonation contour typically dropped right after the qword-related maximum and lacks further prominence-marking pitch events. In some cases (the Dutch study and both studies on Spanish) contour variants are observed where the high pitch in the vicinity of the qword has the form of a plateau that extends beyond the right edge of the qword.

In Tamil, a comparable pattern for qword prominence is observed (Keane 2006a). Qwords in this language may occur in various phrasal positions. They are marked by rising pitch movement which is more pronounced than that on the corresponding word in a declarative. As most content morphemes in Tamil generally carry a rising contour, qwords do not necessarily get the highest phrasal peak, especially when they occur in non-initial position (presumably due in part to declination effects). However, qwords are enhanced compared to their non-qword constituent counterparts.

Korean, finally, does not quite conform to the same pattern of qword enhancement. Korean is analysed as lacking pitch accents and its intonation is instead analysed in terms of sequences of boundary tones (Jun 2005a, cf. Section 2.5.2). In Jun & Oh (1996), various sentence modalities are contrasted with each other, including qword interrogatives (qwords occur in-situ in Korean). Words that are semantically ambiguous between an interrogative pronoun (qword) and an indefinite pronoun were found to be phrased differently depending on the function they performed. Specifically, qwords formed an AP with the following verb, whereas indefinite pronouns were phrased as a single-word AP. This phrasing resulted in the consistent occurrence of a pitch peak around the second syllable of the qword, whereas the indefinite pronoun would get a final peak. It is not clear what this means for an interpretation in terms of qword prominence. Although Jun (2005a, 2014a) assumes that Korean lacks phonological, postlexical prominence, other authors have mentioned the percept of enhanced prominence of qwords in Korean, such as Choe (1995) (further references in Jun & Oh 1996). Even if Korean qwords are not readily interpreted as being prosodically prominent, at least in terms of surface pitch patterns they are not dissimilar to the aforementioned languages in which the main phrasal prominence co-occurs with the qword.

Something to keep in mind about the intonational patterns described in this section is that the findings of experimental studies might well differ from other types of studies including other speech styles. For example, for Dutch, qualitative results from a corpus study involving mostly spontaneous interactions (Chen 2012) yielded somewhat different results from Haan's (2002). In the corpus study, about three quarters of the 90 qwords produced in interrogatives were reported to receive some kind of accentuation (although not necessarily the most prominent phrasal accent). In Haan's (2002) work, almost all qwords carried the most prominent accent in the phrase. It seems likely therefore that speech style will have an effect not only on the type of right edge marking of interrogatives (as mentioned in Section 5.3.1) but also on the intonational prominence of the qword.

5.3.2.2 Descriptions and inventories

This section reports on studies that involved mainly qualitative descriptions of qword interrogatives. In a few cases that did have an element of quantitative analysis, it concerned aspects other than qwords (e.g. Sosa 2003). As a result, most of the descriptions of qword intonation mentioned in this section are supported by only one or two single examples, although there are some sources which are more elaborate (such as Vigário & Frota 2003 and Sosa 2003).

The three aforementioned volumes that contain collections of language-specific descriptions of intonation (i.e. Hirst & Di Cristo 1998; Jun 2005c, 2014b) together yield some two dozen language varieties about which some observations relating to qwords is made.[11] A few additional languages will be discussed here based on available resources outside these volumes.

In line with the observations in the previous section, most of the languages reviewed here have qword questions in which the qword gets the main intonational prominence in the phrase (sometimes called 'focal' prominence, which can, for present purposes, be interpreted as main prominence), and always involve some aspect of high pitch. The following languages conform to this common pattern: Four varieties of Latin American Spanish (Sosa 2003), Brazilian Portuguese (Antônio de Moraes 1998); Standard and Northern European Portuguese (Frota 2002; Vigário & Frota 2003); three varieties of Basque (Elordieta 1998; Elordieta & Hualde 2014); Catalan (Prieto 2014), Italian (Chapallaz 1972); Bulgarian (Misheva & Nikov 1998); Romanian (Dascalu-Jinga 1998); Georgian (Vicenik & Jun 2014); Russian (Svetozarova 1998); Mongolian (Karlsson 2014); Bengali (Khan 2014), Bininj Gun-Wok (Bishop & Fletcher 2005); Egyptian and Lebanese Arabic (Chahal & Hellmuth 2014); and Moroccan Arabic (Benkirane 1998).

Based on the examples and descriptions, the high pitch associated with the qword was variably realised as rising, high, or high-falling. Comparison between these descriptions is complicated by the fact that qwords are often short words occurring in phrase-initial position. This can obscure certain intonational movements like the presence of initial rises, which may or may not be realised depending on factors such as the number of

[11] For some of the languages in these volumes, qword questions are not mentioned at all (e.g. Italian and Swedish in Hirst & Di Cristo (1998)). Tamil and Spanish are also left out of this count as more detailed studies were discussed in the previous section.

syllables preceding the stressed syllable, or the number of voiced segments preceding the peak. With this proviso in mind, a small number of languages nevertheless seems to characterise qword prominence by means of a *rise* (with a peak and/or fall occurring after the qword): Egyptian and Lebanese Arabic, Georgian, Bengali and Mongolian. The remaining languages seem to have high pitch on the qword which is followed by a *fall* which is also (in part) realised on the qword.

The fact that there are apparent differences in the phonetic realisation of 'high F0 on the qword' does not invalidate the observation that qword intonation is highly similar crosslinguistically (at least based on the languages reviewed here). Firstly, alignment differences are not necessarily indicative of meaningful differences in the type of pitch event used to mark qwords as prominent. Alignment may differ predictably as a function of many factors, some of which reflect differences in phonological structure between qwords (such as length in syllables, segmental make-up, position of the stressed syllable), rather than meaningful differences in the nature of prosodic prominence marking. Moreover, even differences in the phonological labels used to describe pitch events in different languages do not necessarily entail fundamental differences, which is why I have refrained from referring to AM pitch accent categories as found in the original sources.[12]

Secondly, irrespective of the precise details of the high region, most languages exhibit similarities in the part of the intonation contour that follows the qword, which typically consists of a low-level flat stretch of F0 until the phrase end. In some cases this is explicitly described as deaccentuation (e.g. for the Basque varieties, and for Greek). In reviewing all the above resources, the phonetic resemblance between example F0 contours from different languages was in fact striking. Two examples of similar contours are given in Figure 5.1 for Northern Bizkaian Basque[13] and Hungarian. The schematic representation given in Figure 5.2 (from Haan 2002: 116) for a set of Dutch questions is also highly similar to these examples.

Interestingly, in some languages a very similar contour is also used for statements with (initial) focus. In those cases, the intonation contour is typically analysed as involving a focal pitch accent on the focused constituent, followed by deaccentuation and/or postfocal compression accounting for the flat F0 stretch (e.g. Elordieta & Hualde 2014 for Basque; Vanrell & Fernández Soriano 2013 for Spanish). These observations form further support for an interpretation of qwords in interrogatives as foci.

Finally, in contrast to the aforementioned languages that have prosodically prominent qwords, there are a number of languages that do not seem to associate the main phrasal prominence with the qword. These include Jamaican Creole (Gooden 2014), Chickasaw (Gordon 2005), Dalabon (Fletcher 2014), Japanese (Igarashi 2014) and Cur-

[12]Differences between languages in the choice of pitch accent are nevertheless suggestive of some general differences: The use of e.g. H*L for qwords in Northern Bizkaian Basque versus L*+H for Lebanese Arabic suggests that it is the fall from a high point which is crucial in the former and the rise to a high point which is crucial in the latter (even though both contours might share a global rising-falling movement)

[13]This variety of Basque contrasts words that have lexical pitch with words that lack lexical pitch accent. Words with lexical accent are identifiable by an acute accent on the relevant syllable.

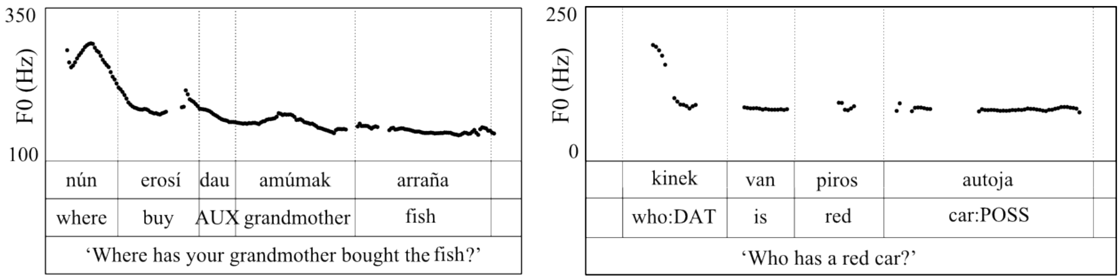

Figure 5.1: F0 contours for Basque and Hungarian qword questions, both with initial qword. Northern Bizkaian Basque reproduced from Elordieta & Hualde (2014: 435) based on original sound file, Hungarian recording author’s own.

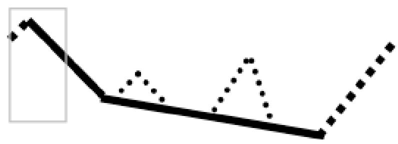

Figure 5.2: Schematised F0 contour for Dutch qword questions with initial qword. Reprinted from Haan (2002: 116). Solid line represents most of the contours in the sample (71%, N = 186) the dotted lines illustrate alternative realisations with further accents or optional final rise.

açao Papiamentu (Remijsen, Martis & Severing 2014). Interestingly, no experimental work could be found that explicitly investigates prosodic properties of qwords in English (or any other Germanic language, excepting Dutch).[14] Nevertheless, the general impression is that qwords in Germanic languages (again, excepting Dutch) are not marked with the main phrasal prominence.

5.3.2.3 Summary and discussion

The above overview suggests that there two main crosslinguistic patterns in prosodic prominence marking of qwords (cf. Ladd 2008: 227 for a division with a slightly different focus):

- Languages that typically mark the qword with the main phrasal pitch prominence: European and Brazilian Portuguese, many if not all varieties of Spanish and Basque, Catalan, Dutch, Italian, Greek, Maltese, Egyptian Arabic, Lebanese Arabic, Moroccan Arabic, Romanian, Bulgarian, Hungarian, Russian, Georgian, Bengali, Tamil, Bininj Gun-Wok;

[14]There are many resources that deal with yes–no question intonation, including Michalsky (2017) for German, but these tend to make little mention of qword questions.

- Languages that either do not typically mark the qword with the main phrasal pitch prominence or variably do so: English, German, Jamaican Creole, Chickasaw, Dalabon, Japanese, Curaçao Papiamentu.

The overall pattern, even if based on limited evidence only, is clear. Some Germanic languages, including English and an English-based creole, and a handful of other languages form the exception in matters of qword interrogative intonation. Most other languages require the main phrasal pitch prominence in the form of some sort of high pitch to coincide with the qword.

As stated before, I have avoided to frame descriptions in terms of pitch accentuation. Consequently, qword pitch prominence patterns could be observed to look very similar across a range of languages. This theory-neutral way of describing qword intonation is less helpful when the aim is to analyse the prosodic properties as either serving prominence marking purposes or edge marking. For this latter purpose, the role of stressed syllables and phrasal edges in determining the location of the pitch movement should prove insightful, and steers right into the domain of intonational phonology.

In AM approaches to intonation, pitch events that co-occur with stressed syllables are considered to be prominence marking, whereas pitch events that occur at the edges of phonological domains are considered edge marking. In the case of qwords, determining whether its pitch properties are prominence marking or edge marking is somewhat complicated by the fact that most languages require qwords to appear at the left edge of some prosodic phrase. This makes it potentially difficult to distinguish phonetically between pitch movement that is to be interpreted as edge marking and pitch movement that serves to mark the prominence of the qword (see also Section 2.4.3.2). Helpfully, most of the previously discussed studies mention that there is a role for stressed syllables in determining the location of specific F0 turning points, and accordingly analyse the pitch events as pitch accents rather than boundary tones. Additionally, in languages where qwords can occur in initial as well as non-initial positions, the high pitch co-occurs with the qword irrespective of its position. This is the case for Romanian and Maltese (and also for Tashlhiyt Berber and Moroccan Arabic, as will be shown in the next chapters).

In conclusion, phonetically and phonologically, it seems that the pitch properties of qwords in many languages reflect intonational *prominence* marking, rather than *edge* marking. This interpretation is moreover supported by the qword's semantic-pragmatic salience as discussed in previous sections.

5.4 Further points of interest

This chapter has highlighted that the intonational properties of qword questions are highly similar across a range of languages, with the main insight that many languages associate the main phrasal prominence-marking event with the qword.

No experimental work has so far been conducted on qword interrogatives in Tashlhiyt Berber, and little is known about qword interrogatives in Moroccan Arabic beyond mere observation. In light of the previous discussion, a number of concrete questions can be asked with respect to qword interrogative marking in both languages:

1. What are the phrasal prosodic characteristics of question word interrogatives in general and the prosodic properties of qwords in particular?

2. What is the prosodic prominence status of question words? How should any prosodic marking be analysed phonologically?

The following two chapters (6 and 7) will address these questions for Tashlhiyt Berber and Moroccan Arabic, respectively.

There is a third question, which will not be addressed directly, but is relevant nevertheless: To what extent do Tashlhiyt Berber and Moroccan Arabic exhibit similarities in question word interrogative marking? This question bears on the language contact situation that characterises these languages (see Section 1.3). At this point, very little is known about qword interrogatives in other varieties of Berber, making it difficult to determine how any characteristics of Tashlhiyt compare to those of other varieties. More information is available about qword interrogatives in other varieties of Arabic (e.g. Chahal & Hellmuth 2014), but, to my knowledge, there are no controlled experiments on the topic. A very promising possibility for future work in this direction lies in the IVAr corpus (Hellmuth & Almbark 2017), which contains comparable data from seven varieties of Arabic, including recordings of the same experiment that yielded the MA qword interrogatives that will be discussed in Chapter 7. Any similarities between TB and MA qword marking could be compared against qword marking in other Arabic varieties in order to find out where the greater similarities lie.

6 Question word interrogative intonation in Tashlhiyt Berber

6.1 Introduction

6.1.1 Prior work on the intonation of Tashlhiyt Berber

Prior to 2011, no work had ever directly addressed Tashlhiyt Berber intonation. Passing remarks had been made in Stumme (1899) and Dell & Elmedlaoui (2002), and a handful of examples were given in Lafkioui (2010). Work on the intonation of Tashlhiyt since then has focused on the tonal events occurring near the right edge of Intonation Phrases (IPs) (cf. proceedings papers Grice et al. 2011; Roettger, Ridouane & Grice 2012), articles Grice, Ridouane & Roettger 2015; Roettger & Grice 2015, and the PhD thesis Roettger 2017, which is based on some of the earlier publications).[1]

A finding shared by all of these past studies is that Tashlhiyt exhibits a great deal of variability in tonal placement in IP-final position. Grice, Ridouane & Roettger (2015) investigate two types of sentences produced in an experimental setting: Yes-no questions (marked with a question particle), and statements with a contrastively focused word in IP-final position. Both are characterised by a rising-falling F0 movement close to the right edge of the phrase and both these movements are analysed phonologically as involving an H tone. Despite their presumed different functions as edge-marking (in the case of the yes-no questions) and prominence-marking (in the case of the declaratives with a contrastively focused word), H tones in both phrase types exhibit similar behaviour in the sense that they can both dock onto either the penultimate or the final syllable of the final word in the phrase. These patterns of variable realisation are moreover very similar to the realisational variants Roettger & Grice (2015) report for declarative questions (yes-no questions not marked by a question particle).

Grice, Ridouane & Roettger (2015) analyse the H tones at the right edge of questions and contrastive statements as having primary association with the phrase edge, and secondary association to a specific syllable. This explanation fits in well with the assumption that the language does not have lexical stress, since this type of association presupposes that there are no predetermined lexical anchors in the form of stressed syllables to which postlexical tones would associate. This is not to say that tonal association was entirely unconstrained: Instead of structural prominence factors, syllable weight and nucleus sonority as other lexical–phonological metrical factors, were invoked to explain the distribution of the H tones.

[1] The present chapter has been published in a slightly different format as Bruggeman, Roettger & Grice (2017).

The present chapter looks at the intonational events that co-occur with question words (qwords) and will provide a detailed investigation of tonal alignment and association in phrase-initial position.

6.1.2 Question word interrogative structure in Tashlhiyt Berber

In Tashlhiyt Berber, most morphologically simple qwords consist of one or two syllables starting with 'ma', e.g. *ma/mad* 'what' and *mani* 'where'. Complex question constituents (henceforth, complex qwords) consist of *man* 'which' followed by a noun, e.g. *man aʜuli* 'which sheep'). The canonical position for qwords in Tashlhiyt is phrase-initial, as it is in Berber in general (Stumme 1899; Kossmann 2012). Many authors have drawn parallels between initial qwords and cleft constructions, with clefts in Berber being defined as phrase-initial focal constructions (Van den Boogert 1997; Stoyanova 2004; Kossmann 2012). Specifically, clefted constituents are followed by the marker *a(d)* 'which, that', from which the /a/ following /m/ in most qwords might stem (Kossmann 2012). While it is implicit in most of the aforementioned work that qwords are not licensed in any other phrasal position than the phrase-initial one also reserved for clefts, this view is made explicit in Stoyanova (2004) for Tamazight and Tarifit Berber. However, a quick glance at qword questions in a corpus of semi-spontaneous speech in Tashlhiyt (Bruggeman & Roettger 2017) reveals that, at least in this type of speech, question words also frequently occur in phrase-final position (see also Section 5.2.1).

6.1.3 Research questions

In Section 5.4 I listed a few questions that are of general interest when looking at qword interrogative intonation crosslinguistically. They are repeated here and will be addressed for Tashlhiyt Berber in this chapter:

1. What are the phrasal prosodic characteristics of question word interrogatives in general and the prosodic properties of qwords in particular?

2. What is the prosodic prominence status of question words? How should any prosodic marking be analysed phonologically?

Because very little is known about qword interrogative intonation in Berber in general, and nothing about Tashlhiyt, a subquestion to the first main question is whether and in what contexts qwords in Tashlhiyt are characterised by prosodic prominence marking. Based on the discussion in Chapter 5, it is likely that qwords are marked by an intonational prominence-marking event in direct interrogatives, as they are in many other languages. The follow-up question would then concern how this postlexical prominence should be interpreted phonologically, and how it stands in relation to the prosodic characteristics of the rest of the interrogative (i.e. Question 1). In section 6.2 I will report on a pilot experiment which compares the qword *managu* 'when' in direct interrogatives, i.e. speech acts expressing a request for information, and in embedding contexts. It is likely that the qword is intonationally prominent by means of a pitch event (which likely conveys focus) in a direct interrogative (see Chapter 5.2.2.2). At the same time, the expectation is that the same intonational marking is not found for embedded qwords.

The second question pertains to how the intonational characteristics of qwords in Tashlhiyt should be interpreted phonologically. After the pilot in section 6.2 has clarified that qwords indeed are marked consistently with a pitch event, the main experiment from section 6.3 onwards will serve to answer this second research question with specific reference to qwords in phrase-initial position. This main experiment investigates the properties of 11 different qwords (five simple and six complex), zooming in on the alignment and scaling characteristics of the F0 event in question.

6.1.4 Data

All speech materials were developed in consultation with native speaker informants. The data were recorded by the author in a quiet room in the Département des études amazighes at the Université Ibn Zohr in Agadir, using a PreSonus Audiobox solid-state recorder at a sampling rate of 44.1 kHz, and a head-mounted AKG C420 III microphone. Pilot recordings took place in November 2014, recordings of the main experiment in March 2015.

6.1.5 Overview of chapter

In Section 6.2, the pilot experiment and results are discussed. After this, in Section 6.3, the methodology of the main experiment is presented, followed by the results (6.4), with simple qwords being discussed in the first few sections and complex qwords afterwards (6.4.5). A discussion follows in 6.5, with the research questions being answered in Section 6.5.2. Finally, the last part of this chapter (6.6) gives a brief summary and conclusion.

6.2 Pilot

6.2.1 Participants

The recordings for both experiments involved the same participants, with the exception of one participant who completed the pilot but not the main experiment. Data from two participants were excluded from analysis in both pilot and main experiment, which in one case was due to reading difficulties and in the other case due to the speaker being unable to finish recording. In the end, the pilot results are based on data from the seven participants that did in both experiments.

Table 6.1 gives more detailed information about the main seven participants (six female, indicated by ‘f’, and one male, indicated by ‘m’) who did both pilot and main experiment. Participants’ place of birth varied, as shown in Figure 6.1. The three participants whose place of birth is marked by an asterisk moved to Agadir sometime

Participant	Age	Born/raised	Repetitions
1f	21	*Tata	4
2m	23	Essaouira	2
3f	22	Ait Baamran	4
4f	22	*Ida-ou-Tanane	4
6f	21	*Essaouira	4
8f	26	Taroudant	2
9f	24	Sidi Ifni	4

Table 6.1: Participant details.

during their youth, the rest had come to Agadir for the purpose of their university education.

All participants were native and dominant speakers of Tashlhiyt Berber, students at the Département des études amazighes at the Université Ibn Zohr in Agadir at the time of recording, and spoke Tashlhiyt regularly both with friends and family. Participants were multilingual: They were all fluent in Moroccan Arabic, and had varying fluency in Modern Standard Arabic, French and English as languages learnt in school (this was expected, see also Section 1.2).

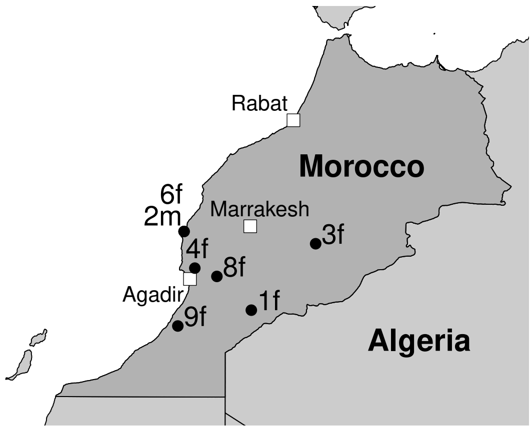

Figure 6.1: Location of participants' place of birth.

6.2.2 Speech materials and methodology

This pilot served to find out whether qwords are marked differentially in different contexts. It compared the qword *managu* 'when' in direct questions, i.e. speech acts expressing a request for information, with cases in which it occurs embedded (in questions and statements).[2] The basic expectation is that the qword is marked by intonational prominence in a short direct interrogative, whereas it is probably not when it is embedded.

In order to elicit natural sounding instances of direct questions as well as embedding contexts, a scripted mock telephone dialogue between two imaginary speakers was used. The full script with gloss and translation is given in Appendix A. Participants were familiarised with the content of the conversation beforehand and were then instructed to read and act out both sides. The Tashlhiyt text was presented in Latin script a few lines at a time on a screen. Pictures and speech bubbles were used to represent the turns of the dialogue participants.

6.2.3 Results and discussion

As mentioned previously, the default position for qwords in Tashlhiyt is phrase-initial. A typical intonation contour of an interrogative with an initial qword is given in Figure 6.2a. Qwords may occur in final position as in 6.2c, in a context where the question is repeated. There was no apparent effect of position: Qwords in interrogatives were characterised by the presence of a peak, with the rising and falling movements surrounding the peak for the most part also being realised on the qword.

In contrast to qwords in interrogatives, embedded qwords that do not contribute to interrogative meaning are not marked by a similar pitch event, as can be seen for medial position in 6.2b and final position in 6.2d.[3]

On the other hand, qwords that are in medial (peninitial) position, preceded by the discourse marker *im:a* 'so', were characterised by the same rising-falling pitch movement as initial and final qwords. Figure 6.3 shows two intonation patterns that are possible on the phrase *im:a managu rad tbdut lxdmt* 'so when will you start work?'. The preferred way to produce these peninitial qwords was with the rise starting prior to the qword, on the discourse marker, as in 6.3a. The pattern in 6.3b, a low level stretch of F0 contour before a rise that is initiated at the left edge of the qword, was produced by one speaker only.

The patterns of qword intonational marking just described were consistently produced by all eight speakers in the context of this dialogue task. It therefore seems that a localised rise-fall is the defining characteristic of qwords in interrogatives, since it occurs in different phrasal positions and across different discourse contexts (target sentences occurred at different parts in the dialogue). Qwords that do not contribute to interrogative meaning are not marked by the same intonational event.

[2] Dell & Elmedlaoui (2002: 185f.) note that the pronunciation of the final segment in *managu* is subject to free variation and may be produced either with a full high back vowel or with a labialised velar stop, i.e. as [managu] or [managʷ]. In informal elicitation with all our participants the word was consistently produced with a clearly identifiable final vowel, which motivated the transcription as 'managu'.

[3] The scripted dialogue also included a test stimulus with an initial embedded qword: *managu a ra n: aʃkʁ, s:nʁ is rad t t:init* 'when (that) I come home, I know that you want to say it'. However, given that most participants struggled to understand this construction upon presentation, this item was excluded from further analysis.

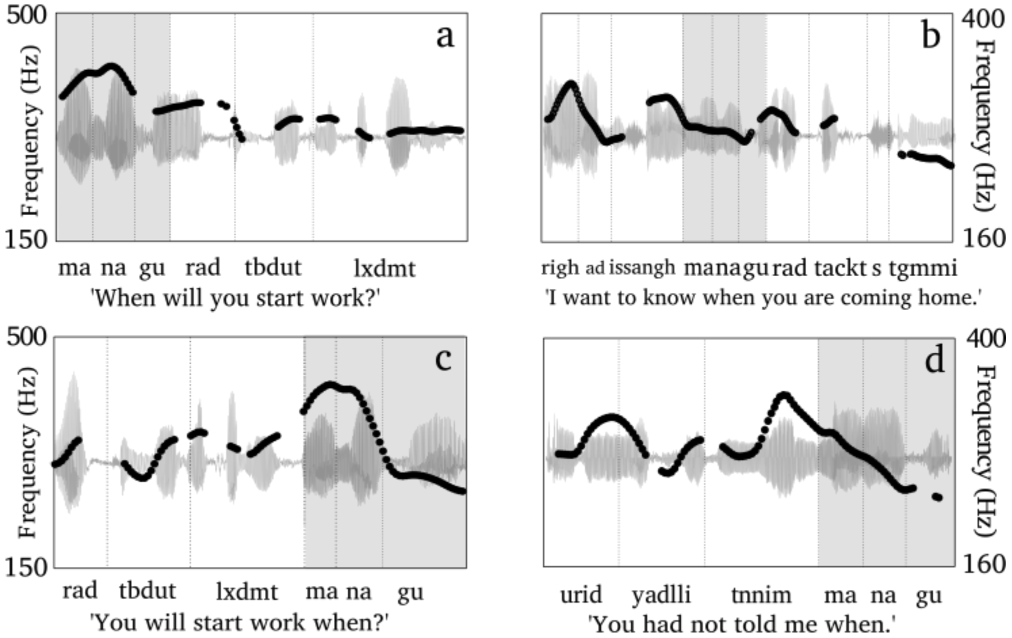

Figure 6.2: Representative F0 contours and waveforms for interrogative qwords (a: initial and c: final) and embedded qwords (b: medial and d: final); target word *managu* 'when' highlighted in grey.

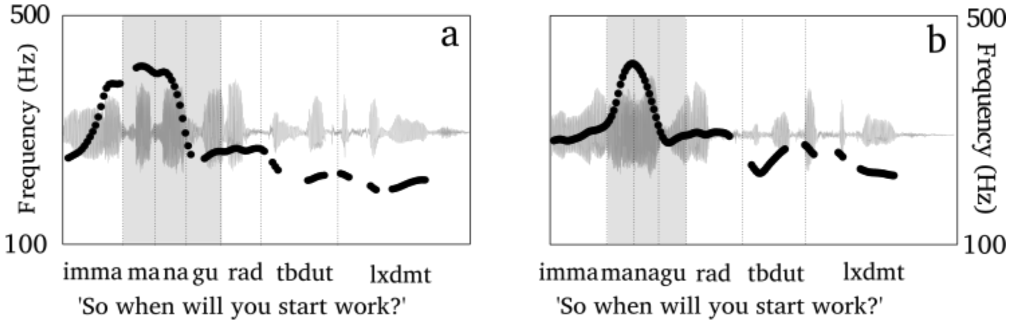

Figure 6.3: Two F0 contours and waveforms for questions with peninitial qword, with target word *managu* 'when' highlighted in grey.

An important follow-up question to these findings concerns the specific realisational detail and phonological analysis of the relevant tonal movement, including whether it should be considered prominence-marking or edge-marking. Turning point candidates for further exploration include the sequence of low, high and low turning points that are all seen to occur on or in the vicinity of the qword. Eyeballing the present data suggests that the first (low) turning point that initiates the rise might be less important than the following two. Firstly, in peninitial position, most speakers' rising F0 movement starts prior to the start of the qword, as exemplified in Figure 6.3a. This suggests that the initial low turning point in the sequence might be linked to an intonational event tied to the left phrasal boundary and is not a target that needs to be realised on the qword itself. Secondly, in final position, as in Figure 6.2c, the rising movement may be realised largely prior to the qword. Following these observations of variability, the initial low turning point appears to be less inherently part of the qword tune than the other two that exhibit less variability in alignment. I will return again to the role of phonetic alignment in determining phonological representation in the discussion of the main experiment in Section 6.5.

The aim is to find out, by means of the main experiment reported in the following sections, what the exact realisational details are of this rising-falling F0 pattern on the qword. In the next section, the methodology of the main experiment will be described.

6.3 Methodology

6.3.1 Participants

Participants in the main experiment were the same as in the pilot, minus one speaker. This resulted in data being analysed from a total of seven speakers. For details see Section 6.2.1.

6.3.2 Procedure

Participants were instructed to act out the role of a primary school teacher doing an exercise with their pupils that involved asking questions about pictures. The experiment was presented to participants as a powerpoint presentation in slide presentation mode with each slide showing a picture and a brief textual description of the scene depicted. The target question asking about the picture was printed at the bottom of the slide underneath the picture. Participants were instructed to read the picture description out loud and then produce the question as if they were asking it to their pupils.

6.3.3 Speech materials

While there are independent reasons to believe that the question word acts as the default focus of the question (see Section 5.2.2.2), questions were also framed in a context in which the question word constituted the phrase's single focus. Short questions with simple word order were used, and lexical items other than the qword were both textually and visually given. In addition, in the context of the task, with subsequent questions having different phrase-initial qwords, qwords were subject to implicit contrast with the qword from the previous qword interrogative. Thus, the following factors ensured that stimuli were treated as constituting a single 'focus domain' (cf. Ladd 2008) with narrow focus on the qword: i) the simple sentence structure, ii) the short length of the phrase, and iii) the textual and visual givenness of all elements other than the qword.

An example context is given below for the target qword *mani*. The first line represents the Latin orthography used in the presentation to the participants, and the second line a phonemic transcription.

Context sentence 1

afrux	ann	iZRa	yan	aHuli	gh	ugharas.
afruχ	*an:*	*izˁra*	*jan*	*aʜuli*	*ʁ*	*uʁaras.*
boy	DIST	3SG.M-see.AOR	one	sheep	in	road

‘The boy there sees a sheep on the road.’

Context sentence 2

tsaqsat	imHDarn	nnm	:
tsaqsat	*imʜdarn*	*n:m*	:
2SG.F-ask.AOR	students	POSS-2SG.F	

‘You ask your students :’

Target question

mani	gh	iZRa	aHuli?
mani	*ʁ*	*izˁra*	*aʜuli?*
where	in	3SG.M-see.AOR	sheep

‘Where does he see the sheep?’

The full set of 11 qwords (five simple and six complex) with their carrier sentences are given in Table 6.2. Target questions had either qword-Verb-Object or qword-Verb-Adverb structure and the number of syllables following the qword was always five. Simple qwords varied in number of syllables from one to three, and included CV and CVC syllables. Complex qwords consisted of the interrogative element *man* ‘which’ followed by either a disyllabic or a trisyllabic noun. Syllabification in Tashlhiyt has been the subject of much previous work and is especially complex in the case of long consonantal sequences (Dell & Elmedlaoui 2002, 2008; Ridouane 2008). In the case of the simple qwords, which all had vocalic nuclei, the syllable boundary location is uncontroversial. It is unclear if resyllabification takes place across the elements of a complex qword constituent, which is why no syllabification is given for those qwords.

Picture slides were presented in blocks of 11 within which each target qword interrogative occurred once. As the recording session involved a number of other tasks, speakers completed a set of two blocks with the stimuli in each block having a different semi-randomised order, followed by another task, and finally the same set of two blocks again. Each set was preceded by five practice items. No fillers were included to minimise the total duration of the session. Of the seven speakers, two completed only one block, so that their number of repetitions per stimulus is two instead of four, as shown in Table 6.1. This led to a total number of 24 repetitions per target word, resulting in a total of 120 tokens of the five simple qwords, and 144 for the complex qwords. After the exclusion of disfluent items (misreading, hesitation) 107 simple (89%) and 120 complex qwords (83%) remained.[4]

[4]The relatively high number of exclusions is probably an artefact of speakers not being used to reading

Simple		Complex	
ma *ma iʃ:ta ʁ umalu?* 'what does he eat in the shade?'	'what'	*man anu* *man anu nzˤra ʁ umalu?* 'which well do we see in the shade?'	'which well'
mad *mad nzˤra ʁ umalu?* 'what do we see in the shade?'	'what'	*man ananas* *man ananas nzˤra ʁ umalu?* 'which pineapple do we see in the shade?'	'which pineapple'
mani /ma.ni/ *mani ʁ izˤra aʜuli?* 'where does he see the sheep?'	'where'	*man aʜuli* *man aʜuli nzˤra ʁ umalu?* 'which sheep do we see in the shade?'	'which sheep'
manwi /man.wi/ *manwi nzˤra ʁ umalu?* 'who do we see in the shade?'	'who'	*man tizi* *man tizi nzˤra aʜuli?* 'what time do we see the sheep?'	'what time'
managu /ma.na.gu/, /ma.nag^w/ *managu nzˤra aʜuli?* 'when do we see the sheep?'	'when'	*man tili* *man tili nzˤra ʁ umalu?* 'which ewe do we see in the shade?'	'which ewe'
		man butili *man butili nzˤra ʁ umalu?* 'which shepherd do we see in the shade?'	'which shepherd'

Table 6.2: Target question words and their carrier sentences.

6.3.4 Analysis

Scaling and alignment measurements were taken from the F0 contour provided by the standard pitch tracking algorithm in Praat (Boersma & Weenink 2015), with manual correction of spurious pitch points and tracking errors such as octave jumps.[5]

Tashlhiyt aloud. The remaining productions, however, were judged to be natural sounding utterances expressing the intended communicative function by two additional Tashlhiyt speakers who did not participate in the experiments.

[5]Utterances were characterised by considerable microprosodic effects at the transition between vowels and nasals. In an attempt to control for these perturbations in the analysis of alignment, measurements were taken both from the raw and a smoothed version of the F0 contour. For this smoothed contour the raw contour (already handcorrected) was manipulated in four additional steps with the customised Praat script *mausmooth* (Cangemi 2015), which included two rounds of smoothing. Results from raw and smoothed contour versions were highly similar and I will only report the results from the raw contour here.

Two measurements were used in the quantification of the properties of the maximum occurring on the qword: A single absolute F0 maximum, and a measure of a 'high region'. It has repeatedly been shown that small and gradual F0 displacement within a dynamic contour does not lead speakers to perceive pitch differences ('t Hart 1976; d'Alessandro & Mertens 1995). Given the possibility, then, of a region in the contour within which pitch is perceptually equivalent, it is also conceivable that a larger region is involved in systematic tune-text association.

In order to quantify the high region, a heuristic measure inspired by earlier work on pitch plateaux was used. Different measurements strategies for plateaux have been used in the past, and some will be quickly reviewed here. Knight & Nolan (2006) and Knight (2008) for example use the start and end of a region delimited by 4% of F0 values in Hertz below the absolute maximum when measuring high plateau accents in British English. A slightly different measure is adopted by Niebuhr & Hoekstra (2015), who use a 1 ST difference criterion around the peak in their discussion of North Frisian plateaux. Both of these sets of authors refer back to more general pitch perception findings by 't Hart (1981), who had to draw the somewhat dissatisfactory conclusion that the perception of pitch movement was highly speaker-specific: For a reliable discrimination between pairs of pitch movements different speakers needed differences that ranged in size from 1 to 6 semitones. Given this variability in perceptual discrimination ability among speakers, the exact definition of what constitutes a perceptually relevant plateau seems somewhat arbitrary. Knight & Nolan (2006) argue that there is little difference between plateaux delimited by 4% and 6% Hz values around the maximum. This motivated the choice to use 6% difference values as plateau measures in the present study, especially because in most of the present speakers' ranges, 6% in Hertz values was very similar to the 1 ST criterion used by Niebuhr & Hoekstra (2015). Figure 6.4 schematically depicts the adopted plateau measurements (start and end point).

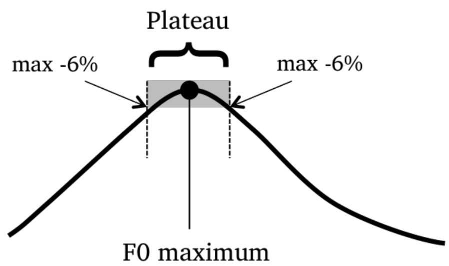

Figure 6.4: Schematic representation of peak measure (absolute F0 maximum) and plateau measures (6% lower values in Hz around maximum).

6.4 Results

6.4.1 Global contours

Figure 6.5 below shows time-normalised phrasal contours for all individual qword interrogatives with a simple qword.

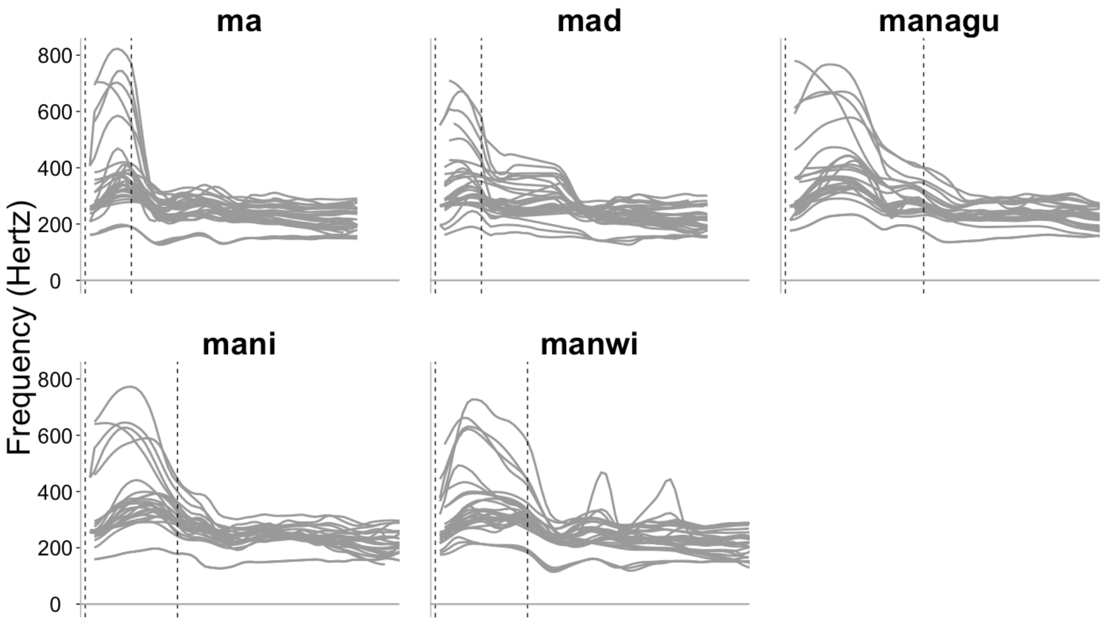

Figure 6.5: Smoothed F0 contours for all target utterances with simple qwords (N = 107). Normalised duration based on ten equidistant measuring points throughout each syllable. Dotted lines delimit target qword.

It can be seen that qwords are characterised by a region of high pitch, with a rise towards a peak on the qword and a fall back to what seems to be a low baseline immediately following the peak, around the right edge of the qword. Only in interrogatives with *managu* the fall is achieved somewhat earlier, before the right of the qword. Interrogatives with *mad*, finally, diverge only marginally from the general pattern in having an additional, further fall, a few syllables to the right of the qword. Taken together, these interrogatives provide a highly coherent picture of a rising-falling F0 movement on the qword and a typical absence of additional intonational prominence anywhere in the rest of the phrase.

6.4.2 Turning point scaling and alignment

This section presents results of peak alignment in the five simple qwords *ma*, *mad*, *mani*, *manwi* and *managu*. Figure 6.6 shows the alignment of the absolute F0 maximum per word (normalised duration), with word boundaries indicated by solid lines.[6] Each dot represents an utterance, i.e. one specific peak as realised on that qword.[7]

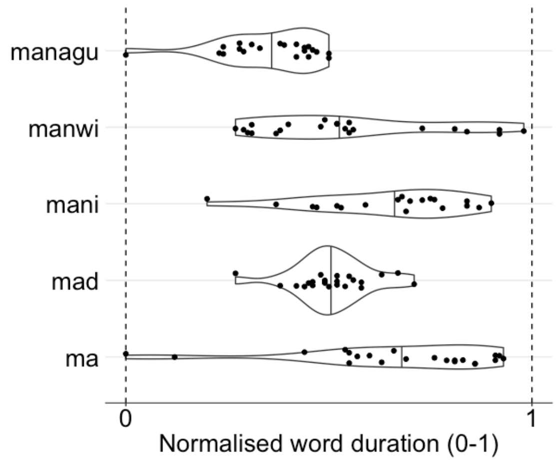

Figure 6.6: Alignment of absolute F0 maximum relative to normalised word duration for the five simple qwords *ma, mad, mani, manwi* and *managu*.

While maxima are not commonly realised in the very first part of the word, the temporal domain across which they are realised is surprisingly large, and spans the largest part of the word (excepting *managu*, which exhibits an absence of peaks in the second half of the word).

In order to investigate peak alignment in more detail, the location of F0 maxima was considered in relation to individual segments. Figure 6.7 illustrates this for all five qwords and for each speaker separately. The differing number of tokens per speaker is a function of the speaker's original number of repetitions, which was either two or four (see Table 6.1), and subsequent exclusions.

As expected, the distribution of peaks over a large part of the word translates into maxima that variably occur on different segments. In absolute terms, monosyllabic qwords *ma* and *mad* 'what' exhibit the greatest degree of uniformity. For these words, all speakers tend to produce maxima that occur on the vowel /a/ across multiple repetitions. In the polysyllabic words, maxima are spread across different segments as well as different syllables. Peaks are observed on any of the segments from the second to the last segment in *mani* and *manwi*, and from the first to the fourth in *managu*. The apparent restriction on the occurrence of F0 maxima at the start of the second syllable (i.e. around /n/ in *managu* and *mani*, and /w/ in *manwi*) can be explained in terms of microprosody: There is a predictable dip in the pitch contour on those segments. It is

[6]The realisation of *mad* often underwent assimilation to the following consonant, whereby the target sequence *mad nzˤra* became [man(:)zˤra]. As it was impossible to distinguish between individual /d/ and /n/ segments, the whole sequence [man(:)] is taken to reflect orthographic *mad*.

[7]Four peaks in the case of *manwi* were realised just milliseconds after the qword end, but this was in cases where the absolute maximum was difficult to identify as a single point that was higher than the rest. In the following I will simply assume that peaks align systematically on the qword.

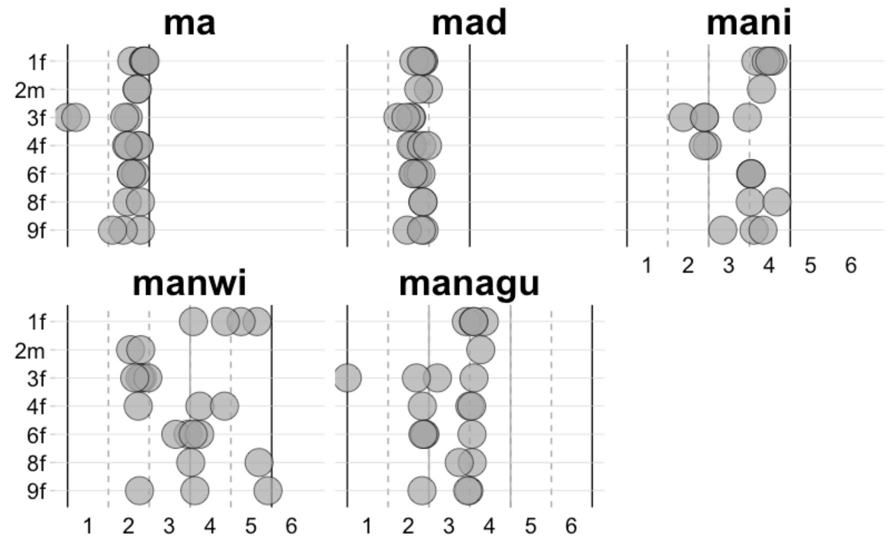

Figure 6.7: Alignment of absolute maximum with respect to segments (numbered and separated by dashed lines) for all five simple qwords.

nevertheless clear from the above that there is no stable segmental anchor for the peak.

Peak distribution overall can moreover be characterised as exhibiting a gradient spread rather than a categorical distribution, with most speakers producing a multitude of alignment patterns. The attested variability can be classified along a number of parameters:

- peaks that align with different syllables (e.g. 3f's peaks in *mani*, and 4f's peaks in *managu*)
- peaks that align with different segments within the same syllable (e.g. 1f's and 9f's peaks in *manwi* and *mani*, respectively)
- peaks that align with different syllables *and* with different segments within one of these syllables (e.g. 3f's peaks in *managu*, 9f's peaks in *manwi*)

The alignment patterns across words also display similarities, especially between *managu* and *mani*, both of which exhibit a general preference for maxima on the second vowel. In both words, the earliest peaks occur around the segment boundary between /m/ and /a/, and the latest peaks halfway through the second vowel. It appears, therefore, that *managu* behaves like a disyllabic word in which the final syllable *gu* does not count. A plausible explanation is that the word, despite being produced with a clear final vowel, is treated as if it does not have a final phonological vowel that participates in tune-text association, i.e. as *manag*^{w} (see also 6.2.2).

Compared to *mani* and *managu*, *manwi* exhibits slightly more categorical peak alignment, reflected also in the greatest within-speaker variability (e.g. speakers 4f and 9f producing discretely different peak alignment, i.e. on different syllables, across repetitions). Still, single speakers may produce any combination of 'categorical' peaks and other more gradiently different peaks (excepting 3f). This suggests that *manwi*, too, is characterised by a gradient distribution of peaks, which is simply obscured by the large microprosodic perturbation of the labiovelar approximant in the middle. Additionally, a comparison between late peaks in *manwi* and the right edge peaks in the complex qwords (see Section 6.4.5) shows that the late peaks on *manwi* are not aligned as late as the peaks marking the right edge of a complex qword. This suggests that the *manwi*-peaks are qualitatively more like the peaks on the other simple qwords than like the possible edge-marking strategy seen in complex qwords.

In sum, then, the alignment results presented so far indicate that there is little systematicity both within and across speakers in alignment of F0 maxima in TB qwords. While peaks may occur on most segments in the word, and variably align within these segments, the main consistent feature of all peaks is that they occur *within* the domain delimited by the boundaries of the qword.

This degree of variation in alignment is unlike that in languages for which segmental anchoring has been invoked as an explanation for the relatively stable alignment of (specific) tonal targets in relation to the segmental string. To name one example of consistent alignment results, Atterer & Ladd (2004) found that in rising L* + H accents in two varieties of German, the low turning point characterising this accent was significantly later aligned in one variety compared to the other. Crucially, cast in absolute terms, this alignment difference roughly spanned half a segment. Speakers of each language variety behaved so uniformly that the resulting difference, realised on a very small temporal scale, functioned as a significant predictor of the variety spoken. In relation to these findings, the patterns discussed here are of an altogether different nature. While it might be objected that the present data come from speakers with perhaps a less uniform background than Atterer & Ladd's (although this is debatable given that their dialect regions were rather broadly defined as 'Northern' and 'Southern'). Two further points may be raised in defence of considering the present study's speakers together as representing a single variety. Firstly, alignment patterns from the two speakers with the same birthplace (2m and 6f, both of whose parents are also from that same town), are not more similar than those of any two other speakers. In a similar vein, the three speakers who grew up in Agadir (1f, 4f, 6f) do not behave more alike than any other grouping of speakers. Secondly, even if variability in alignment between speakers could be explained by attributing differences to specific subdialects, we would still not expect to observe the degree of intraspeaker variability exhibited in the present data if alignment patterns were characterised by segmental anchoring. It can be concluded that the alignment of the qword-related F0 maximum exhibits genuinely variable alignment, both within and across speakers.

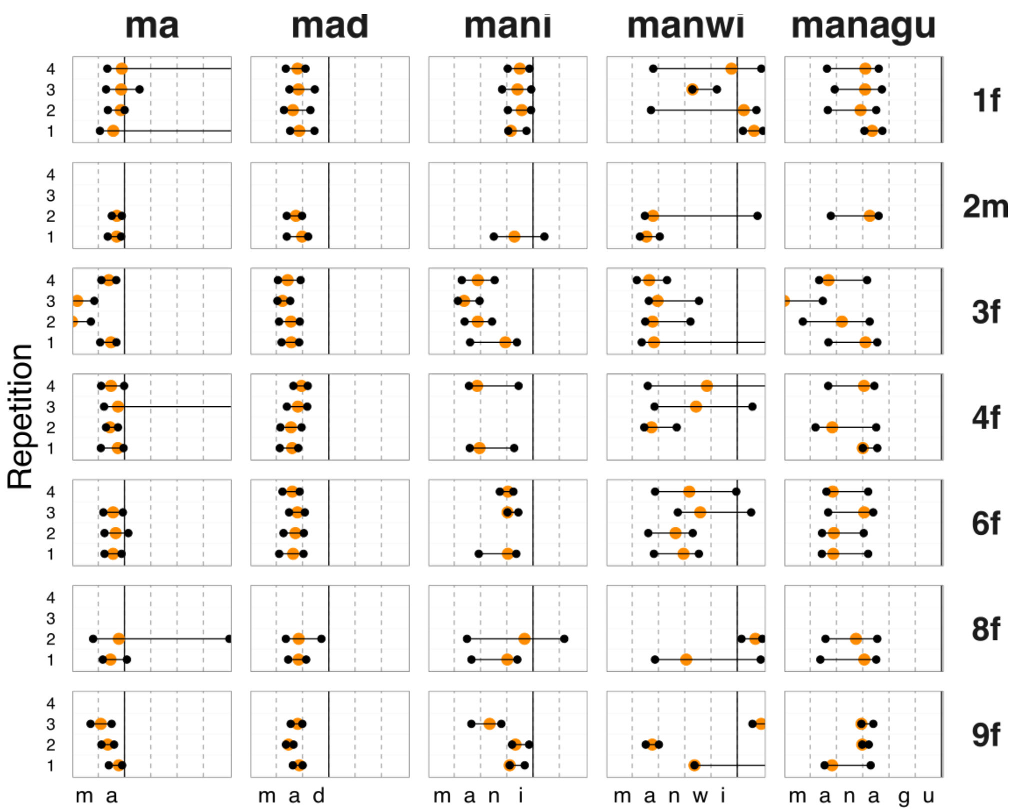

Figure 6.8: Plateau alignment for all individual repetitions of simple qwords by qword and by speaker. Plateau onset and offset indicated by black dots, location of absolute maximum within plateau indicated by orange dot, segments (duration normalised) separated by dashed lines and qword boundaries indicated by solid lines.

6.4.3 High plateau alignment

The variability seen in the alignment of individual maxima is reflected in the results from the plateau measures. Figure 6.8 depicts, for the simple qwords, the plateau onset and offset (black dots linked by line) and the location of the absolute maximum within it (orange dot). A pattern observed across all qwords is one whereby the plateau starts in the second segment (the vowel /a/), and extends across a number of segments within the qword, but usually does not cross the right qword boundary.

Within this general pattern, plateaux nevertheless exhibit considerable within-qword variability in alignment of both onset and offset. Additionally, some of the qwords seem to exhibit idiosyncratic alignment patterns. *Mad* exhibits what are arguably the most consistent alignment patterns across repetitions, with the end of the plateau reached well before the end of the word-final segment /d/ (realised as [n], see above).

Ma differs from the other qwords in that it may be marked by a high plateau that extends far to the right of the qword boundary (in repetitions by speakers 1f, 4f and

8f), matching the observations about an additional fall (or later fall) in this particular interrogative phrase (Figure 6.5). Further research will have to show what exactly the location of this fall reflects, and specifically whether this apparent categorical difference between early and late falls is meaningful. Other speakers produce plateaux on *ma* that are more similar to the narrower plateaux characteristic of other qwords. Finally, the qword *manwi* is characterised by another kind of variability in the realisation of its high region, with both very wide plateaux and very narrow plateaux.

Considering all qword plateaux together, every single plateau parameter investigated here seems to be characterised by gradient variation (with the exception of *ma* with two categorically different patterns): The plateau onset, the plateau offset, the duration of the plateau and the peak location within the plateau region. The latter measure is crucial to a discussion that takes into account peak shape as an important perceptual cue to listeners' categorisation of contours (cf. Barnes et al. 2012a,b). In the present data, the varying peak location within the plateau indicates that different peak shapes are being produced, even for plateaux characterised by similar spans. Speaker 6f, for example, produces plateaux on *managu* that are similarly aligned in all her productions. In one case, however, the peak is aligned differently from the rest, namely near the end of the plateau, which indicates that there is a relatively shallow rise to a peak and a steeper fall following it. Similar and more extreme variability is seen across repetitions of the same word by most individual speakers. The different realisations of the high region and, by extension, different peak shapes, once again suggest that considerable variability is an intrinsic and non-spurious characteristic of intonational events in Tashlhiyt Berber.

6.4.4 Summary of alignment in simple qwords

So far, the search for systematic alignment has proven unfruitful for qwords in TB. Alignment of the rising-falling event did not reveal a high degree of systematicity, neither with reference to different levels of phonological structure below the word (segment or syllable), nor in terms of different measures characterising the contour (a single F0 maximum and a more broadly defined high region).

In many languages, alignment of turning points in the contour is clearly defined with respect to phonological units below the lexical level, like the syllable or mora. In some cases, pitch accents are moreover characterised by consistent segmental anchoring of (combinations of) turning points. This is not what was found for the high region occurring on qwords in TB.

At the same time, one finding was consistent: A local F0 peak systematically occurred *on* the qword. In this sense, it appears that firstly, the peak is a defining component of the qword tune, and secondly, that the qword forms the domain that constrains its location. Together, these observations could be interpreted as suggesting that an H(igh) tonal target is phonologically associated with the qword domain. Its phonetic alignment, as reflected by the exact location of a peak or high region within this qword domain, on the other hand, is variable in seemingly all possible ways.

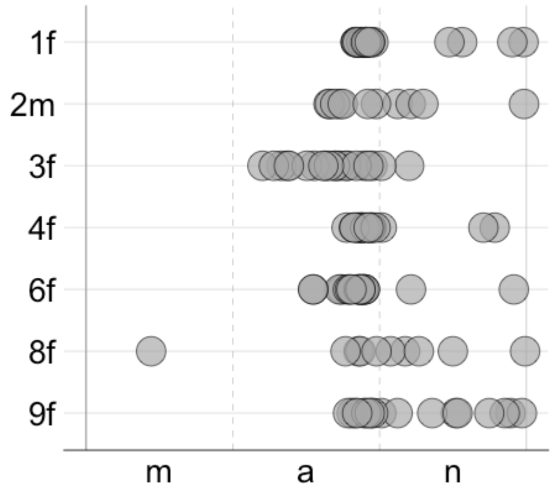

Figure 6.9: Alignment of absolute maximum with respect to segments (separated by dashed lines) of the interrogative element *man* for all six complex qwords.

6.4.5 Alignment in complex qwords

Much in line with the patterns attested for simple qwords, complex qwords are also marked by a rise-fall of which the peak always occurs on the interrogative element *man*. Figure 6.9 shows yet again that the alignment of the maximum within these words exhibits a considerable spread, in this case across /a/ and /n/. The vast majority of maxima are realised towards the end of the vowel, but early alignment within the vowel and later peaks on the following segment /n/ are also attested. What all speakers' productions moreover have in common is that the peak is followed by a steep fall, with most of the falling movement achieved before the right edge of the word *man*. In line with the previous observations about simple qwords, a post-peak fall seems to be an integral aspect of the qword tune. In complex qwords, which exhibited little microprosody, a low turning point can be observed to occur somewhere around the right edge of *man*, or early in the following noun.

In addition to the rise-fall on *man*, there may also be a rise at the right edge of the qword constituent as a whole (i.e. at the right edge of the noun), which is then followed by a fall on the next word, the verb. As an example, contours for the constituent *man butili* 'which shepherd' are shown in Figure 6.10. Three out of seven speakers (1f, 6f, 9f) are rather consistent across repetitions in producing a rise at the right edge of the noun, indicated with the vertical solid line. The number of constituent-final rises by speaker are shown in Table 6.3. Based on this information, speakers can be allocated to two groups: One group produces final rises (1f in 100% of her questions, 6f and 9f in most cases, and 4f and 8f variably), the other group never does (2m and 3f). The factors determining the occurrence of a final rise near the end of the question constituent remain unclear, but consultations with two native speakers suggest that the presence of a rise is simply a stylistic variant which possibly reflects phrasing.

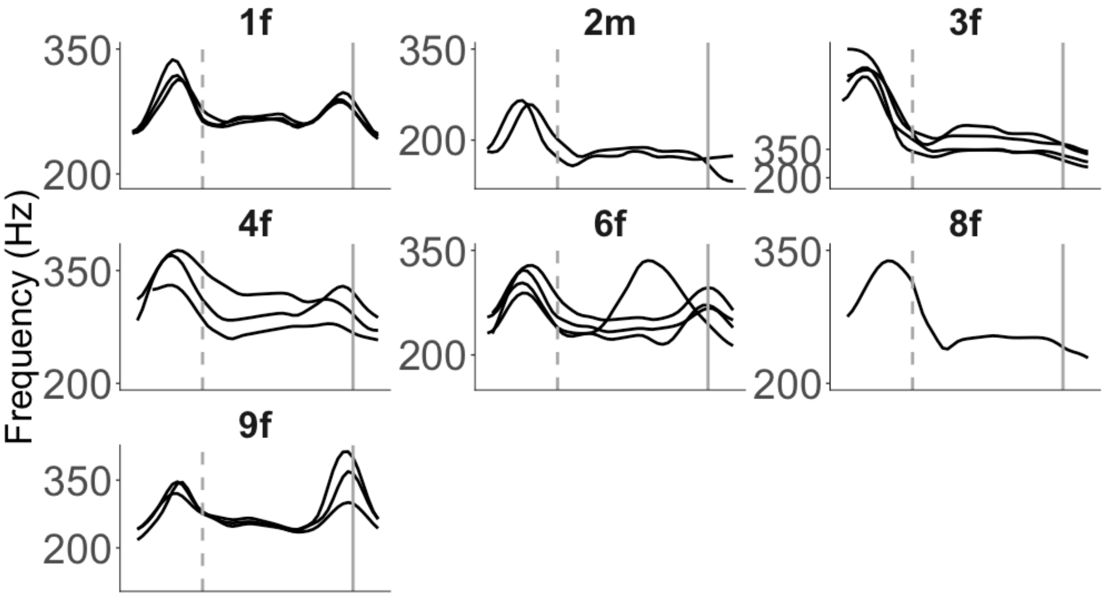

Figure 6.10: Time-normalised smoothed F0 contours for all individual repetitions of *man butili*, the first constituent of the question *man butili nzˁra ʁ umalu?* ‘Which shepherd did we see in the shade?’. F0 is shown in Hertz with reference points of 200 and 350. Dotted vertical line indicates the right edge of *man*, solid line the right edge of *butili*.

	1f		2m		3f		4f		6f		8f		9f	
	N	w/ rise	N	w/ rise	N	w/ rise	N	w/ rise	N	w/ rise	N	w/ rise	N	w/ rise
man anu	4	4	2	0	4	0	4	2	2	1	2	2	2	1
man ananas	4	4	2	0	2	0	1	1	3	3	1	1	4	4
man aʜuli	4	4	2	0	4	0	4	0	3	3	2	0	3	2
man tizi	4	4	1	0	3	0	4	1	2	2	2	2	3	2
man tili	4	4	2	0	3	0	3	0	2	2	2	1	4	4
man butili	3	3	2	0	3	0	3	1	4	3	1	0	3	3
sum	23	23	11	0	19	0	19	5	16	14	10	6	21	16

Table 6.3: Number of tokens produced per speaker and number of these produced with final rise.

6.4.6 Pitch scaling and cue interaction

Given the variable alignment of peaks discussed in the previous sections for both simple and complex qwords, the question arises whether variability in the alignment domain can be related to facts about the scaling of potential tonal targets. Possible trade-offs and interactions between the two domains are well documented (e.g. Gussenhoven 2004) and have the potential to shed light on the reasons why peak alignment is so variable. As was clear from Figure 6.10, scaling of the qword peak varies considerably across subjects, with at the lowest end the single male speaker who produces maxima around 200 Hertz, and at the highest end one of the female speakers, 3f, who produces maxima in falsetto up to 850 Hertz. Such high values are not unusual for speakers of TB (or speakers of MA for that matter), as large pitch excursions and the use of falsetto are a recurrent feature of Moroccan speech (observed in semi-spontaneous recorded data from Roettger & Grice (2015) and in general daily interactions).

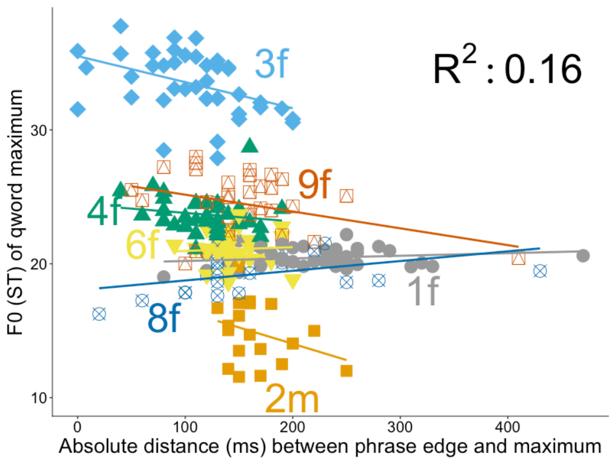

Figure 6.11: Relatively stable scaling of F0 peak on qword for all simple and complex qwords (N = 227).

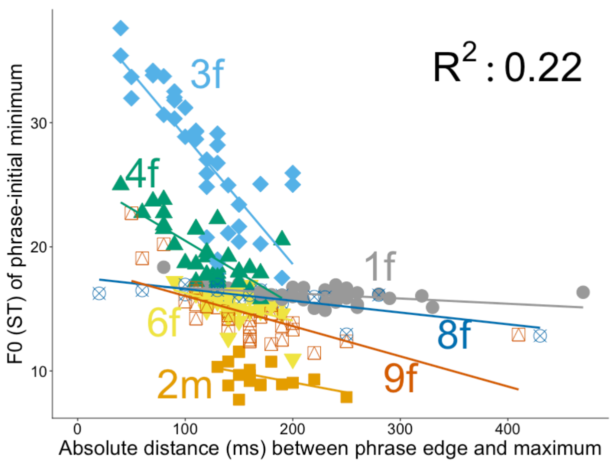

Figure 6.12: Somewhat lower scaling of F0 values at phrase onset with increasing peak distance for all simple and complex qwords (N = 227).

Returning to pitch scaling of the individual turning points, Figure 6.11 shows F0 height for the peak on the qword as a function of peak distance from the left phrasal edge (which coincides with the start of the qword). The correlation is rather weak overall: $R^2 = 0.16$. Speakers moreover appear to have individual preferences for the scaling of this peak: Four speakers (2m, 3f, 4f and 9f) produce somewhat lower maxima with increasing distance from the left phrasal edge. A further two speakers produce more or less the same peak height irrespective of its alignment (1f and 6f) and the remaining speaker (8f) produces marginally higher peaks with later alignment.

Cases in which decreased peak height correlates with later alignment are especially interesting with respect to speaker 3f, who produces the highest but also the earliest aligned peaks. In order to produce a rise from the phrasal edge to a peak shortly after, this speaker must produce very steep pitch rises, or limit rise excursion. While the other speakers exhibit varying interactions between alignment and scaling of the peak, peak height in absolute terms is not influenced much by its alignment. It seems that whereas alignment is relatively unimportant, there is more of a requirement to produce a high target (H) with some specific peak height.

In the main experiment, qwords were phrase-initial, so that the left qword boundary coincided with the left IP boundary. In these cases the minimum value preceding the peak on the qword marked the start of the rise. This stands in contrast to the pilot experiment, where a minimum or turning point usually occurred before the left edge of the qword when it was non-initial (Section 6.2.3). For phrase-initial qwords, the scaling of the F0 minimum at the phrase onset can serve to shed light on whether a compromise is made on the excursion of the rise.

Figure 6.12 shows that when the peak on the qword is aligned later, the preceding low F0 value is typically realised lower, with the strongest effects for the two female speakers with the highest ranges. The overall correlation of $R^2 = 0.22$ is not much higher than for the scaling of the F0 peak but the trend is the same for all speakers. This indicates that alignment of the peak and scaling of the low(er) F0 value at phrase onset are in a trading relation: In cases that require a steep rise (when the peak is early), the rise may be truncated by starting the rise at an initial F0 value that is somewhat higher. The relative stability of peak scaling and the more variable scaling of the starting point of the rise suggest that the peak is the more important turning point in the qword tune.

The properties of the low turning point *following* the peak on the qword were not quantified here as microprosody would make elbow or minimum detection in the relevant region unreliable. Based on observation, the fall following the peak usually entailed a drop to a level similar to that of the phrase-onset F0 value, which is then maintained until the end of the phrase, as illustrated with an example in Figure 6.13.

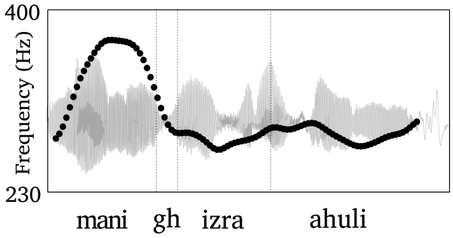

Figure 6.13: Example of an utterance with post-peak fall to baseline in *mani ʁ izˤra aħuli?* as spoken by speaker 4f.

6.5 Discussion

6.5.1 Towards a phonological analysis

The results reported in the previous sections have shown that qwords in direct interrogatives in TB are consistently marked by a localised pitch event on the qword (see especially Figure 6.5). There were no other noteworthy intonational events in the sentence and this section will therefore be concerned primarily with the analysis of the intonational movement that co-occurs with the qword.

All 227 tokens of qwords, 107 simple and 120 complex, were characterised by a local maximum that occurred on the qword, or on the interrogative element *man* in the case of the complex qwords. These peaks often aligned on vowels and tended not to be realised at the very edges of the qword, especially in the case of the simple qwords. There were no apparent phonological or other structural factors that systematically governed peak alignment in qwords: Peaks did not consistently align with a specific syllable in polysyllabic qwords, nor did they align in general with specific structural units such as syllable rhymes or even specific segments. These observations about alignment, together with the context in which qwords were produced (with narrow focus on the qword), strongly suggest that the peak reflects an H target that is somehow linked to the qword, and that it should be interpreted as a prominence-marking intonational event rather than an edge-marking event. This interpretation in terms of a prominence-marking event is further supported by the arguments brought forward in Chapter 5: Qwords are pragmatically prominent, and intonational events co-occurring with qwords in other languages are also usually interpreted as prominence-marking.

The variability in peak alignment in this experiment however presents a puzzle. On the one hand, its consistent occurrence on the qword, and its rather consistent scaling, suggest that it is a central component of the qword tune. On the other hand, its alignment exhibits a type and degree of variability that poses problems for an analysis in terms of association to a specific TBU.

As a first step to arrive at a phonological analysis, the peak alignment in the present study can be compared with the variability observed in previous work on intonation in TB. Grice, Ridouane & Roettger (2015) showed that in phrase-final position, both in yes-no questions and statements with a contrastively focused phrase-final word, the variability in alignment of a local high turning point in the contour could be captured probabilistically by appealing to a number of constraints that favoured the peak to occur on heavier and more sonorous syllables. This raises a number of questions with respect to the present data. Why is the apparent tonal attraction effect of 'prosodically privileged' syllables in final position, as previously reported, not observed in initial position? Unfortunately a direct comparison is not possible, as stimuli in Grice, Ridouane & Roettger (2015) were designed to vary in terms of syllable structure and sonority, factors that only vary in a limited way in the present qwords. Among the present qwords, syllable weight only distinguishes two minimal word pairs, namely *ma/mad* and *mani/manwi*. There is nevertheless little reason to believe that any differences in peak alignment between members of these minimal word pairs should be attributed to syllable weight, as the peak distribution in the present study and the one in Grice, Ridouane & Roettger (2015) is fundamentally different. Grice, Ridouane & Roettger (2015) find a discrete peak distribution, with the maximum being aligned either on the penultimate syllable or on the final syllable, with an attraction effect for peaks to occur on the heavier or the more sonorous of the two syllables. This categorical tonal placement was additionally supported by the authors' auditory impression that the peak was located either on one syllable or the other, and by the peak's systematic alignment in the rhyme of the syllable it occurred on. In the present experiment, peaks on qwords exhibited a far more gradient distribution, with peaks also occurring on intervocalic consonants in onset or coda position, and lacking systematic alignment relative to any specific segment or sub-lexical structural unit. While many of the target words in Grice, Ridouane & Roettger (2015) had obstruents in syllable onset position, which would have prevented peaks from occurring on onsets, an explanation in terms of segmental make-up can account for only some of the differences in alignment between experiments. Other target words in the earlier study had liquids in onset position, and no peaks occurred on any of these onsets either. In contrast, in the present data nasals in syllable onset position did carry peaks.

Differences in tonal alignment between experiments could be due to many factors, including the difference in phrasal position of target words (initial in the present study, versus final in Grice, Ridouane & Roettger 2015), as well as differences in the function of the intonational event (narrow focus in the present experiment versus yes-no question modality marking and contrastive focus). Some of the differences in alignment patterns might nevertheless still be due to differences in the segmental make-up of the words.

At this point, peak alignment patterns in qwords do not provide enough evidence to argue for the association of the peak, in the form of an H tone, to a specific sub-lexical unit. There are of course other potential tonal targets that might form part of the qword tune and are perhaps more systematically aligned, such as potential L tones on both sides of the H, which might be invoked to represent the potential requirement to have a local rise or fall. Their alignment could unfortunately not be investigated with the present data. The results from the pilot however can be used to argue that the rise is not as integral a component of qword intonation as the peak. Rises in interrogatives with non-initial qwords preferentially occurred prior to the qword, with the start of the rise coinciding with the left phrasal edge. Alignment in these cases suggests that any low turning point does not co-occur with the qword. Instead, if an L tonal target should be posited to account for the presence of a steep rise, it could be interpreted as seeking association to the phrasal edge.

Even more speculation applies to the realisation of the low turning point marking the end of the fall. There were some differences between qwords, notably with *ma* exhibiting a stepped contour down to some baseline level. All other words, including the complex qwords, showed a pattern in which a local minimum was produced shortly after the peak, suggesting that there does appear to be an L tonal target that accounts for the steep fall. This will be left for future work, as the decision to include tonal targets in a phonological representation should ideally also take into account paradigmatic contrast. To this end, it would be useful to know whether a shallower fall, or a less complete fall to a baseline, results in a different pragmatic interpretation from the steep fall observed here. If so, this would form evidence to support an interpretation of an HL tonal sequence marking the qword in the present case, in contrast to for example a single H target to characterise cases with a shallower or no immediate fall.

In sum, the high turning point can be interpreted as representing an H tonal target by virtue of its systematic occurrence on the qword. Since no further association to a TBU below the level of the word could be determined, I suggest to analyse the H tone as a 'non-metrical pitch accent'. The 'non-metrical' part of this characterisation follows from the variable alignment of the H tone paired with the absence of lexically stressed syllables, and the 'pitch accent' part follows the AM tradition to distinguish between delimitative 'edge tones' and culminative prominence-marking 'pitch accents'. Even if the analysis in terms of an H accent would have to be revised to include further tonal targets, the absence of lexical stress in the language would still justify the use of the term 'non-metrical pitch accent'.

In the context of the discussion about the nature of the tonal event, it should be mentioned that rather different accounts of intonational movements have been proposed for several other languages that are considered to lack stress. Intonational systems in such languages are typically analysed in terms of predetermined tonal strings that associate sequentially within small phrasal domains such as APs or PPs (for Korean: Jun 2005a, for French: Post 2000; Jun & Fougeron 2002, for Mongolian: Karlsson 2014), or even the IP (for Ambonese Malay: Maskikit-Essed & Gussenhoven 2016). The present results, on the other hand, indicate that the intonational event marking qwords in TB can be analysed rather straightforwardly as serving the purpose of prominence-marking, not edge-marking. This makes the intonational event in question much like a pitch accent, which by definition is prominence-marking, but unlike a pitch accent in the sense of its lacking further association to a TBU, specifically a stressed syllable.

In sum, the present qword data exhibit a type of intonational marking that does not refer to sub-lexical metrical structure. This means on the one hand, that Tashlhiyt Berber clearly differs from languages with lexical stress and postlexical pitch accent, and, on the other hand, from languages like Mongolian, French and Korean, in which there is a role for the mora or the syllable in determining the location of intonational tones.

Finally, it is possible that an analysis in terms of non-metrical tonal association is appropriate for other languages as well. A recent case has been made for Ambonese Malay, which is another language that lacks lexical stress, as also exhibiting intonational tonal alignment that does not involve reference to a "word-internal synchronisation point" (Maskikit-Essed & Gussenhoven 2016). The intonational tones in question were analysed as boundary tone complexes associating with the right edge of an IP, but the phenomenon of variable tonal alignment appears very similar to what is observed in the present case.

6.5.2 Answers to the research questions

The research questions asked in 6.1.3 are repeated here:

1. What are the phrasal prosodic characteristics of question word interrogatives in general and the prosodic properties of qwords in particular?

2. What is the prosodic prominence status of question words? How should any prosodic marking be analysed phonologically?

Question 1 can be answered as follows: Qword interrogatives in TB are characterised by a main phrasal intonational event which co-occurs with the qword. This tonal event takes the form of a rising-falling pitch movement with a peak or high region, with the F0 maximum always aligned on the qword. Both peak and high region were unsystematically aligned, suggesting that the realisation of the intonational event is rather unconstrained. Following the fall from the maximum there are no further intonational events: The contour stays low until the right edge of the interrogative phrase. An exception to this pattern are complex qwords, which are optionally characterised by an additional, edge-aligned rise at the right edge of the noun following the interrogative word.

As for question 2, it is clear that qwords attract the main phrasal prominence, whether the qword is in initial position (as in the main experiment) or in non-initial (medial/final) position (as in the pilot). In the context of the main experiment, qwords were produced under narrow focus, so that on a semantic-pragmatic account, qwords are prominent. Consequently, the intonational event marking this constituent was considered to serve a prominence-marking rather than an edge-marking function. This interpretation is supported by the alignment of the peak, which occurred relatively freely within the domain of the qword, rather than near one of the edges.

While the role of low turning points in shaping the qword tune could not be analysed in great detail, it was clear that an H target is among the defining aspects of the qword tune in TB. This followed from i) alignment and scaling of the peak (which occurred on the qword and was scaled at a consistent height for each speaker), ii) the scaling of the start of the rise (which was more variable and depended in part on the alignment of the peak), and iii) the alignment of the start of the rise (with the pilot suggesting that the rise starts prior to the qword when enough segmental material is available).

In conclusion, an AM phonological analysis of the qword tune in Tashlhiyt Berber involves at a minimum an H target associating to the qword, and an optional H associating to the right edge of the noun in a complex qword, as schematised in Figure 6.14. The rise towards this high target can be accounted for by positing a low target that reflects an L intonational tone associating with the left phrasal edge, i.e. an L%.

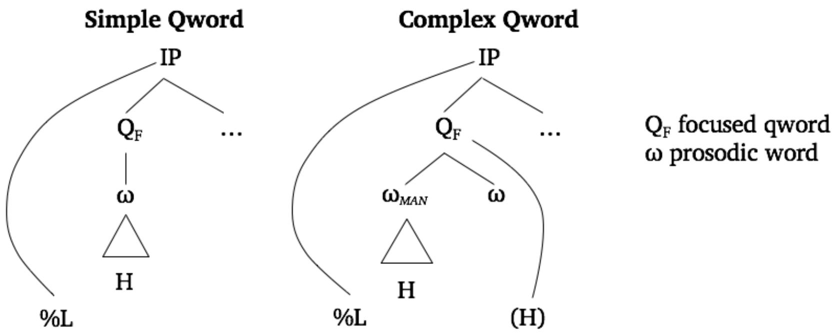

Figure 6.14: Tonal association in TB qword interrogatives.

6.6 Summary and conclusion

This chapter provided a first investigation of question word (qword) interrogative intonation in Tashlhiyt Berber, which is also the first investigation of qword interrogative intonation in Berber in general.

The results of a pilot experiment revealed that qwords attract the main intonational event in a direct interrogative phrase. When they are embedded, in contrast, qwords are not prosodically prominent.

The main experiment served to investigate the prosodic marking of qwords in more detail. To this end, data were analysed from seven participants who read scripted questions in context, with qwords that occurred in their default position phrase-initially, and under narrow focus. It was shown that question word interrogatives in TB are characterised by a rising-falling F0 movement that invariably occurs on the qword. Following this movement there are no further intonational events and the intonation contour typically remains flat, continuing as a level stretch of low F0 up until the right edge of the phrase (an IP).

The main rising-falling event on the qword was interpreted as a prominence-marking intonational event that associates with the qword. This interpretation followed on the one hand from the pragmatic status of the qword as focused, and on the other hand from the alignment of the high turning point, which did not behave like an edge tone in terms of alignment. The alignment of the high turning point (F0 peak) specifically was prone to a great deal of variability. Peaks always occurred on the qword, but did not align with word edges, nor did they systematically occur on specific segments or syllables. The alignment of the starting and end points of a high plateau, as a potential alternative characterisation of the intonational event, exhibited similar variability. Finally, the presence of a rise preceding the peak was argued to follow from the likely presence of a low phrase-initial boundary tone, and was provisionally accounted for by positing a %L marking the left edge of the interrogative IP.

It was argued that the best representation of the qword tune is in terms of an H, non-metrical pitch accent that associates with the qword, but has no further association to a tone-bearing unit below the lexical level.

7 Question word interrogative intonation in Moroccan Arabic

7.1 Introduction

7.1.1 Prior work on the intonation of Moroccan Arabic

Most of what is currently known about Moroccan Arabic intonation is limited to qualitative observations, found primarily in Benkirane (1998) and Maas (n.d.: Ch. 10). The former's analysis includes a concise inventory of the prosodic properties that are characteristic of various sentence types, including yes–no questions, declaratives, imperatives and question word (qword) questions. Whereas Benkirane's (1998) claims are based on read sentences from various speakers and his observations as a native speaker, Maas's () work is corpus-based, describing, on a case by case basis, examples of various sentences in context.

Experimental work on MA is limited, too, and can be divided into two main thematic areas: Prosodic marking of focus, and prosodic marking of yes–no question intonation.

For the former, two papers discuss MA focus marking with reference to other language varieties, namely Burdin et al. (2015) and Yeou, Embarki & Al-Maqtari (2007).[1] While both sets of authors consider their work to look at contrastive focus, results of the two studies are not directly comparable. Burdin et al. (2015) investigate focus marking in a game setting, contrasting three focus conditions in noun phrases consisting of a noun + adjective: i) noun–only focus, ii) adjective–only focus and iii) full noun phrase focus. Yeou, Embarki & Al-Maqtari (2007) look at contrastively focused single words in read speech consisting of question–answer sentence pairs.

For the latter, the intonation of yes–no interrogatives, there is one study based on data from both elicitation and semi-spontaneous interaction between speakers (Hellmuth et al. 2015).

What is currently known about the topic of this chapter, qword interrogatives, is limited to observations. According to Benkirane (1998: 354), the qword attracts the phrasal pitch peak, and this peak is followed by "rapidly falling pitch on the rest of the utterance" . This is in line, he observes, with the description of qword interrogatives given by Rhiati-Salih (1984). Maas (n.d.: Ch. 10) also reports similar patterns, supported by example F0 contours.

[1]The same dataset is reported on, from a slightly different angle, in Yeou et al. (2007).

7.1.2 Stress and tune-to-text association in Moroccan Arabic

The role of word-level stress in determining intonational tune-to-text association remains rather elusive. Benkirane (1998) makes comments in passing about the role of stressed syllables in shaping intonational patterns. As discussed in Chapter 4, he assumes that either the penultimate or final syllable of a word is stressed, depending on the weight of the final syllable: As a rule, the stressed syllable is the penultimate syllable of a word, unless the final is heavy, in which case stress targets this final syllable. For a more detailed discussion and experimental results that shed doubt on this conception of stress see Chapter 4. In any case, from Benkirane's (1998) general description it can be deduced the default main sentence accent in a declarative, as well as focal pitch accents, are characterised by a high target. He further states that default accents typically have a rise on what he considers the stressed syllable, and that contrastive accent peaks align variably with respect to the stressed syllable depending on the word's position in the phrase.

Burdin et al. (2015) adopt Benkirane's view on stress, but find that the location of focus in the noun phrase (i.e. on the adjective, on the noun or on the noun phrase as a whole) is not consistently differentiated by means of a different kind or alignment of pitch prominence. They argue that MA instead uses phrasing to mark focus prosodically, and conclude that MA lacks pitch accents altogether. This seems a claim that extrapolates beyond the available evidence, since it is well known that there are differences among languages in terms of whether focus *within* the noun phrase is marked prosodically, and the absence of focus marking in this specific context does not imply an absence of focus marking altogether (cf. Krahmer & Swerts 2001; Swerts, Krahmer & Avesani 2002; Swerts 2007 on this type of difference between Dutch and Italian, among other languages.)

Yeou, Embarki & Al-Maqtari (2007) also assume that there is lexical stress in MA, although no information is given about how they determine the stress position in their target words. The authors find that both broad and contrastive focus are signalled by a rising–falling movement, where in broad focus "the F0 peak is aligned within the end of the stressed vowel in closed syllables, but it is aligned after the stressed vowel in open syllables" (Yeou, Embarki & Al-Maqtari 2007: 321). In the contrastive focus condition, the rise–fall (which is realised as a localised pitch event in Yemeni and Kuwaiti Arabic) may span the entire word in Moroccan Arabic. Finally, they find that four out of their five MA speakers deaccented the words preceding the contrastive focus. Judging from the single example contour they provide, neither does there seem to be any pitch movement in the postfocal region.

Finally, Hellmuth et al. (2015) provide the only detailed investigation of yes–no interrogative intonation in MA. The authors observe that the alignment of the rise–fall which consistently marks the right edge of yes–no questions is neither fully predictable with reference the right phrasal edge alone nor with reference to the stressed syllable alone (their definition of stress is taken from Benkirane 1998). Instead, the rise–fall is aligned with reference to both of these domains simultaneously, in that it aligns consistently with the final foot of the phrase-final word.

In sum, the role of stress in determining patterns of postlexical pitch prominence in MA is currently not clear. Given the above sets of findings it seems unlikely, however, that MA intonational prominence will be characterised by highly systematic tune-to-text mapping, in the way we have generally come to expect turning points in the F0 contour to align to well-defined segmental landmarks (cf. Section 2.4.3.2).

7.1.3 Question word interrogative structure in Moroccan Arabic

Qwords in Moroccan Arabic typically occur in phrase-initial position and thus belong to the syntactic category of languages that require qwords to occur in a position that differs from the location of a corresponding non-qword constituent (cf. Section 5.2.1). Qword questions with an initial qword in MA appear to exhibit structural similarities to clefts (i.e. focus constructions). Ouhalla (1999: 339) for example argues that these constructions both involve movement to a specifier position in the left periphery. Although typically found in phrase-initial position, qwords may occur phrase-finally in some contexts as shown with an example in Section 5.2.1. They are also found phrase-medially, for example as in one of the target sentences used in this experiment, which was created with the help of native speakers and independently confirmed to be felicitous by Jamal Ouhalla (p.c.).

As for the morphological form of qwords, MA has qwords like *imta* ‘where’ and *fin* ‘when’, which are commonly used in other varieties of Arabic. MA also frequently uses *ʔaʃ*, which may function as a question particle, or as a qword meaning ‘what’. In the present study it is used in one interrogative in conjunction with a preposition, namely *fʔaʃ* ‘in what’.

7.1.4 Research questions

The main research questions are identical to those in Chapter 6 on qword intonation in Tashlhiyt Berber:

1. What are the phrasal prosodic characteristics of question word interrogatives in general and the prosodic properties of qwords in particular?
2. What is the prosodic prominence status of question words? How should any prosodic marking be analysed phonologically?

Based on the aforementioned observations there are a number of predictions for the answer to Question 1. The first is that the qword is most likely to be prosodically prominent and that it can be expected to carry the main pitch event in the sentence, specifically in that it takes the form of a peak followed by a fall. The right phrasal edge of the qword interrogative as a whole is expected to be low rather than rising.

The possible answer to Question 2 remains more open. Does the phrasal maximum align on the qword, and how is this high turning point best represented? These questions are concerned with the nature of the relevant pitch event as edge marking (a phrasal or edge tone) or prominence-marking (some type of accent), or even a combination of both (see also Chapter 2).

In addition to these main questions, there is one more area of interest, concerning the effect of bilingualism. As discussed in Chapter 1, language contact in Morocco has apparently led to convergence between the (Tashlhiyt) Berber and (Moroccan) Arabic phonological systems on a macro-level. On a lower level, this chapter can contribute an answer to the question as to what extent bilingual and monolingual speakers of these languages show speaker-level differences (based on whether they speak both or one language only):

3. Are there any differences in the patterns of MA qword intonation produced by 'monolingual' MA speakers on the one hand and 'bilingual' MA/TB speakers on the other? If so, might any differences be explained by appealing to bilinguals' transfer of patterns found in TB?

7.1.5 Data

This chapter reports on an experiment that involved questions being read aloud by participants as part of a scripted mock dialogue, the transcription and translation of which is given in Appendix B. As in Chapter 4, data are part of the IVAR corpus (Hellmuth & Almbark 2017) and were kindly made available to me by the authors. Further details about each specific interrogative are given in Section 7.2.3.

7.1.6 Overview of chapter

In the next section, Section 7.2, the methodology of the present experiment is explained. This is followed by the results in Section 7.3. The discussion in Section 7.4 starts with possible phonological analyses for qword interrogative intonation in Section 7.4.1, followed by a more general discussion of intonational categories in MA in Section 7.4.2, and answers to the research questions in Section 7.4.3. Finally, Section 7.5 gives a brief summary and conclusion.

7.2 Methodology

7.2.1 Participants

There were 24 participants, identical to those in Chapter 4 on correlates of stress in MA (for details see Section 4.2.1). All speakers were university students in Casablanca, with the majority being born and raised there. Participants belonged to one of two groups with 12 speakers each: One was a 'monolingual' group with participants who grew up speaking only MA at home, and one 'bilingual' group with participants who spoke Tashlhiyt as a first language in addition to MA. All speakers have some to near-native proficiency in MSA and French.

7.2.2 Procedure

The present task was performed by pairs of participants as part of a set of recordings for the IVAR corpus. Recordings took place in a quiet room at the Université Hassan II in Casablanca, with participants each wearing a headset microphone and being recorded on a separate channel. Participants were presented with a scripted dialogue (Appendix B) printed on paper, and were instructed to play one of the roles. The dialogue was performed twice by each pair of speakers, so that each participant performed each role once. All sentences in the scripted dialogue are therefore available for each participant.

7.2.3 Speech materials

The dialogue served to elicit a number of other constructions in addition to qword interrogatives. There were in total six target qword questions which used five different qwords: *ʃnu* 'what' (occurring in two different interrogatives), *ʃkun* 'who', *imta* 'where', *fin* 'when' and *fʔaʃ* 'in which'. Target questions and their glosses are given below, in order of their IVAR code (not the order in which they occurred in the dialogue).

whq1 *ʃnu (1)*

شنو سميت هاد الراجل اليمني؟

ʃnu	smit	ɦad	lraʒəl	ljamani
what	name.POSS.3SG.M	this	man	Yemeni

'**What** is the name of this Yemeni man?'

whq2 *ʃkun*

شكون للّي غيشهد على العرس الرومي؟

ʃkun	l:i	ʁajʃɦad	ʕala	lʕars	r:umi
who	REL	FUT.witness.IPFV.3SG.M	on	wedding	civil

'**Who** is it that is going to witness the civil wedding?'

whq3 *imta*

إيمتا عرس بنت عمك دينا؟

imta	ʕars	bint	ʕam:ək	dina
when	wedding	daughter	uncle.POSS2SG	Dina

'**When** is the wedding of your cousin Dina?'

whq4 *fʔaʃ*

العرس غيكون فأش من مدينة؟

lʕars	ʁajkun	**fʔaʃ**	min	mədinə
wedding	FUT.be.IPFV.3SG.M	in.which	from	city

'**In which** city will the wedding be?'

whq5 *fin*

فين شافت دينا نبيل؟

fin	ʃafit	dina	nabil
where	see.PFV.3SG.F	Dina	Nabil

'**Where** did Dina see Nabil?'

whq6 *ʃnu (2)*

شنو غتجيبي لدينا من حانوت دلال؟

ʃnu	ʁatʒibi	ldina	min	ħanut	dalal
what	FUT.bring.IPFV.2SG.F	Dina	from	shop	Dalal

'**What** will you bring Dina from Dalal's shop?'

Target questions differed in a number of respects. A summary of the most relevant differences is given in Table 7.1. The number of syllables is based on the phonological transcription.

IVAr code	Qword	N syll qword	Syntax of qword	N syll in phrase
whq1	*ʃnu (1)*	1	phrase-initial	8
whq2	*ʃkun*	1	phrase-initial, clefted	9
whq3	*imta*	2	phrase-initial	8
whq4	*fʔaʃ*	1	phrase-medial, in-situ	8
whq5	*fin*	1	phrase-initial	7
whq6	*ʃnu (2)*	1	phrase-initial	11

Table 7.1: Details of target qword interrogatives used in the experiment

In addition to the factors listed in the table, whq4 additionally differs from the other questions in that its qword is morphologically complex, in that *fʔaʃ* is historically decomposable into *fi* 'in' + *ʔaʃ* 'what, which'. Whq4 also differs from the rest with respect to the type of qword, because *fʔaʃ* is part of the complex question constituent 'in which city'. [2] The relevance of these factors to the realisation of qword interrogative intonation will be discussed in Section 7.4.

7.2.4 Analysis

The theoretical total number of utterances was 144 (6 qwords * 12 speakers * 2 groups = 144). Out of this total, 11 utterances were non-targetlike due to mispronunciations or major disfluencies and excluded. A further four utterances were excluded due to smaller disfluencies in the vicinity of the qword. This resulted in a final number of

[2] Similarly, *fin* can historically be decomposed into *fi* 'in' and *ʔajnə* 'where', but this form has long been grammaticalised.

65 utterances analysed for the monolingual group and 64 utterances for the bilingual group (total N = 129).

Acoustic measurements were taken in Praat (Boersma & Weenink 2015). Initial segmentation of utterances into words and segments was performed automatically by Prosodylab-Aligner (Gorman, Howell & Wagner 2011). Subsequently all utterances were manually checked and segmentation was adjusted where needed. F0 measurements are based on a version of Praat's automatically generated pitch contour which was manually corrected for pitch tracking errors and smoothed with 15Hz bandwidth. Local minima and maxima were detected automatically and verified manually.

Statistical analysis was performed in R (R Core Team 2016). For comparisons between speaker groups, linear mixed models were run with the package *lme4* (Bates et al. 2015). Significance for individual predictors or interactions between predictors were calculated by means of LRTs between a main model and a corresponding null model lacking the relevant interaction or predictor. The R syntax of the main model is in each case given in a footnote.

7.3 Results

This section consists of three parts. First, in Section 7.3.1 an overview is given of the phrasal intonational patterns of the read sentences is given. What is observed here serves to motivate the choice to focus on the alignment and scaling of specific turning points, the initial low and the following peak, as discussed in Section 7.3.2, as well as the time course of the fall, discussed in Section 7.3.3.

7.3.1 Global contours

Figure 7.1 below shows time-normalised phrasal contours for all individual qword interrogatives (total N = 129). It is clear that qwords are characterised by a region of high pitch, with for some qwords a visible rise to a peak, and for others a high onset (notably *ʃkun*).[3] The shape seems to vary from peak-like in interrogatives with *ʃnu* to almost plateau-like high regions in *ʃkun*. A local maximum seems to occur either on or shortly after the qword, followed by a considerable drop. Taken together, these interrogatives provide a highly coherent picture of a rising–falling contour around the qword and an absence of additional intonational prominence later in the phrase.

In terms of right-edge marking, most questions end low, which is in line with the observations by both Benkirane (1998) and Maas (n.d.: Ch. 10). There is a handful of cases with final rising F0 movement (N = 5 / N = 64 for the bilingual group and N = 8 / N = 65 for the monolingual group). The five rises in the bilingual group were

[3]There was one exception in the whole dataset concerning a voiceless rendering of *fʔaʃ* as [fʔʃ] followed by a rise–fall on the final word *mədinə*. This particular final contour is typical of yes–no questions (Hellmuth et al. 2015) and I will assume that it was wrongly interpreted as a yes–no question (especially as *fʔaʃ* can also be a question particle in a yes–no question). This particular utterance will further be ignored.

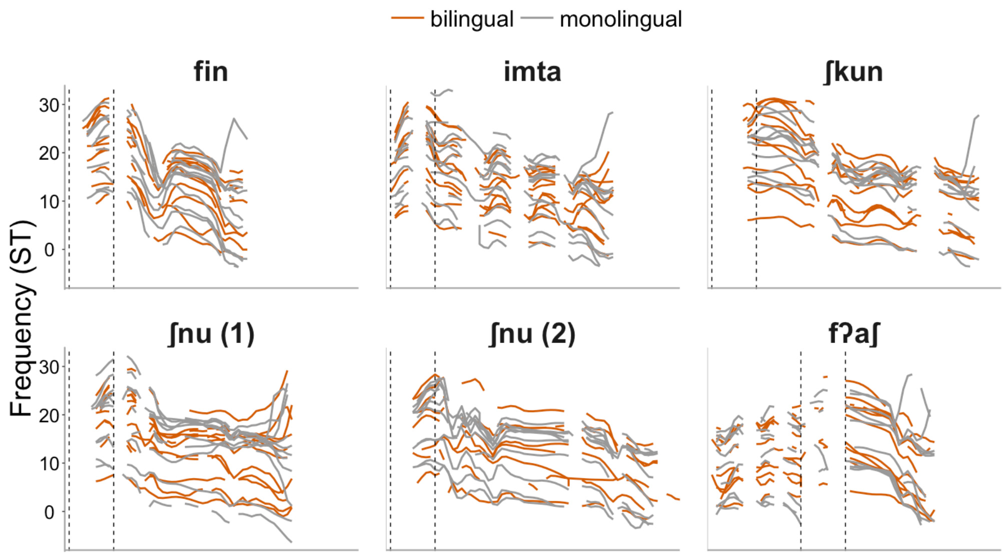

Figure 7.1: F0 contours for all target utterances (N = 129; normalised duration based on 10 equidistant measuring points throughout each word. Dotted lines delimit target qword.

due to four different speakers (two male, two female), but in the monolingual group there was one female speaker, f3, who produced five of the total eight rises. Finally, rises were spread across most utterances, although *ʃnu (2)* lacked final rises altogether. Interrogatives with *ʃnu (1)*, on the other hand, exhibited relatively many rises (27%, or six out of a total 22).

In sum, in this experiment, final rises for qword questions are firstly optional, and secondly, representing only 10% of all qword questions, the more marginal option dispreferred to level or falling final intonation. The distribution of final rises also suggests that there might be something specific about the interrogative with *ʃnu (1)* as opposed to all other interrogatives (the attested six rises among the total 22 renderings for this sentence represent a likelihood 0.014, given the overall frequency of 10% rises). Plausible explanations could appeal, on the one hand, to the notion of epistemic bias: This particular question involved only given constituents that were all overtly mentioned in the preceding turns, and the speaker would presuppose that the interlocutor knows the answer (cf. Warren 2016 on similar contextual effects influencing final rising intonation or 'uptalk' in English). On the other hand, final rises might be linked to the higher social cost associated with asking a detailed follow-up question requesting highly personal information (see Chen 2012 for the suggestion that higher social cost of asking certain questions might result in deviant accentuation of qwords in Dutch). In any case, the presence of final rises does not in any obvious way correlate with any structural linguistic factors, including information structure of the phrase.

For the sake of a clearer between-group and between-qword comparison, the averaged time-normalised F0 contours for are shown in Figure 7.2 (N = 13 with final rise and N = 1 with final rise–fall excluded). Just as in the above figure, these contours were based on frequency sampling at ten equidistant measuring points per word, but are plotted retaining information about average word duration as well. It can be seen that the average duration of the interrogative phrase with *imta* was shortest, and that among the qwords, *ʃnu (2)* had the shortest duration.[4]

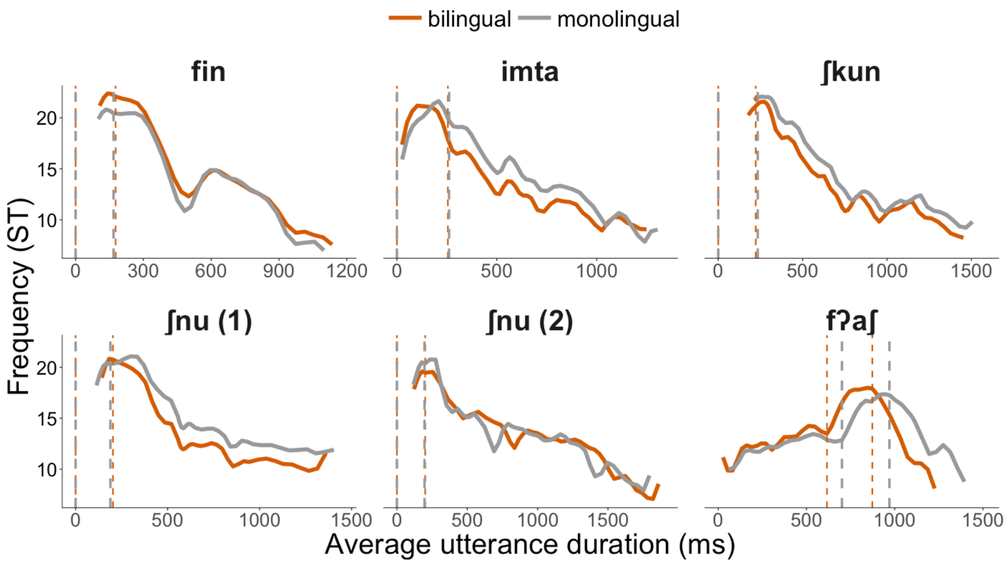

Figure 7.2: Averaged and time-normalised interpolated F0 contours across all target utterances without final rise–fall) (N = 105). Dotted lines delimit target qword.

Two main observations can be made based on this figure. The first is that the global contours for the different interrogative phrases look very similar to each other (excepting, for the moment, the phrase with medial qword *fʔaʃ*). Across the board, the qword carries what seems to be the one and only intonational prominence in the phrase, and it takes the form of a rising–falling pitch movement in most cases. The second point is that these mean contours highlight the apparent absence of differences between the two groups of speakers. There are no categorical differences, and the overall contours look near-identical, suggesting that any differences in qword question intonation linked to the linguistic background of speakers are expected to be subtle at best.

These preliminary findings are compatible with Benkirane (1998)'s observation that the qword attracts the utterance maximum and is followed by "a rapid fall". Clearly, a rise–fall in the vicinity of the qword is a defining intonational characteristic of qword questions.

[4]The considerable dip in the F0 contour occurring around 500 ms. in interrogatives with *fin* is due to microprosody (the mostly voiced sequence /t d/ in /ʃafit dina/) and does not represent a linguistically meaningful intonational movement.

In the next sections I focus on the intonational movement characterising the qword, or qword tune, trying to define any regularities defining the relationship between the tune and the text for this particular word.

7.3.2 Turning point alignment and scaling

As seen in the previous section, qwords in MA are characterised by a rising–falling F0 contour, with the peak of this movement occurring on or in the vicinity of the qword. Its exact alignment will be looked at in detail in Section 7.3.2.1.

Figures 7.1 and 7.2 also suggested that the magnitude of the rising movement towards the peak varies across words (greatest in *imta* and *fʔaʃ* and smaller in *ʃkun*). These, incidentally, are also the interrogatives that have qwords whose peaks do not co-occur with the left phrasal boundary, which can straightforwardly be explained by appealing to the presence of preceding segmental material that *can* carry intonational movements (for *imta* because its first few segments are voiced and for *fʔaʃ* because it is phrase-medial). Given that there is a rise where one is possible, a low and a high turning point demarcating this rise should be identifiable. In order to determine whether either one or both of these low and high turning points can be considered to constitute phonetic reflexes of AM tonal targets I will consider the relationship between them, in terms of the scaling and alignment of the peak and the scaling of the left-edge minimum (Section 7.3.2.2). The start of the rising movement (i.e. the low turning point) is predictable, as it typically co-occurs with the onset of voicing. Given this predictability in alignment, it should not be considered to yield a meaningful alignment measure on its own, but I will still refer to its scaling in relation to the high turning point. Similarly, to investigate properties of the fall, I will look at the temporal progression of the fall as a whole rather than at the alignment of an individual low turning point (Section 7.3.3).

7.3.2.1 Peak alignment

Table 7.2 shows the distribution of qword-related maxima.[5] A first observation concerns the fact that in 83 out of the 129 utterances (64%) the absolute maximum of the qword-related rise–fall occurred *on* the qword itself. Among the initial qwords the percentage is just a bit lower, with 62 out of 108 (57%) attracting the peak.

In the following, I will discuss whether alignment of these peaks can be characterised systematically by appealing to to several segmental landmarks: Absolute distance from the qword start, absolute distance from the qword end, and relative position within the qword.

Absolute distance from word start

Firstly, peak alignment was calculated with reference to the word start, or left qword edge, to chart the temporal domain across which maxima are spread. There were no

[5]These frequency counts come from the smoothed contours but are almost identical to the peak distribution in absolute contours.

	Peak on qword	Peak not on qword
ʃin	15	8
imta	23	1
ʃnu (1)	11	11
ʃnu (2)	5	13
ʃkun	8	13
fʔaʃ	21	5
Total	83	46

Table 7.2: Distribution of peaks on qword.

differences in absolute alignment relative to the qword start between the two groups of speakers (LRT: $\chi^2(1) = 1.8$, $p = 0.17$),[6] and results are therefore shown pooled in Figure 7.3. For all six qwords, the absolute distance between the word start (indicated by dotted line) and the F0 maximum varies considerably. Mean peak alignment varies between 174 ms. for bilingual *imta* to 244 ms. for monolingual *ʃkun*, but these varying means are not very informative given large alignment differences within single words. In fact, the smallest range of variation in alignment (latest minus earliest) for a single word comprises 185 ms. for *ʃnu (2)*.

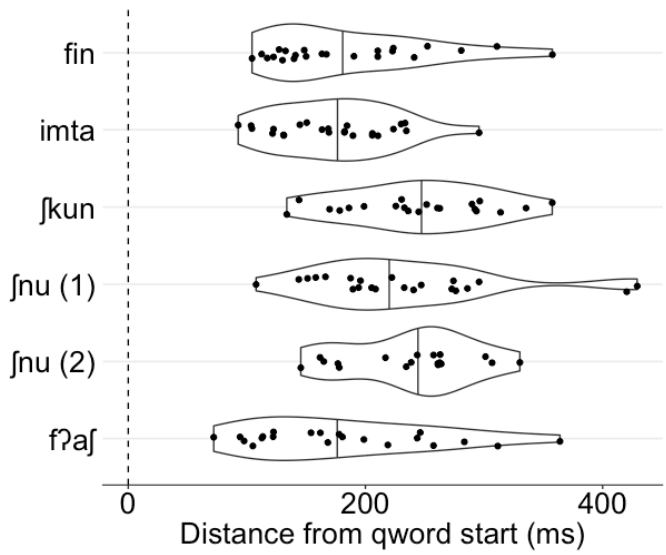

Figure 7.3: Alignment of qword-related peaks for all six qwords relative to the left edge of the qword (dotted line).

Nevertheless, maximum alignment is highly correlated with the word start. Predicting maximum alignment (absolute values) based on the qword start revealed a signific-

[6] peak alignment from qword *start* ~ group + (0 + group|qword) + (1|speaker)

ant correlation coefficient of r = 0.96 (LRT: $\chi^2(1) = 19.3$, $p < 0.001$).[7]

Absolute distance from word end

For the sake of comparison, peak alignment was also calculated relative to the word end (right qword edge). Figure 7.4 shows, for all six qwords, the absolute distance between the F0 maximum and the word end (indicated by dotted line). Again there were no group differences and data are shown pooled (LRT: $\chi^2(1) = 1.7$, $p = 0.2$).[8] Negative values indicate that the peak was on the qword, positive values that it occurred after. Maxima are not quite consistently aligned, with peaks occurring before and after the qword boundary (mirroring the numbers in Table 7.2). Again, there was an overall correlation, with maximum alignment significantly correlated with the qword end, with r = 0.77 (LRT: $\chi^2(1) = 22.8$, $p < 0.001$).[9]

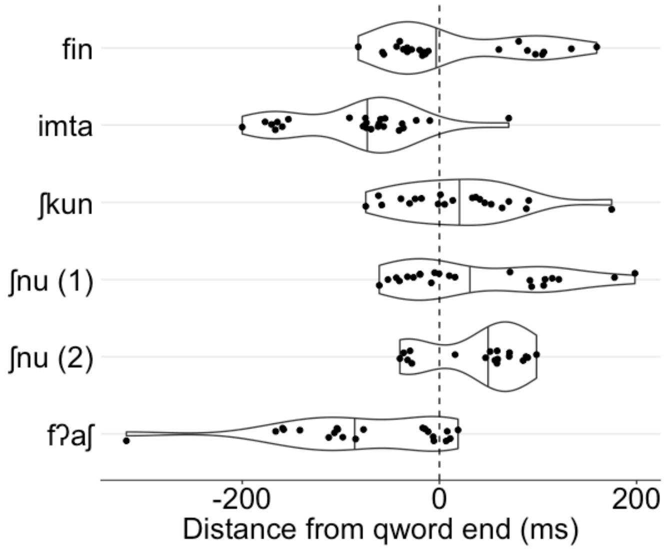

Figure 7.4: Alignment of qword-related peaks for all six qwords relative to the right edge of the qword (dotted line).

The qword *ʃkun* in *whq2* occurred in a cleft, and it seemed possible that peak alignment for this qword might be better correlated with the end of the larger constituent that includes clefted /l:i/. A quick comparison shows that this is not the case. Maximum alignment was better predicted with respect to the (simple) qword edge than with the extended qword constituent edge (correlation coefficients of r = 0.54 and r = 0.34, respectively).[10]

[7] peak alignment ~ qword start + group + (0 + qwordstart|qword) + (0 + qwordstart|speaker)

[8] peak alignment from qword end ~ group + (0 + group|qword) + (1|speaker)

[9] peak alignment ~ qword end + group + (0 + qword end|qword) + (0 + qword end|speaker)

[10] peak alignment ~ qword end + group + (1|speaker) and peak alignment ~ extended qword end + group + (1|speaker)

Comparing word edge alignment

The above two ways of characterising peak alignment resulted in the observations that there is, for every single qword, a large temporal domain within which peaks may occur. At the same time, the phonological domain of the qword clearly plays an important role in determining peak location. This was reflected in the high overall correlations between peak alignment and both word edges.

	relative to *word start*	**relative to *word end***
fin	185 (67)	13 (71)
imta	174 (50)	-84 (65)
ʃkun	244 (61)	19 (60)
ʃnu (1)	232 (90)	36 (86)
ʃnu (2)	237 (53)	38 (49)
fʔaʃ	181 (96)	-78 (112)
mean	209 (72)	-9 (86)

Table 7.3: Mean alignment values of maxima, relative to qword start and qword end (ms.), SD (ms.) in brackets.

In order to directly compare the magnitude of this variation with reference to each edge, table 7.3 lists, per word, the mean alignment values across all participants. Two main patterns are observed in the data, the first to do with average alignment and the second with the degree of variability in the data.

Firstly, peak alignment is *on average* very close to the right qword edge, much closer than to the left qword edge. In case of a global analysis only, this might be taken to argue for right edge-alignment of the peak and subsequent interpretation as association to the qword edge. However, given the overall variation and different behaviour of individual words, this grand mean falsely suggests that all peaks align with the right qword edge.

Secondly, the large standard deviations for both measures for each qword indicate that the means are not very informative. Since there were high correlations for peak alignment with both word edges, it seems that the edges play some role in peak alignment. The qword edges nevertheless do not constitute strong absolute limits on the domain within which peaks are realised, nor do peaks align systematically within the qword domain when they do occur within it. The range of absolute alignment values seems especially large in comparison with results reported in earlier studies that set out to test predictions of the segmental anchoring hypothesis (see Chapter 2).

In order to make this comparison explicit, I will quickly review the details on alignment from two much-cited studies. In Ladd, Mennen & Schepman (2000), the authors investigate rising accents in Dutch by looking at the alignment of low and high turning points marking the start and end of a rising movement on a stressed syllable. For the high turning point, mean speaker alignment values ranged from 10.8 ms. (short vowel condition) to 19.9 ms. (long vowel condition) (based on their Table 1). Similarly, in a study on rising L* + H accents in two dialects of German, Atterer & Ladd (2004) found that the temporal domain across which high turning points occurred spanned 50 ms. with standard deviations of 13 and 17ms. (values reflecting means per speaker/group in their table 1). The high turning point is subsequently interpreted as the phonetic reflex of a H trailing tone (and the accent thus being L* + H). Trailing tones are generally analysed as such by virtue of being *less* systematically aligned than the preceding, starred, tone. This suggests that the range of variation for H alignment attested in Atterer & Ladd (2004) is already relatively large compared to tonal alignment of starred tones.

While these measures of variation in alignment concern small groups of speakers, and controlled speech under laboratory conditions, it does seem that the alignment patterns in the present data are of an altogether different nature. In the present data, the temporal domain across which maxima are realised is, as mentioned previously, 91.8 ms. in the cases of the 'smallest' domain, and standard deviations for the variation in alignment for individual words start at 49 ms. Even if not directly comparable, these numbers are indicative of substantially different degrees of variability in alignment.

The next section will delve into alignment just a bit further, to see if relative measures might still reveal a somewhat more systematic pattern otherwise missed. The phrase-initial peak in MA qword interrogatives is apparently not systematically aligned to any part within the qword, but the qword edges nevertheless delimit the domain within which peaks are realised. So far, this is in line with the initial hypothesis that there is no segmental anchoring for the peak in the qword rise–fall in MA. It could still be conjectured, however, that some of the variation seen here can be explained away by accounting for inter-speaker differences in speech rate. Additionally, the segmental make-up of the words in the experiment varied and some words were more likely to have post-qword peaks than others, see table 7.2. This suggests that some more variation can be accounted for by investigating peak alignment relative to segments.

In order to address whether one or both of the above factors might play a role in determining peak alignment, the next subsection will investigate peak alignment relative to segment duration.

Relative alignment within qword

In order to control for the potential confounds of speech rate and segmental make-up, alignment was also considered in relative terms, for which peak alignment relative to the duration of individual segments within and immediately following the qword was calculated.[11] Alignment of F0 maxima relative to segments is shown in Figure 7.5. No peak was aligned more than six segments to the right of the start of the qword, and for the phrase-medial word *fʔaʃ* no peaks were realised prior to the qword start.

Firstly, it is clear that there is considerable similarity in maximum alignment for bilingual and monolingual speaker groups on this measure too, confirming that there are no group differences (LRT: $\chi^2(1) = 0.28$, $p = 0.6$).[12]

[11] Since maxima were extracted automatically based on smoothed and interpolated F0 contours, maxima may also occur on voiceless segments.

[12] normalised peak alignment ~ group + (0 + group|qword) + (1|speaker)

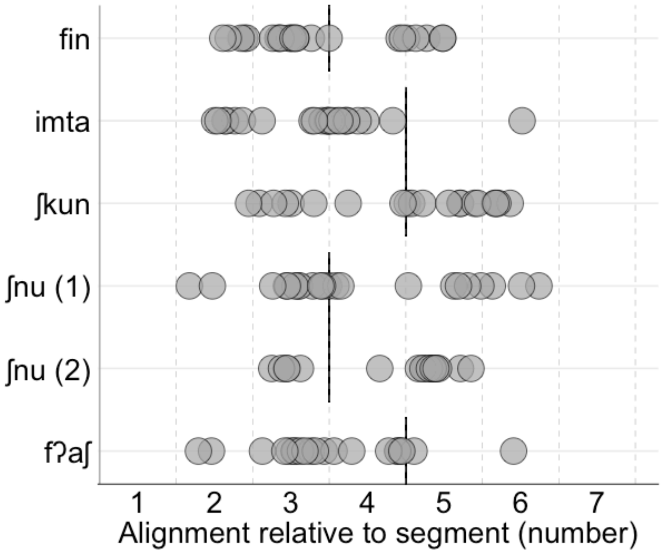

Figure 7.5: Alignment of F0 maxima in all utterances, relative to normalised segment duration (dotted lines) and word boundaries (solid lines).

Secondly, there is a word-specific effect for *imta* and *fʔaʃ* in that peaks in these words typically align before the right edge of the qword. An obvious explanation would appeal to the fact that *imta* is disyllabic word and has a first segment that is phonetically voiced (both unlike all other qwords). Interestingly, in none of its 21 renderings does the peak actually occur on the initial vowel, with the /i/ instead being used to carry a rising pitch movement. This suggests that if enough voiced segmental material is available, both a rise and a peak will be realised on the qword. Under suboptimal circumstances, as in all other qwords, the peak might be realised to the right of the qword, or the initial rise gets truncated. To what extent either one of these possible strategies is more consistent with the observed data will be explored in the next section in which scaling is investigated.

Returning to the issues mentioned at the end of the previous section, it seems that even if speech rate and segmental make-up are considered, peak alignment remains highly variable. Based on both absolute and relative peak alignment results I conclude that F0 peaks associated with MA qwords are not aligned systematically with reference to segmental anchors. F0 peaks lacked consistent alignment in terms of:

- absolute distance from the start of the qword (left edge) (Figure 7.3)
- absolute distance from the end of the qword (right edge) (Figure 7.4)
- relative position within the qword (Figure 7.5)

Nevertheless, peak location does seem to be governed to some extent by the edges of the qword domain, as peak alignment exhibited high correlations with both edges.

From a pragmatic point of view (see discussion of prominence and focus in qword interrogatives, Chapter 5) the qword is a good candidate of a domain to which a pitch event might seek to associate. If the qword domain is indeed relevant in governing tune-to-text association, it is worth exploring whether properties of the qword tune beyond alignment can explain why peaks in some cases are realised outside this domain.

7.3.2.2 Scaling of initial low turning point and peak

The finding that peak alignment is not systematic while still being linked closely to the qword raises the possibility that this variability is caused by other, potentially competing requirements on F0 realisation. For example, later peak alignment might stem from the need to realise an initial rising movement of a given size, where a rise and a peak on the qword compete for limited time.

Here I will explore whether it is the slope of the rise which forms a consistent feature of qword intonational marking. For this, duration of rise (distance between maximum on or after the qword and the preceding minimum) and rise excursion (difference in ST between this maximum and minimum) were calculated. For all qwords, the qword minimum was identified as the minimum F0 value between the qword onset and the qword maximum.

The location of the minimum typically coincided with the onset of voicing. For the qword *fʔaʃ*, which is largely voiceless, the minimum was identified in the interpolated F0 contour, and typically yielded a point at the start of the qword. As mentioned previously, I do not take this alignment measurement to yield an informative insight into the alignment of a low turning point, but it is useful for purposes of rise duration.

The results show that scaling of both turning points exhibited much variation within the dataset, with a range of around 30 ST for both the low and the high turning points. There were no differences between speaker groups for either the excursion size of the rise (LRT: $\chi^2(1)=0.12$, $p=0.72$) nor for the temporal distance between start and end of rise (LRT: $\chi^2(1)=2$, $p=0.16$). The following analysis is therefore based on the full dataset of $N=129$, pooled across the speaker groups.

For the excursion size of the rise, values ranged from 0.2 to 11.3 ST with a mean of 4.0, excluding the seven utterances for which there was a complete absence of a rise (three renderings of *ʃnu (1)* and four of *ʃkun*). Additionally, the number of utterances for which the excursion size was 2 ST or less was 32, equal to 25% of the total, which suggests that the presence of a clear rise is a common but not a mandatory aspect of qword intonational marking.[13]

Details of the rise are variable across qwords, as is also suggested by the mean contours shown in Figure 7.2. Within each of the six utterances there was predictable variability in alignment of the low turning point (which coincided with the onset of voicing), and any correlation between its scaling and its alignment will be influenced by this. The alignment and scaling of maxima, in contrast, seems to be independent:

[13] 2 ST is chosen here as a somewhat arbitrary cut-off point for perceptible changes in F0, see also Section 6.3.4).

Maxima that are aligned later are not convincingly higher (weak correlations with R^2 varying from 0.0001 for *ʃkun* to 0.19 for *ʃnu (2)*). Correlations can be expected in cases where there are temporal constraints on the realisation of a series of tonal targets, especially if low and high targes occur in quick succession. In such cases the scaling might be compromised or alignment delayed (see also Chapter 6). The lack of a correlation between scaling and alignment of the high turning points in MA qwords suggests no such bidirectional effects on realisation.

To consider the rising movement more holistically, Figure 7.6 serves to show the relation between the scaling of the start and the end of the rise (minimum F0 value near left phrasal edge and peak on or after qword, respectively). There is a correlation between these two, manifested as a clear linear relationship indicating that the ratio of maximum relative to minimum is constant: For any given minimum, the following maximum is scaled roughly 42% higher (overall $R^2 = 0.86$). Similarly, a linear model taking into account by-item and by-speaker variation also finds a strong dependency between the turning points in the scaling domain (LRT: $\chi^2(1) = 32.7$, $p < 0.001$, $\beta = 1.01$).[14] Of course, this correlation on its own does not give any information about the directionality of the effect.

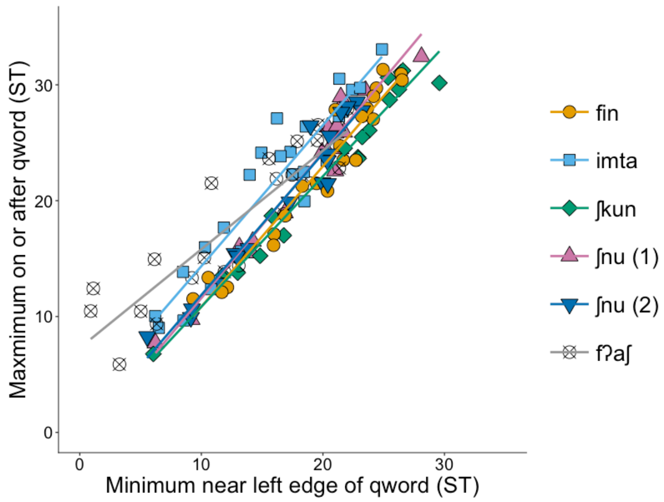

Figure 7.6: Scaling of qword minima and maxima, regression lines given for each qword separately.

In terms of possible causation and directionality, the consideration of speech planning mechanisms suggests that the low turning point (i.e. the start of the rise) might best be interpreted as providing the reference scaling in the relationship. If the reference were the high turning point instead, the scaling of the low turning point would have to be determined in hindsight. In any case, the link between the two points in the scaling domain suggests that both are involved in the qword tune together, which can be interpreted as evidence for an analysis that posits an LH bitonal sequence to

[14] maxF0 ~ minF0 + group + (0 + minF0|qword) + (0 + minF0|speaker)

represent the rising movement. Moreover, if one of the defining prosodic characteristics of qwords involves the presence of a rising F0 movement, and the rise starts at the left edge of the qword (or soon thereafter at the onset of voicing), then variable peak alignment can be analysed to follow from the requirement to realise some degree of a rise.

Similar argumentation, in which an L tonal target is posited based on a scaling dependency with a following H target, is found in Ladd & Schepman (2003). This set of authors argues for the existence of an L target based on a scaling dependency between a low and high turning point in the context of a so-called sagging transition between two H targets in British English. The particular low turning point they discuss had previously been considered a mere transition effect with a phonetic explanation rather than a phonological target.

7.3.2.3 Interim summary and implications for analysis

In sum, this section on turning point scaling and alignment showed the following results:

- in most cases there is an identifiable rise on the qword towards a peak
- the F0 maximum marking the peak is aligned variably: It may occur on or after the qword
- there is a consistent scaling relationship between F0 minimum and maximum for each qword rise

I interpret these findings as providing evidence for the qword tune involving both an L and an H target, in the form of an L + H bitonal accent. No starred tone is included for two reasons, which are to a large extent overlapping with the reasons given for the absence of a starred tone in the TB qword tune (see Chapter 6).

Firstly, there is some unclarity about the alignment of the L, and much attested variability in the alignment of the H. If the alignment of the high turning point had been consistent with respect to a specific sub-lexical phonological unit, then this would have provided some evidence for phonological association to this unit. There is however no apparent moraic or syllabic unit that serves the purpose of tonal anchoring.

Secondly, and more fundamentally, there is unclarity surrounding the existence of stress in MA (see Chapter 4). If MA has no lexical–metrical positions in the form of stressed syllables, then these cannot be expected to play a role in determining tonal association, including starredness, in the first place.

A possible analysis of the L + H tonal complex is that it seeks association to the qword, and the fact that the alignment of the peak is not always aligned within the domain of the qword can then be explained by the requirement to realise at least some rising movement. This discussion will be taken up in more detail in Section 7.4.

7.3.3 Properties of the post-qword fall

So far this chapter has focused on the properties of the rise towards the peak and the peak itself. As could be seen in Figure 7.1, and judging from Benkirane's description of "rapidly falling pitch", qword peaks are also followed by a steep fall. This section will explore the properties of this fall. The low turning point following the peak could not reliably be extracted from either the raw or interpolated and smoothed contour due to microprosody, and therefore the methodology employed here considers the overall trajectory rather than a single low turning point. The trajectory of the fall was quantified as follows: For each utterance, one F0 measure was taken at the midpoint of each vowel. Then, the F0 span was calculated based on these values (maximum as found on the the qword vowel or the next vowel, minus minimum value as typically found on the phrase-final vowel).

Figure 7.7 shows the mean trajectory of the fall for each utterance, the vowel with the maximum F0 serving as reference point.[15] The trajectory for *fʔaʃ* is noticeably shorter because the qword and peak occur only 4 syllables prior to the end of the phrase.

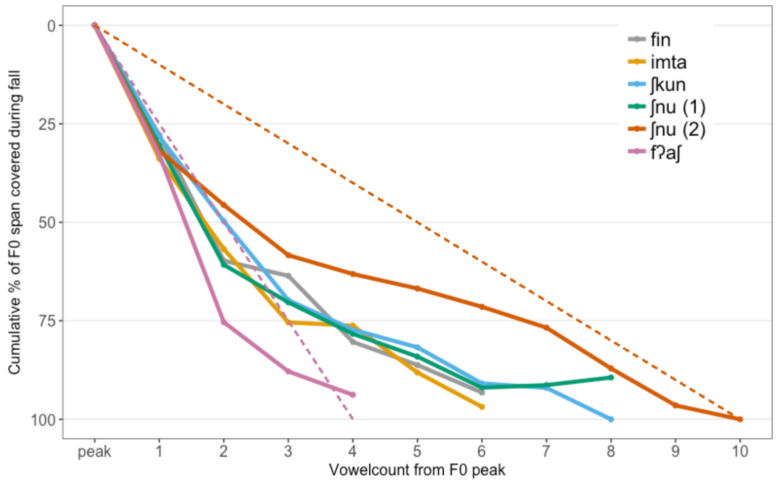

Figure 7.7: Trajectory of fall as cumulative percentage of utterance F0 span covered, starting with the vowel carrying the maximum. Dashed lines illustrate hypothetical linear interpolation between peak and final vowel (for shortest and longest interrogatives only).

It is clear that there is a steep fall following the peak in all cases: On average 28% of the entire range is covered by the F0 distance between peak and immediately following vowel, and the 50% mark is reached around the second vowel after the peak in all utterances. The trajectories then start to diverge. This seems best explained by appealing to the number of vowels between peak and end of utterance. *Fʔaʃ* is followed

[15]The reason why the 100% mark is not reached for all utterances is due to averaging across all individual fall trajectories, including those that exhibit smaller rising movements at any point after the peak (utterances with a final rise were excluded from the mean calculation).

by only four vowels and has a much steeper fall than the longer utterances with *ʃnu (2)*. The other four utterances have different lengths but nevertheless display highly similar trajectories, suggesting that the differences in steepness of the fall for the short and long utterances are indeed best interpreted as adjustments from a default.

The results for the fall are informative with respect to the earlier finding that peak alignment was not systematic. The fall trajectory shows that once the peak is considered relative to what *follows* (i.e. in controlling for varying alignment by taking the peak as reference point), the observed contour is comparable across utterances, with the trajectory of the fall being highly similar for the first two vowels following the peak. Results also indicate that there is some effect of utterance length in that there are adjustments in steepness of fall after the second vowel following the peak: The longer the utterance, the shallower the fall. Utterances are nevertheless alike in exhibiting a similarly steep fall immediately following the qword peak.

On the one hand, this could be taken to argue in favour of the presence of another low tonal target shortly after the peak. Further support for this comes from the idealised scenario with a linear fall between the peak and the phrasal minimum at the end of the utterance. This linearly-interpolated fall scenario is illustrated for the longest and shortest utterances in Figure 7.7 with a dashed line. It can be seen that the trajectories for the falls as they are attested in the present dataset deviate considerably from a linear fall. While no interpolation between intonational tonal targets is ever really linear, the difference seems appreciable enough to look like there should be another low target causing the rather steep fall.

On the other hand, the inflection points in this figure do not necessarily represent turning points in the actual F0 contours. At the very best, they give an indication as to where steeper falls turn into shallower falls. Further work will be needed to determine the location of an actual low turning point following a qword peak in MA, and to investigate the possibility that there is a second L phonological target involved in the MA qword tune.

7.4 Discussion

7.4.1 Towards a phonological analysis

The above findings showed that the MA qword tune consists of a sharp rise, starting at the left qword edge, to a peak reached on or shortly after the qword, and a relatively steep fall following this peak. Earlier in this chapter (in Section 7.2.3), it was mentioned that there were differences between the interrogatives investigated in terms of phonological and syntactic structure, and there might have also been pragmatic differences (since each sentence occurred at a different point in the scripted dialogue). None of these factors had an obvious effect on the intonational structure of the interrogatives, since i) qwords in all interrogatives received the main phrasal pitch prominence, ii) there were no other prominence-marking events in the utterance, and iii) the degree of variability in phonetic realisation within and between utterances suggest that there are no categorical differences between utterances dependent on information structural differences. Any small differences in alignment between utterances could be linked to phonetic constraints on the realisation of pitch.

The result of main phrasal pitch prominence coinciding with the qword provides in first instance strong evidence to posit some sort of high accent associating with the qword. Further claims about the phonological interpretation of qword prominence in terms of specific tonal targets will be discussed here. There are two sets of issues to consider: The first has to do with the nature of the intonational movement, building on observations about the domain it associates with, the second with the event's phonological or symbolic representation in terms of tonal targets.

The first issue deals with the question of what the nature of this intonational event is. Presumably the rise–fall, or high region, serves a prominence-marking as opposed to edge-marking function, an interpretation supported firstly by the inherent prominence of the qword itself (see Chapter 5), and secondly by the fact that the phonetic maximum of the F0 event is clearly linked to the qword domain. Very similarly to what was observed for TB in Chapter 6, the peak does not consistently appear at either of the qword edges, and instead occurs roughly anywhere on or shortly after the qword. The peak also does not seek out a phrasal edge, as the phrase-medial qword illustrated: The qword still carried the peak in this case. This behaviour would typically classify the tonal event as a pitch accent, if it were not for a marked absence of association to specific syllable (and specifically a stressed syllable). Metrical prominence as a criterion for phonological association in MA is moreover problematic in itself, since the existence of stress is generally doubtful, and unresolved at best (see Chapter 4). Any claims about the event's association to a specific syllable, stressed or otherwise, is hampered further by the fact that qwords in this experiment, as in MA in general, are mostly monosyllabic. The results for the only disyllabic word do not resolve the question of association. Peaks in *imta* were just as well characterised by unsystematic phonetic alignment, and peaks could occur on either one of the two syllables.

The second possible criterion to determine association are patterns of phonetic alignment. This criterion is problematic, too, because the alignment of turning points did not reveal much systematicity. The location of the first low turning point was largely determined by phonetics (the start of the rising movement, and thus F0 minimum, occurred at the onset of voicing), and the high turning point was shown not to align systematically with any phonological domain below the qword. The second low turning point, marking the transition from a fall to the level F0 characterising the rest of the phrase, was difficult to locate reliably. An interpretation that could follow from the present findings is one that invokes association of the pitch event to the qword domain. Its prominence-marking function, which is highly similar to that of pitch accents in other languages, paired with an absence of evidence for association to a TBU below the word level, would qualify it as a non-metrical pitch accent. This is much like what seems to happen in Tashlhiyt. In the case of MA, the non-metrical pitch accent's domain of association also is the qword, which in direct interrogatives is likely to act as a focus domain.

The second issue deals with the F0 event's symbolic–phonological representation. Specifically, in an AM analysis, one could interpret the above findings as suggesting that there is a series of low, high and low turning points that are candidates for a phonological analysis in terms of L, H and L tonal targets, respectively. The maximum appears to play an important role because it is pivotal in accounting for the relatively high F0 and because it tended to phonetically align with the qword. If there is indeed a role for an H target, the next question is whether it forms part of a monotonal or bitonal (or even tritonal) sequence involving targets for the low start- and end points of the intonational movement. The following observations constitute reasons to believe that there is a role for L targets on one or both sides of the H target:

- For a preceding L: The local character of the rise towards the peak. There was a steep rise preceding the peak on the qword. This is especially informative in the case of whq4 where the rise towards the phrase-medial qword peak could theoretically have started much earlier than at the onset of the qword.
- For a trailing L: The relatively steep fall immediately following the peak (Figure 7.7), and the notable absence of contours in which the high region on the qword was extended into a high plateau, as in the case of TB (Chapter 6), or as reported for Spanish (Prieto 2004; Henriksen 2014).

These observations could be taken as support for various analyses, for example along the lines of an LH, HL, or even LHL non-metrical pitch accents.[16]

7.4.2 Intonational categories in Moroccan Arabic

No AM-style accentual categories have been proposed for MA as of yet, although two recent experimental studies provide contributions to the topic of intonational categorisation. Yeou, Embarki & Al-Maqtari (2007), who investigated read MA sentences that contained a single contrastively focused word, concluded that MA uses a rising–falling F0 contour on these specific words. Due to its occurrence on pragmatically focused words, the relevant tonal event seems a typical candidate for an analysis in terms of a prominence-marking tonal event, but the authors do not offer a further phonological analysis.

Hellmuth et al. (2015) (see also Section 7.1.2) investigate the intonational movement found at the right edge of yes–no questions (a rise–fall). Their alignment results reveal the hybrid behaviour of the tonal event in question, which exhibits some degree of edge-alignment, while its location is also influenced by what is assumed to be a position of metrical prominence (the stressed syllable as conceived of by Benkirane 1998).

Returning to the use of AM-style labels for analysis of the present data, one of the main difficulties is the absence of information about paradigmatic contrast of intonational events, i.e. about which accents may occur in the same context or position. At

[16] Tritonal pitch accents are typically avoided in AM-style analyses but are proposed in a few cases, e.g. El Zarka 2011 for Egyptian Arabic.

this point it is simply not known whether there is an interpretative difference between a gradual and a steep fall following the qword peak, and by extension, if a trailing L should be posited to distinguish one pragmatic meaning from another.

As for syntagmatic contrast, on the other hand, all work so far has noted the existence of local rising–falling movements (rather than, for example, extended high regions, rises without sharp falls or falls without sharp rises). It is not clear however whether Yeou, Embarki & Al-Maqtari's (2007) contrastive focus movement is any different from the present qword rise–fall either in phonetic realisation or in terms of the domain of association (although it seems that both rise–falls take words rather than individual syllables as the domain over which they are realised, see also Section 7.1.2).

7.4.3 Answers to the research questions

Of the questions below, the first two questions were also answered for qword intonation in Tashlhiyt Berber in the previous chapter. The third question is relevant to the comparison of the two participant groups in this study.

1. What are the phrasal prosodic characteristics of question word interrogatives in general and the prosodic properties of qwords in particular?
2. What is the prosodic prominence status of question words? How should any prosodic marking be analysed phonologically?
3. Are there any speaker group differences between Tashlhiyt/Moroccan Arabic bilinguals and non-Berber Moroccan Arabic speakers?

Question 1. can be answered as follows: Qword interrogatives in MA are characterised by main phrasal pitch prominence on the qword, irrespective of where the qword occurs in the phrase. This intonational prominence takes the form of a rising–falling pitch movement, followed by a flat stretch of F0 until the end of the phrase. The right edge of the phrase is typically low but may end rising.

To answer Question 2. it was argued in Section 7.4 that the prosodic marking of qwords reflects the pragmatic prominence of the qword rather than serving the purpose of prosodic phrasing. The intonational event is clearly linked to the qword both in terms of its phonetic realisation and in terms of its phonological association to the domain of the qword, which for independent reasons may be assumed prominent if not focused (see Chapter 5). As for its AM-analysis in terms of level tonal targets, I argued that further tonal targets than H cannot be posited with certainty. While there are good reasons to assume that either a leading L (i.e. an L + H accent) or a trailing L (i.e. H + L), or both leading and trailing L targets will be needed to account for the local character of both the rising and the falling movements, further insights in how intonational categories in MA might contrast with each other would be required to specify which, if any, of these L tonal specifications are needed. In sum, at this point, the tune is best analysed simply as an H(igh), non-metrical pitch accent.

The answer to research question 3. is negative. There were no obvious differences in the overall shape of the contour nor were there statistical differences in the realisation of the intonational event associated with the qword. The present results and the results for Tashlhiyt Berber in the previous chapter are highly similar, in the sense that both languages exhibit the co-occurrence of the phrasal pitch prominence, in the form of a rise–fall, with the qword. They moreover exhibit similar phonetic properties such as variable peak alignment. The absence of a group difference between monolingual (MA) and bilingual (MA and TB) speakers is therefore unsurprising.

7.5 Summary and conclusion

This chapter showed that question word interrogatives in Moroccan Arabic are characterised by a rising–falling F0 contour on the question word. Following this movement there are no further intonational events that could be considered to serve prominence-marking functions, as the intonation contour typically consists of a low stretch of F0 that continues right up until the right edge of the interrogative phrase as a whole. This intonation contour (main high prominence on the qword) is in line with Benkirane's (1998) description of MA qword interrogative intonation. This experiment additionally shows that there is an optional final rise at the right edge of the phrase (i.e. the right edge of the IP). The choice of a rise over a continued low F0 seemed to reflect speaker-specific preferences, but is also likely to depend on pragmatic factors. In general, the present results did not reveal any differences between bilingual and monolingual speaker groups.

As for the main rising–falling event on the qword, the alignment of the maximum (a peak) was not systematic. Pitch maxima could occur on or after the qword, and overall alignment patterns per qword spanned large temporal domains. Peak alignment was however also highly correlated with both qword edges, suggesting that there is a role for the qword as a domain to which the intonational event seeks association, even if alignment is not systematic in the sense of exhibiting segmental anchoring. The start of the rise was scaled with reference to the scaling of the peak, indicating that there is an additional role for the relationship between the start- and end point of the rise. The trajectory of the fall was similarly steep across all different interrogative phrases, suggesting that there is also a requirement for a steep fall in the prosodic prominence marking of qwords in MA.

In terms of tonal targets, I argued that the rising–falling qword tune is best analysed in terms of an H target only. Even if there are reasons to assume that the local rise and fall are best accounted for by positing additional L targets in the accent's representation, too little is known about MA intonational phonology at this point to be sure about their inclusion in a phonological representation.

In terms of the nature and association of this H target, I argued that the accent is best conceived of as a non-metrical pitch accent. Firstly, the function of high pitch on qwords in MA can be appropriately analysed on a par with prominence-marking pitch accents in other languages. There was also no evidence for an edge-marking interpretation of the tonal movement. Secondly, the absence of a role for stressed syllables in determining tune-to-text association strongly suggests that phonological tune-to-text association does not need to refer to TBUs below the lexical level. In conclusion, qword interrogatives in MA are interpreted as involving a non-metrical H accent that associates to the qword.

Part IV

Prominence perception

8 Prominence deafness in Tashlhiyt Berber and Moroccan Arabic speakers

8.1 Introduction

8.1.1 Prior work

8.1.1.1 The phenomenon of stress deafness

Over the last two decades, a very productive line of research has investigated the perception of prosodic contrasts reflecting lexical stress in participants with different language backgrounds. Much work has been done by Dupoux, Peperkamp and colleagues (Dupoux et al. 1997; Dupoux, Peperkamp & Sebastián-Gallés 2001; Dupoux & Peperkamp 2002; Peperkamp & Dupoux 2002; Dupoux et al. 2008; Skoruppa et al. 2009; Dupoux, Peperkamp & Sebastián-Gallés 2010; Peperkamp, Vendelin & Dupoux 2010). More recently, other groups of researchers have also conducted experiments inspired by this work (Correia et al. 2015; Rahmani, Rietveld & Gussenhoven 2015; Hellmuth, Muradás-Taylor & Karrinton to appear).

The main shared finding across all experiments is that certain groups of participants (i.e. native speakers of specific languages) exhibit 'stress deafness'. A prototypical stress deaf group is formed by native French speakers, in contrast to for example Spanish natives. Stress deaf listeners struggle with reliably distinguishing between stimuli that vary in the location of prosodic prominence. Their relative inability to deal with a prosodic prominence contrast is all the more robust because this behaviour is not only different from other listener groups, but also stands in contrast to the same listeners' ability to deal with segmental contrasts. Stress deaf and non-stress deaf participants typically perform similarly well on segmental-phonological contrasts.

Over the years, stress deafness has been tested by in slightly varying ways, and the perceptual tasks and stimuli have varied considerably between experiments. In the following, I will describe a typical stress deafness experiment as conducted in the later studies (i.e. Dupoux et al. 2008; Dupoux, Peperkamp & Sebastián-Gallés 2010; Rahmani, Rietveld & Gussenhoven 2015).

The task that has proved to reliably yield stress deafness effects is a so-called sequence recall task (SRT). In an SRT, listeners are presented with sequences of words (typically nonce words in the native language of the participants) that differ only in terms of where the main prominence is located. Two example word pairs that have been used to test the 'stress' or *prosodic* contrast are /ˈnumi/ ~ /nuˈmi/ and /ˈmipa/ ~ /miˈpa/ (as in Dupoux et al. 2008; Dupoux, Peperkamp & Sebastián-Gallés 2010). The exact phonetic correlates of the difference between these words will be discussed below. Example word pairs exhibiting a *segmental* contrast as used in SRTs are /ˈfiku/ ∼/ˈfitu/ and /ˈmunu/ ∼/ˈmuku/.

The first part of the task involves a learning phase that serves to ensure that participants correctly associate the contrasting members of a minimal stress pair with designated keyboard keys (e.g. key ‘1’ for initial stress, and key ‘2’ for final stress). In this phase, individual words are presented to the listener, and the answer involves a forced choice between the two categories.

The actual experiment, the SRT, follows the learning phase and involves longer sequences of these words, ranging from two to six words. An individual experimental trial thus could involve a three-word sequence such as /ˈnumi/ /nuˈmi/ /ˈnumi/, or /ˈmuku/ /ˈmunu/ /ˈmuku/, followed by the word “OK”. The playing of the word “OK”, typically in another voice, serves to prevent listeners from using acoustic memory, so that the task taps into short-term memory and by extension reflects listeners’ categorical representations. Listeners respond to what they have just heard by typing in what they think are the appropriate keys in matching order (e.g. 121 for the aforementioned sequences).

The crucial finding is that on the trials involving more than two words, different groups of listeners exhibit markedly different behaviour, where those that exhibit low accuracy on the prosodic contrast are considered to be stress deaf. This low accuracy is understood in two ways: i) low compared to the same participants’ scores on the SRT involving the segmental contrast, and ii) low compared to the scores of other groups of participants’ (with a different native language) on the prosodic SRT. A second finding of these experiments is that response accuracy decreases at longer sequences for all groups of listeners, irrespective of whether they are considered stress deaf or not, and irrespective of whether the contrast is segmental or prosodic. The overall difference in performance between stress deaf and non-stress deaf participants on the prosodic SRT is nevertheless maintained even at longer sequences. Both types of effects (asymmetry between groups of participants and decreasing performance with increasing sequence length) are shown for the five groups tested in Rahmani, Rietveld & Gussenhoven (2015), in Figure 8.1.

This graph illustrates a typical stress deafness effect as occurring in some participant groups as opposed to others. Firstly, all listener groups give correct responses for more than half of the trials on the *segmental* SRT, irrespective of sequence length (although accuracy does decrease somewhat with increasing sequence length). Secondly, clear differences between the groups arise in the case of the prosodic, but not the segmental SRT. Of the five groups that featured in the study by Rahmani, Rietveld & Gussenhoven, participants whose native language is Persian (Farsi), French or Indonesian are considered stress deaf, whereas the Dutch and Japanese participants are not. Dutch and Japanese participants give correct responses to more than half of the trials with sequence lengths of three and four words, and to about half of the trials with a sequence length of five words. In contrast, at no sequence length do the other three groups reliably recall sequences of words that differ in the location of prosodic prominence.

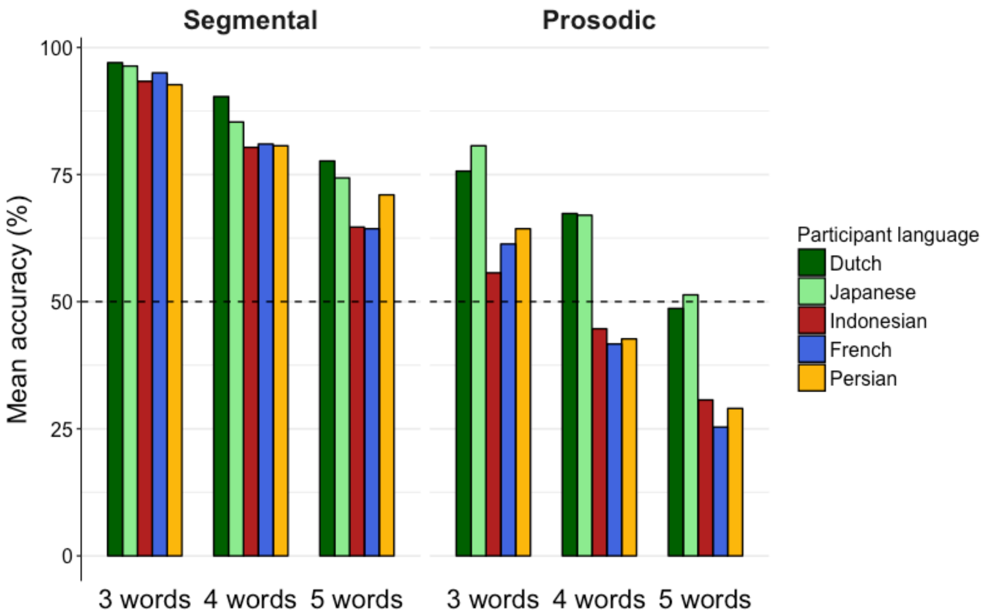

Figure 8.1: Mean accuracy scores for segmental and prosodic SRTs, per native language group and per sequence length, based on average participant scores from Rahmani, Rietveld & Gussenhoven 2015 (Appendix). Dotted line reflects 50% or half of the responses correct.

It should be noted that stress deaf participants can nevertheless *perceive* the prosodic difference between stimuli such as /ˈnumi/ and /nuˈmi/: This is why most of them pass the learning phase of the task, where they learn to associate prosodically differing words with a specific key. Words presented for discrimination in this phase are presented in isolation, which removes the element of memory storage seen in the sequence recall phase. Additionally, an earlier study also showed that French (stress deaf) participants' performance on a task that appeals to purely phonetic–acoustic distinction skills, such as an AX discrimination task, is clearly better than on SRTs (Dupoux et al. 1997). These observations highlight that stress deaf listeners are not incapable of hearing a phonetic difference between prosodically different stimuli. Rather, they struggle to encode the difference at a more abstract level, so that in more demanding tasks that tap into categorical–phonological representation their performance deteriorates. Exactly how linguistic background contributes to causing this stress deafness effect will be discussed in section 8.1.1.3.

8.1.1.2 Nature of the prosodic contrast causing stress deafness

In order to avoid confusion, some terminological clarification is necessary at this point: The term 'stress' as used in most stress deafness papers refers to the *surface prosodic contrast* between syllables of a polysyllabic, typically disyllabic, word. This is not the same definition of stress as the one given in this thesis (‘the phonological property of a syllable within a word of being special’). In fact, it is crucial to note that the surface prosodic contrast between stimuli used in most stress deafness experiments is only *partly* due to lexical stress proper. Specifically, typical stimuli used across all experiments are (modified versions of) words spoken in isolation by a Dutch native speaker. Isolated words form an IP on their own and are, therefore, subject to sentence-level prosody. Isolated words thus exhibit what I will refer to as ‘double enhancement’: Characteristics of lexical stress proper paired with characteristics of phrase-level prosody. Lexical stress, in the case of Dutch, manifests itself as differences between the syllables in terms of spectral tilt, vowel quality and duration. Phrase-level prosody involves a nuclear pitch accent associated with the stressed syllable, and phrase-final lengthening primarily targeting the final rhyme of the final syllable (cf. Cambier-Langeveld 2000). ‘Double enhancement’ defined this way can be contrasted with ‘single enhancement’, referring to acoustic differences caused by lexical stress proper, in the absence of phrase-level prominence. This single enhancement is often referred to as ‘acoustic correlates of stress’, as also discussed in Section 2.3.3.

In sum, the stimuli used in SRTs are typically of the double enhancement type.[1] The acoustic properties of one such set stimuli are exemplified in Figure 8.2 for one realisation of the words /nuˈmi/ and /ˈnumi/ each, spoken by a Dutch male speaker. These particular stimuli are from Rahmani, Rietveld & Gussenhoven, but very similar Dutch stimuli were used in a number of earlier studies (Dupoux et al. 2008; Dupoux, Peperkamp & Sebastián-Gallés 2010; Peperkamp, Vendelin & Dupoux 2010).

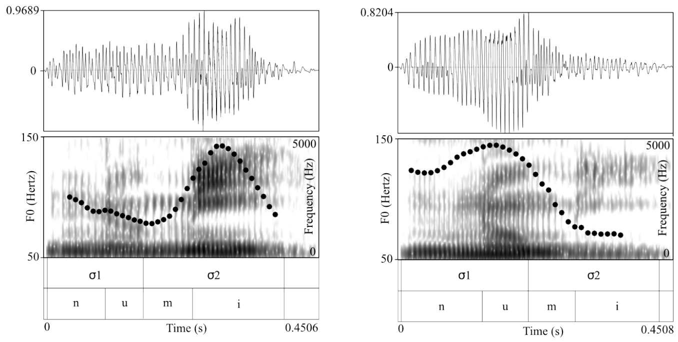

Figure 8.2: Spectrogram, waveform and F0 contour for Dutch pseudo-words /nuˈmi/ (left) and /ˈnumi/ (right) spoken by a male Dutch native speaker. Stimuli from Rahmani, Rietveld & Gussenhoven (2015)

Clearly, the surface acoustics of these stimuli exhibit big differences: The stressed syllable is longer than its unstressed counterpart in the same position, the stress syllable carries most of the pitch prominence on the word and it is spectrally enhanced. Nevertheless, the individual words are by no means mirror images of each other in terms of prominence, i.e. it is not the case that initial stressed syllables are marked by exactly the same type of enhancement as final stressed syllables. As this is relevant for

[1]The main exception being Correia et al. (2015), which will be discussed below.

the discussion in Section 8.4.3, a brief dissection of the prosodic properties associated with both words is given here.

Firstly, while the nature of the pitch accent associating to the stressed syllable is the same for initial and final stressed syllables (according to the ToDI annotation system for Dutch intonation it would be H*L in both cases, cf. Gussenhoven 2005), this pitch prominence is phonetically realised differently in the two words, with a much steeper rise-fall on the word with final stress than on the word with initial stress. This is mainly due to the phrasal position of the syllable the pitch accent belongs with. Final lexical stress in /nuˈmi/ corresponds to absolute phrase-final position, while penultimate stress in /ˈnumi/ coincides with penultimate position in the phrase. This results in the difference in the pitch movement, which in /nuˈmi/ is a rise-fall fully realised on the stressed syllable, as opposed to in /ˈnumi/ where the fall is realised mostly on the non-stressed, final syllable. This difference can be accounted for if we consider that the phrase-final syllable also carries the modality marking L% IP boundary tone used for declarative utterances. This results in a quick succession of tonal targets (H*L L%, i.e. a pitch accent followed by boundary tones) on the final syllable in the case where it is stressed (i.e. in /nuˈmi/).

Secondly, the durational asymmetry between the first and second syllable in /nuˈmi/ is more pronounced than the one in /ˈnumi/, a fact that can be explained by appealing to the interaction of phrase-final lengthening and accent-induced lengthening: In /nuˈmi/, the position of the pitch-accented syllable is phrase-final, which means the final syllable is subject to both of the just mentioned lengthening effects. In /ˈnumi/, on the other hand, the lengthening effects are separated: The initial syllable receives most of the accentual lengthening while the final syllable receives the bulk of phrase-final lengthening, resulting in both syllables ending up with similar duration (cf. Cambier-Langeveld 2000).

Based on the preceding discussion, it would thus seem that stress deafness as it is typically conceived of really might better be called ‘failure to represent a phrase-level prominence contrast parasitic on lexical stress’.[2] Alternatively, one could say that ‘stress deafness’ in its common use refers to a perceptual insensitivity relating to the ‘double enhancement’ type of stimuli. Why it is so crucial to distinguish between ‘double’ and ‘single’ enhancement is highlighted by the results reported in Correia et al. (2015). These authors report on two SRTs performed with native Portuguese speaking participants. In the first SRT (their experiment 2), there were two sets of stimuli (all spoken by Portuguese speakers): One set consisted of isolated words (double enhancement stimuli) while the other set consisted of words excised from a phrase, where the word in question was postfocal and non-IP-final (single enhancement stimuli). For the single enhancement type stimuli, the enhancement consisted primarily of longer dur-

[2]Although stress deafness effects have repeatedly been shown for unmodified or little-modified Dutch stimuli, several stress deafness studies used stimuli that were strongly manipulated in terms of duration and/or pitch, or stimuli spoken by native speakers languages other than Dutch (including Spanish, French, Persian and English). Since similar effects were found in most cases, the phenomenon of stress deafness seems rather robust, although it must be admitted that these manipulations make it impossible to directly compare the relative strength of the deafness effects in each case.

ation for the stressed syllable (vowel quality differences might also be a correlate of stress in Portuguese, but the vowels to which this applies were not used in this set of stimuli, which incidentally also involved the words /ˈnumi/ ~ /nuˈmi/). Notably, these stimuli were not characterised by the presence of a pitch accent, phrase-final lengthening, or boundary marking. The double enhancement stimuli, in contrast, additionally did exhibit this latter type of enhancement.

Results indicated that among the prosodic stimuli, the presence of postlexical prominence resulted in considerably better SRT performance. Specifically, participants gave about 9% correct responses for the single enhancement type target words, and around 22% correct responses for double enhancement type words. They had more correct responses on the phoneme contrast (50%), but here there were no differences as a function of enhancement type. Despite the relatively better performance on the double enhancement type stimuli for the prosodic contrast, the error rates are nevertheless very high for a group of participants who speak a language with variable stress and thus should not exhibit scores that are reminiscent of stress deafness. The authors follow this result up with a second SRT (their experiment 3).

In contrast to the high vowels in the /numi/ word pair, other Portuguese vowels exhibit reduction when unstressed, and the authors' follow-up experiment aimed to test the relative contribution of vowel quality as a cue to stress position. The words used for the prosodic contrast in this task were /ˈnemi/ ~ /neˈmi/ (with /e/ realised as [e] when stressed and [ɨ] when unstressed). The words for the segmental contrast were the same as in the previous experiment. The average percentage correct on the segmental contrast, as expected, was the same as in the first SRT, namely around 50%. The average percentage correct on the prosodic contrast, however, was much higher than in the previous SRT: 45% for the double enhancement type stimuli, and 25% for the single enhancement type. Clearly, response accuracy on the prosodic SRT improves when vowel quality differences are present, which is interpreted by the authors as indicating that vowel quality is a necessary cue to reliably encode stress position in Portuguese, whereas the presence of double enhancement as opposed to single enhancement plays a secondary role. One observation however remains to be explained: Portuguese participants still make about 75% errors on the single enhancement type stimuli in which vowel quality cues are present. While this is a clear improvement with respect to the 91% errors (9% correct) without the vowel quality cue, a 75% error rate still seems worse than the ca. 50% error rate on the phonemic contrast. This difference is also not explained by the authors.[3] It seems, therefore, that it cannot be concluded from Correia et al. (2015) that vowel quality is a sufficient cue for reliable stress encoding in Portuguese participants. Instead, it is safe to say that vowel quality in Portuguese is an *important* cue to stress, but the same can still be said for the presence of postlexical enhancement.

These results once more highlight two important points that have perhaps not received due attention in the literature. The first is that the use of the term 'stress' in

[3] I have so far refrained from reporting any statistics here since they might be problematic to interpret: The authors used ANOVAs on non-normally distributed data that were not transformed.

'stress deafness' might, to some, rather misleadingly suggest that the relevant participants are deaf to a *lexical stress* prominence asymmetry. This is clearly not the case, since, all else being equal, the presence versus absence of *postlexical prominence* can account for a considerable part of the stress deafness effect: Single enhancement type stimuli result in considerably worse accuracy on SRTs than double enhancement type stimuli, as the Portuguese results showed. Portuguese participants were in fact rather deaf to correlates of lexical stress alone: 25% correct scores on stimuli that exhibit only correlates of lexical stress (duration and vowel quality) is rather poor. The second point to take away is the importance of considering the native language's cues to lexical stress in relation to the specific acoustic–phonetic properties of the stress contrast participants are presented with. Portuguese listeners could be expected to perform worse on a Spanish lexical stress contrast (which is not cued by vowel quality) than on a Germanic stress contrast (which is cued by considerable vowel reduction in unstressed syllables). Past studies have not always kept these points in mind.

In the following, I will discuss some of the past work and the explanations that have been brought forward for 'stress deafness'.

8.1.1.3 Explanations for stress deafness

Over the years, many (partially overlapping) explanations have been brought forward as to why stress deafness occurs in some listener populations and not others. All of them appeal to properties of the phonology of the native language, and the most important ones are reviewed here briefly:

- The status of lexical stress as fixed or contrastive. This explanation is given in Peperkamp & Dupoux (2002) to account for the fact that Hungarian and Finnish subjects exhibit stress deafness, in contrast to Spanish speakers. Stress is fixed in Hungarian and Finnish, since it always occurs on the initial syllable of a word, whereas the position of the stressed syllable in Spanish varies depending on the word, so that it may function contrastively. Stress deafness is expected to occur in participants with native languages that do not require the representation of stress in the lexicon (such as Finnish and Hungarian).

- The status of lexical stress as predictable. This explanation overlaps with the previous one. In Peperkamp, Vendelin & Dupoux (2010), it is argued that *southern* French listeners exhibit stress deafness due to the fact that they are native speakers of a language with predictable stress (in this variety, stress is presumed to be final, but penultimate in the case of a final schwa). According to this explanation, their stress deafness could be explained in terms of the absence of a need to represent stress in the mental lexicon. An issue with this explanation concerns the general consensus that lexical stress in French (whichever Hexagonal variety) does not exist, in the sense that prominence asymmetries are not lexically assigned, but postlexically. See also Section 2.3. Better examples of languages with predictable *lexical* stress would be most varieties of Arabic, in which the stressed position can be derived by rule based on syllable weight. None of the stress deafness studies so far have tested uncontroversial predictable stress languages.

- Statistical tendencies towards stress position and the presence of exceptions to a stress rule. Dupoux & Peperkamp (2002), Peperkamp & Dupoux (2002) and Peperkamp, Vendelin & Dupoux (2010) appeal to this explanation in their discussion of Polish listeners' behaviour. Polish stress is mostly, but not entirely predictable due to it having fixed penultimate stress with a number of exceptions. The very presence of these exceptions supposedly accounts for the fact that Polish listeners are not stress deaf. The authors argue that due to the exceptions, stress in Polish is unpredictable enough for native speakers to include stress in their lexical representations of words, and therefore they would not exhibit stress deafness.

- The domain of 'stress' assignment. If the native language assigns 'stress' postlexically (or in my use of the term: If lexical stress is absent), listeners will exhibit stress deafness. In later papers, Dupoux, Peperkamp and colleagues (e.g. Dupoux et al. 2008) use this explanation to explain findings with regard to French listeners. The 'domain of stress' explanation is also used by Rahmani, Rietveld & Gussenhoven (2015) and accounts for the stress deafness effect they find for Persian, Indonesian and French listeners as opposed to the absence of the effect in Dutch and Japanese participants.[4]

The common denominator of the (partly) successful explanations is the idea that a lack of a grammaticalised contrast for 'stress' in the native language leads to stress deafness. It is important to note, however, that the absence of a lexical stress contrast in a language may have various underlying explanations, with different phonological properties leading to the same surface lack of the contrast. This part of the explanation for stress deafness has so far not been made explicit. In what follows, I will briefly sketch three different 'reasons' why languages may lack a surface lexical stress contrast.[5]

Firstly, languages like Finnish and Hungarian, whose speakers exhibit stress deafness in some of the earlier experiments, are commonly considered to have lexical stress. In these languages, the issue seems to relate to the fact that lexical stress is fixed (and thus by definition not contrastive).

Secondly, languages like French and Indonesian, whose speakers also exhibit stress deafness, lack contrastive lexical stress in a different way. In this case the absence of contrast is due to the simple fact that lexical stress is absent altogether.

Thirdly, native speakers of Japanese were found not to be stress deaf (Rahmani, Rietveld & Gussenhoven 2015). Several varieties of Japanese are commonly considered to lack lexical stress in the sense that lexical prominence asymmetries are not of the stress, but of the lexical pitch accent type. Thus, Japanese lexical phonology does encode lexical prominence asymmetries.

[4] Japanese has lexical prominence asymmetries in the form of lexical pitch accent rather than lexical stress.

[5] Presumably, more reasons, such as the presence of lexical tone, can also be included, but no work has so far been conducted to test speakers of lexical tone languages.

In sum, it appears to be the presence of contrastive prominence asymmetries at the word level, then, which seems to be the property that prevents 'stress' deafness. I suggest to use 'lexical prominence asymmetry' as the term that best represents the phonological property whose absence results in stress deafness. It should not be 'lexical stress' because usage of the term 'stress' would imply that Japanese listeners should be stress deaf. The property should also not be called 'lexical *prosodic* markings' (as in Rahmani, Rietveld & Gussenhoven 2015: 5). Findings by Correia et al. (2015) showed that participants can, to some extent, appeal to representations of lexical prominence asymmetries even if prosodic cues are missing from the acoustic input. Finally, the cause of Hungarian and Finnish listeners' deafness is not entirely clear, but the experiments that were performed with these listeners did use different stimuli from the later experiments and thus results are not entirely compatible.

8.1.2 Research question

The preceding discussion of stress deafness has highlighted that native speakers of languages that lack lexical stress are expected to exhibit lexical prominence deafness.

This thesis so far has accumulated evidence from production suggesting that lexical stress is absent in TB and MA. This raises questions about the perceptual status of lexical prominence asymmetries for speakers of these languages. The main question that this chapter aims to answer, therefore, is: Do native speakers of TB and MA exhibit a perceptual insensitivity to acoustic prominence contrasts?

A positive answer to this question would lend support to the idea that the languages lack inherent lexical prominence asymmetries, in this case lexical stress.[6] If listeners do not exhibit prominence deafness, this would be suggestive of the possibility that the languages do in fact have some sort of prominence marking for individual words in the lexicon.

The question is tested empirically by means of an experiment that consisted of two SRTs (one on a segmental and one on a prosodic contrast), replicating the design of and using the exact same stimuli as Rahmani, Rietveld & Gussenhoven (2015).

8.1.3 Overview of chapter

Section 8.2 will sketch the methodology, starting with a detailed description of participant selection. The following section, Section 8.3, will report the results, and the discussion after that, Section 8.4, will compare the present results in detail with the study it replicates (Rahmani, Rietveld & Gussenhoven 2015). It will also deal with the unexpected effect of one of the target stimuli. A brief summary and conclusion in Section 8.5 will round this chapter off.

[6] Although it does not, of course, exclude other possible causes for a stress deafness effect.

8.2 Methodology

8.2.1 Participants

In order to test effects of prominence deafness as a function of a given native language, participants are ideally monolingual speakers of this language. This is an impossible task when recruiting among university students in Morocco, so a number of measures was taken to arrive at target groups of participants that would, despite not being monolingual, still reflect perception effects that could be attributed to a single native language.

In addition to the diglossic situation that characterises present-day Morocco, with Berber varieties being used alongside Moroccan Arabic, there was the added complication of foreign languages learnt in school. These include French, English, and Modern Standard Arabic, the latter of which Moroccans have semi-regular exposure to on radio and TV, and in written form for the purpose of official communications. Among these three languages, only familiarity with French was deemed inconsequential, given the past results on stress deafness in native French speakers (e.g. Dupoux, Peperkamp & Sebastián-Gallés 2001; Dupoux et al. 2008; Peperkamp, Vendelin & Dupoux 2010). On the other hand, familiarity with Modern Standard Arabic and English, as languages that do have lexical stress, is potentially more problematic. Care was therefore taken to select participants with minimal proficiency and exposure to these two (and other stress) languages.

Two target groups of participants were recruited for the present study: One group of native speakers of Tashlhiyt Berber and one group of native speakers of Moroccan Arabic. While it was possible to find MA natives with zero proficiency in Berber, the reverse was not possible. Care was taken therefore to find those TB speakers that almost uniquely spoke TB and very little MA. To this end, prior to participation all participants filled out a questionnaire which included questions about their proficiency in various languages as well as the regularity with which they spoke these. No participants were recorded who did not meet the criteria of clear dominance in either Tashlhiyt Berber or Moroccan Arabic respectively. Participants also had to have no more than intermediate proficiency in any language with lexical stress (these were English, Spanish and German; Participants all understood MSA but did not speak it often).

All participants were recruited and recorded at the universities at which the experiments took place (Université Ibn Tofail in Kenitra for the MA group and Université Ibn Zohr in Agadir for the TB group).

In the TB group (N = 39) none of the participants spoke MA as a first language (age of onset for learning MA was six to eight) and all participants primarily spoke Tashlhiyt with their family, their friends, and at university or work. Participants' self-reported use of Moroccan Arabic typically involved around 20% of their social interactions. Most of the participants in this group were either students in the Département des études amazighes, where the main language of instruction is Berber, or, in case of a few graduates, primary school teachers of Standard Berber (see also Section 1.2). For the MA group (N = 37), participants were selected from the French and physics departments, where the main language of education is French (as opposed to MSA or English).

In both groups, a number of participants had to be excluded, resulting in a grand total of 62 participants, with 31 in each group. See Section 8.2.3 for the reasons why these were excluded prior to analysis. Mean age in the resulting Tashliyt-speaking group was 22 (range: 18 to 33), with 13 female and 18 male participants. Mean age in the Moroccan Arabic group was 20 (range: 17 to 36), with 16 female and 15 male participants.

8.2.2 Stimuli

The stimuli used in the experiment are the exact same ones as those used in the experiment by Rahmani, Rietveld & Gussenhoven (2015).[7]

The experiment consisted of two parts, a segmental and a prosodic part, which were separated by a voluntary break. The order of presentation was counterbalanced across participants, see Procedure below. The segmental part tested listeners' ability to discern the contrast in the segmental minimal pair /ˈmuku/ ~ /ˈmunu/. These are non-words in both Tashlhiyt Berber and Moroccan Arabic. The prosodic part tested listeners' ability to discern the contrast in the prosodic minimal pair /ˈnumi/ ~ /nuˈmi/. The word /numi/ is a non-word in both Tashlhiyt Berber and Moroccan Arabic. Several phonetic variants of these four words were used in the experiment: 12 per word as spoken three times each by a female and male speaker of Dutch, and a female and male speaker of Persian. The tokens were time-compressed so that all stimuli had a comparable duration of around 450 ms. The word "OK" which concluded every sequence of target words (see 8.1.1.1) was spoken by a different female speaker.

The acoustic differences between the members of the prosodic minimal pair /ˈnumi/ ~ /nuˈmi/ were discussed in Section 8.1.1.2 and were visualised for one example pair in Figure 8.2.

The exact order of words as used in the sequences were the same as in Rahmani, Rietveld & Gussenhoven (2015), i.e. there were five different sequences for each sequence length:

- Sequence length 3: 112, 121, 122, 211, 212
- Sequence length 4: 1211, 1221, 2112, 2122, 2212
- Sequence length 5: 12112, 12122, 12212, 21211, 21221

Each sequence was made up phonetic variants from a single speaker, and no phonetic token occurred more than once in each sequence. Sequences such as 1211, with three tokens of the same word, thus exhausted the three phonetic variants of that word as uttered by a single speaker. Each of the above 15 sequences occurred twice in each SRT: Once as instantiated by tokens from one of the two Dutch speakers and once by tokens from one of the two Persian speakers. The total of 30 test trials (per SRT) was

[7] I am very grateful to the authors for letting me use their stimuli.

thus made up of 3 sequence lengths * 5 sequences * 2 stimulus languages. The total number of responses was 30 responses * 31 participants * 2 participant groups * 2 SRTs (i.e. prosodic/segmental) = 3720.

8.2.3 Procedure

The two parts of the experiment, the prosodic and the segmental part, had the same structure, with four subparts. In the segmental part, participants had to perform discrimination on the stimulus pair with the segmental contrast /ˈmuku/ ~/ˈmunu/. In the prosodic part, participants did the same for the stimulus pair with the prosodic contrast /ˈnumi/ ~/nuˈmi/. The four phases were the following, with 4. being the main experimental task, or SRT test phase:

1. Word presentation phase
2. Word identification phase
3. Sequence recall practice phase
4. Sequence recall test phase

In phase 1, the word presentation phase, participants were presented first with all 12 phonetic tokens of a member of a contrasting word pair (upon pressing a designated key, ‘1’), and then with all 12 phonetic tokens of the other member of the pair (upon pressing the other designated key, ‘2’). After this, participants could press either key as often as they wanted in order to listen to single instances of the words. They could choose to continue to the next phase whenever they felt they had learned the key-to-word association.

In phase 2, the word identification phase, participants had to achieve a number of correct responses before they could proceed to the next phase. They were presented with randomly selected tokens of the 24 phonetic variants for the contrast. For each token they had to press the matching key, after which they received feedback as to whether their response was correct or not. Only once a participant had given eight correct responses in a row were they allowed to move on to the next phase. This proved to be very difficult for some participants in the case of the prosodic contrast. The average number of tokens participants listened to in order to get eight correct was 34 for the MA group, and 41 for the TB group. In order to get eight correct on the segmental contrast, these groups required only 14 and 10 tries, respectively. Similarly to Rahmani, Rietveld & Gussenhoven (2015), participants who needed more than 150 tokens before reaching eight correct in a row were excluded (six in the MA group, eight in the TB group).

Phase 3, the SRT practice phase (‘warm-up block’ in Rahmani, Rietveld & Gussenhoven 2015), served to train participants in doing the recall task. Participants heard eight sequences of two words with each of the four possible sequences (i.e. 11, 12, 21 and 22) occurring once in each stimulus language (Dutch/Persian). Each sequence was followed by the word "OK" as spoken by a different female speaker. All words (target word 1, target word 2 and "OK") were separated from each other by a pause of 120 ms. Participants gave their response by entering the numbers (keys) in a dialogue box. They could only enter their response after the word OK, and they could check if they had entered their intended response. They pressed enter to confirm their response. A new sequence was presented following a 1500 ms. pause. Participants received feedback as to whether their response was correct or not, and trials that were responded to incorrectly were presented again until the correct response was given.

Phase 4 was the main experimental SRT and used the same procedure as the practice SRT. Sequence lengths in this phase however involved three, four or five words, and participants received no feedback on the accuracy of their response. Sequences were randomly selected from the pool of 30 trials, 15 of which were Dutch speaker stimuli and 15 Persian speaker stimuli. Phonetic variants for each target word were selected randomly among the realisations from a single speaker. The order in which the prosodic and segmental task were performed was counterbalanced: Half of the participants performed the prosodic part first, followed by the segmental part, and the other half performed the segmental part before the prosodic part.

The perception experiment took place in subsequent weeks in November 2017 in quiet rooms at the Université Ibn Tofail in Kenitra for the MA group, and at the Université Ibn Zohr in Agadir for the TB group. The experiment was run with *E-Prime 3.0* (2017) on a laptop computer, and with headphones to listen to the sound. Pre-experiment instructions were given orally in TB or MA, and on screen during the experiment in French. The average duration to complete the whole experiment was comparable in both groups and was around 30 minutes (including an optional break).

8.2.4 Analysis

Responses were classified as either correct or incorrect. Incorrect responses included reversals such as 121 for a target response 212.[8]

Statistical analysis was performed with binomial Generalized Linear Mixed Models (GLMMs) in R (R Core Team 2016) with the package *lme4* (Bates et al. 2015). The accuracy of individual responses was modelled as a function of the predictors GROUP (MA/TB), CONTRAST (segmental/prosodic), SEQUENCE LENGTH (3/4/5) and STIMULUS SPEAKER (DutchF/DutchM/PersianF/PersianM). Specific interactions between main effects were included, as well as random slopes allowing for interactions of main effects with participant. The R syntax for each model is given in a footnote. The inclusion of a

[8]The decision to exclude reversals along with incorrect responses was motivated by the relative number of reversals per participant. Three participants (two MA and one TB) produced more reversals than correct responses on the prosodic SRT (eight reversals versus only six correct (and the rest incorrect), and seven versus four, and two versus zero, respectively). Six participants produced more reversals than correct responses on the segmental SRT (differences of one, two, two, two, seven and ten). It is not clear from these numbers whether some cases involve the incorrect, but systematic, association of words with keys, as suggested in some of the earlier studies. Exclusion seemed to be the the most sensible treatment of these patterns, and mirrors what was done in most previous studies.

main effect of STIMULUS SPEAKER in the model followed from participants' comments, as well as my own impression, that the difficulty of some of the contrasts depended on the speaker.

The binomial models used here are rather different from analysis methods in previous work, which are typically ANOVAs on averaged accuracy per participant per SRT, expressed in percentages. There are some disadvantages to this approach, including the fact that such data is bounded (0-100%). This means that the variance is not constant across the scale and by consequence, that ANOVAs are a suboptimal, if not inappropriate method of analysis. Moreover, the aggregation of data achieved by averaging error scores per participant, rather than using raw responses (correct/incorrect), leads to an increased Type I error rate, i.e. an increased likelihood of reporting false positives. These issues occur in all previous studies, with the exception of Rahmani, Rietveld & Gussenhoven (2015), who performed arcsine transformation to the data before running ANOVAs, eliminating the equal variance issue (although not the aggregation issue).

In any case, the comparison between the present and previous results should be made carefully. The approach taken here is to compare the present data with the *raw* data from Rahmani, Rietveld & Gussenhoven (2015) (after a reconstruction of single responses based on the aggregated scores per participant). The model that serves the direct comparison between the TB/MA and the Rahmani, Rietveld & Gussenhoven data (the 'combined model'), has the same structure as the model run on the present data only ('the TB/MA model'), with one difference: The combined model has the predictor STIMULUS LANGUAGE (Dutch/Persian) instead of STIMULUS SPEAKER, since the exact speaker for each trial is not available for the other dataset. Further details will be discussed in Section 8.4.1.

In order to compare accuracy on the *prosodic* SRT, many previous studies check whether the behaviour of groups is similar on the *segmental* SRT. Presumably, the segmental contrast is equally easy for different participant groups as long as it reflects a phonemic difference in the relevant native languages. If scores on the segmental SRT are considerably lower for one group this might form an indication that groups are not comparable in terms of working memory or other cognitive factors. In practice, however, there are many reasons why the contrast might still more difficult for one group than the other, relating to e.g. linguistic factors such as frequency of occurrence of the relevant phonemes in different native languages. Previous studies (Peperkamp, Vendelin & Dupoux 2010; Hellmuth, Muradás-Taylor & White n.d.) have used a so-called stress deafness index, defined as the error rate in the prosodic SRT minus the error rate in the segmental SRT to control for participant variability in terms of such general performance factors. The present way of accounting for participant- and group-specific differences is by including random slopes for the interaction of the main effect of CONTRAST with individual participants.

8.3 Results

Mean scores for both participant groups, averaged across the four stimulus speakers, are shown in Figure 8.3.

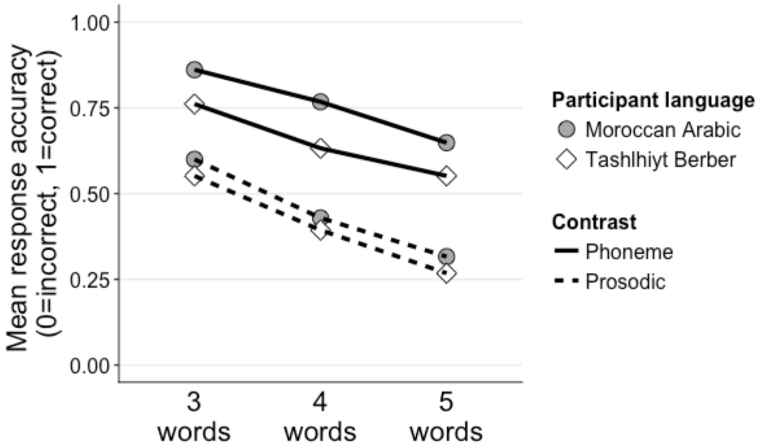

Figure 8.3: Mean accuracy per participant group for both contrasts and for each sequence length.

Mean accuracy is, on average, higher on the segmental SRT for both groups. The difference between the participant groups seems minimal for the prosodic SRT, but on the segmental SRT the TB group scores somewhat lower than the MA group. This difference is however not statistically significant, see below. As was discussed in the previous section, a difference on the segmental SRT precludes the reliable use of a stress deafness index for which the segmental SRT serves as a baseline. Since Rahmani, Rietveld & Gussenhoven (2015) also did not make use of difference scores, I decided to base the comparison between the present and their study on the raw scores on the prosodic SRT only. Most importantly, Tashlhiyt participants score well over half of their responses correct on the segmental SRT (which is better than pure chance performance given that a mistake in any of the three to five items in a sequence results in an incorrect response for the entire sequence). Finally, cognitive differences between groups are highly unlikely given the recruitment among university students in both cases.

I will turn now to the statistical comparison of accuracy on both the SRTs. A binomial model was run with the aforementioned main effects, interactions between GROUP and CONTRAST, and STIMULUS SPEAKER and CONTRAST, and random slopes for interactions of CONTRAST and STIMULUS SPEAKER with participants.[9] The 'mixed' function from the *afex* R package (Singmann et al. 2015) and LRTs were used to obtain main effects in the case of interactions, and the *lsmeans* package (Lenth & Hervé 2015) was used to perform multiple comparisons (Tukey), yielding the following results:

[9] score ~ Group + Contrast + StimulusSpeaker + SequenceLength + Contrast:Group + Contrast:StimulusSpeaker + (0 + Contrast|Participant) + (0 + Contrast|StimulusSpeaker)

- A main effect of CONTRAST, in the sense that overall accuracy was lower on the prosodic SRT: ($\chi^2 = 38.21$, $p < 0.001$);
- A main effect of STIMULUS SPEAKER (but see interaction below): ($\chi^2 = 23.04$, $p < 0.001$);
- A main effect of SEQUENCE LENGTH: ($\chi^2 = 228.1$, $p < 0.001$). Post-hoc comparisons indicate that all pairwise comparisons are significantly different at $p < 0.001$ (i.e. accuracy is lower for five words than for four words, which in turn is lower than for three words);
- An interaction of CONTRAST with STIMULUS SPEAKER: ($\chi^2 = 38.97$, $p < 0.001$). Post-hoc comparisons show that the Dutch female speaker's stimuli result in significantly lower accuracy than all other three speakers' stimuli in the prosodic SRT ($p < 0.001$). There is an additional difference concerning the two Persian speakers in the segmental SRT, where the Persian male speaker's stimuli cause lower accuracy than the female speaker's ($p < 0.01$).

Importantly, there is no effect of GROUP, that is the scores of TB and MA participants do not differ overall. There was also no interaction of CONTRAST with GROUP, that is, TB participants do not perform significantly worse on the segmental SRT than the MA participants ($\chi^2 = 1.29$, $p = 0.26$).

Since the main effect of STIMULUS SPEAKER was also involved in an interaction this requires some further scrutiny. Post-hoc comparisons showed that of the four speakers, the Dutch female speaker in the prosodic SRT carries most of this effect, for both participant groups. The difference is visualised in terms of model predictions and confidence intervals (on the prosodic SRT) for the different speakers in Figure 8.4.

8.4 Discussion

One of the crucial results so far is that there are no differences between the TB and MA participant groups. That is, both groups exhibit the same degree of 'deafness' to lexical prominence. Exactly to what degree this term applies will be discussed in the following. I will compare the present results to earlier work on stress deafness (Section 8.4.1), and specifically to Rahmani, Rietveld & Gussenhoven (2015), whose experiment was replicated here and therefore provides comparable reference groups. Following this comparison I will sketch the implications for the interpretation of lexical prominence structure in TB and MA (section 8.4.2), and finally I will try to explain the effect of STIMULUS TYPE by comparing the phonetic properties of the stimuli that cause this effect (Section 8.4.3).

8.4.1 Comparison with earlier work

The present results on the main effect of CONTRAST are in line with those by Rahmani, Rietveld & Gussenhoven (2015), who found that the prosodic SRT was on average more

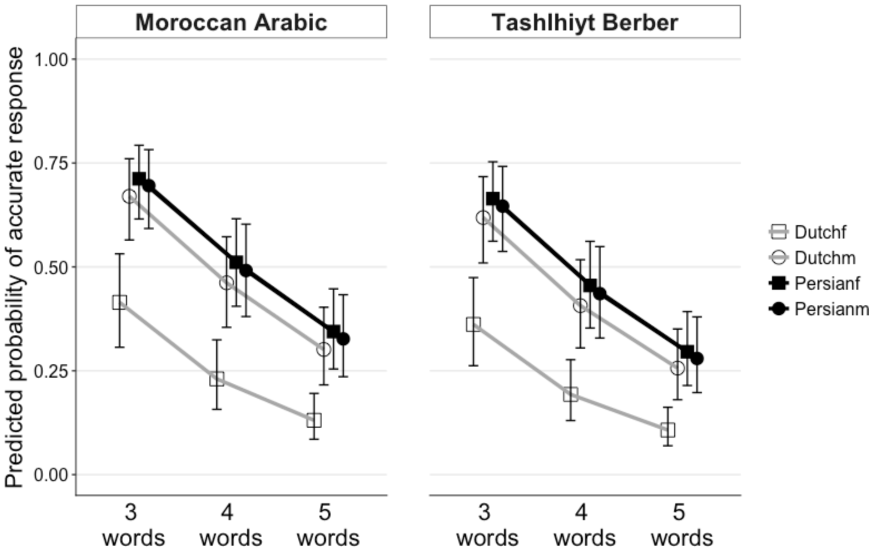

Figure 8.4: Predicted probability of an accurate response, with 95% confidence interval (based on model fixed effects only), per stimulus length and speaker group, for the prosodic SRT. The Dutch female speaker's stimuli result in reliably lower accuracy than the other three speakers'.

difficult, with participants scoring fewer trials correct than on the segmental SRT. In their analysis, this effect held irrespective of group, that is, stress deaf and non-stress deaf groups performed alike in having worse scores on the prosodic SRT (i.e. there was a significant main effect of CONTRAST in their ANOVA, in addition to an interaction which meant that all groups performed even worse on the prosodic SRT than others). The groups who performed relatively badly on the prosodic SRT were consequently considered the 'stress deaf' groups: Native speakers of French, Indonesian and Persian.

These observations are somewhat different to results by Dupoux et al. (2008), who found that Spanish participants, representing a non-stress deaf population, performed a little better on the prosodic contrast (scoring 80% correct) than the phoneme contrast (75% correct). The phonemic contrast in this study concerned /t/ and /k/ in the segmental minimal pair /fiku/ ~/fitu/, a segmental-phonological contrast only of place of articulation, which is arguably more difficult than the contrast between the voiceless stop and the (voiced) nasal in present /muku/ ~/munu/, a difference in both manner and place of articulation, as well as voicing. In any case, despite the small differences in means, the Spanish scores did not differ significantly as a function of contrast (prosodic/segmental) in their model.

The next important question is how, in absolute terms, the TB and MA groups compare with 'stress deaf' and 'non-stress deaf' participants. Since the present study was designed to be comparable with the Rahmani, Rietveld & Gussenhoven (2015) study for this very purpose, I will consider their raw scores in relation to the present ones. A similar model to the above was run on this combined dataset (MA/TB plus Dutch/Japanese/Indonesian/French/Persian), although there are two important differences: i) The predictor STIMULUS SPEAKER was changed to STIMULUS LANGUAGE (since STIMULUS SPEAKER is not available for the Rahmani dataset) and ii) an interaction is added for the main effects STIMULUS LANGUAGE and PARTICIPANT LANGUAGE.[10] This latter decision follows from the hypothesis that there might be performance differences as a function of whether the stimulus contrast is in one's native language. This was the case for neither TB and MA groups with Dutch and Persian stimuli, but it becomes relevant when considering that two of the participant groups in the combined dataset (the Dutch and Persian participant groups) each give half their responses to stimuli that exhibit native acoustic properties.

Figure 8.5 shows the model's predicted accuracy for all seven participant groups per contrast and sequence length (collapsing across both stimulus languages).

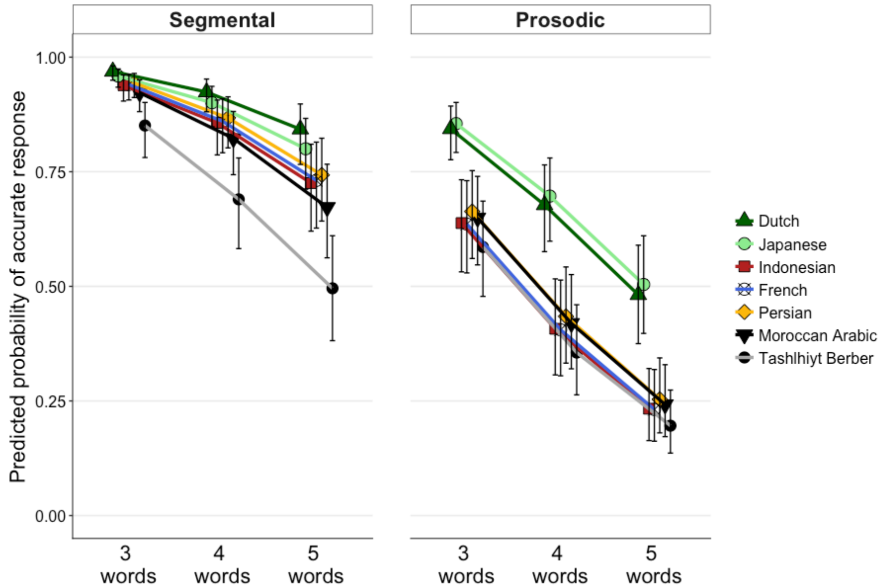

Figure 8.5: Predicted probability (dots) and 95% confidence intervals of accurate response, separated by contrast and sequence length (collapsed across stimulus language Dutch/Persian), for all native participant groups tested in Rahmani, Rietveld & Gussenhoven (2015) (Dutch, Japanese, Indonesian, French, Persian) and the present experiment (Tashlhiyt Berber, Moroccan Arabic).

It is clear that while there are some differences among these groups on the segmental contrast, notably with Tashlhiyt participants scoring lower than the other participants, the great divide concerns, as expected, the prosodic contrast. On the prosodic contrast, the TB and MA participant groups score no different from the Indonesian, Persian and

[10] score ~ ParticipantLanguage + Contrast + StimulusLanguage + StimulusLength + ParticipantLanguage:Contrast + StimulusLanguage:Contrast + StimulusLanguage:ParticipantLanguage + (0 + Contrast|Participant) + (0 + StimulusLanguage|Participant)

French participant groups, as suggested by the overlap in these groups' confidence intervals. In addition, at no sequence length is there any overlap in confidence intervals between the 'stress-deaf' groups and the 'non-stress deaf' groups, confirming the original finding of a similarly clear-cut split by Rahmani, Rietveld & Gussenhoven (2015) (despite the differences in statistical approach).

In sum, it is clear that TB and MA participants exhibit 'lexical prominence deafness'. The present results confirm many previous findings in showing how robust the differences are between groups, as a function of their native language, in their ability to categorise a prosodic prominence contrast.

8.4.2 Implications for analyses of word prosody in TB and MA

An obvious explanation for the above results would be that the word prosody of TB and MA influences participants' ability to carry out the prosodic SRT, causing an effect whereby they score much lower than some other participant groups. This explanation would appeal to the absence of lexical prominence asymmetries (in any form) in these languages. As discussed in Chapters 1, 3 and 4, there is currently no evidence for the presence of lexical stress, the most likely prominence asymmetry in the phonology of these languages (as opposed to lexical tone or lexical pitch accent). The fact that TB and MA native speakers struggle with the encoding of prominence asymmetries, as required during an SRT, is compatible with an interpretation along the lines of the absence of stress. In this respect they are similar to speakers of French, Indonesian and Persian, other languages known to lack lexical prominence asymmetries.

8.4.3 Phonetic properties of stimuli

So far, the exact nature of the prominence contrast exhibited by test stimuli has received little attention, with the exception of Correia et al. (2015). As mentioned above, Rahmani, Rietveld & Gussenhoven (2015) did not find an effect of stimulus language on overall accuracy. The present study, in contrast, found an interaction between the stimulus type with CONTRAST, but this interaction did not concern STIMULUS LANGUAGE per se, but rather an individual speaker's stimuli, i.e. an interaction with STIMULUS SPEAKER: The effect specifically concerned the Dutch female speaker's renderings of the prosodic stimuli, which were responded to with significantly lower accuracy than all other speakers' prosodic stimuli. In order to understand this effect, Figure 8.6 shows examples of the prosodic stimuli /ˈnumi/ ~/nuˈmi/ for all four speakers.[11]

There is a striking difference between the Dutch female speaker's realisations and the ones of the other speakers, related to the durational differences between the two syllables. While most speakers produce a difference in duration of the first syllable as a function of stress status, no such clear asymmetry is seen for the Dutch female speaker. Syllable durations per speaker and per token are given in Table 8.1.

[11] The male Dutch speaker's stimuli shown here are the same as those in Figure 8.2.

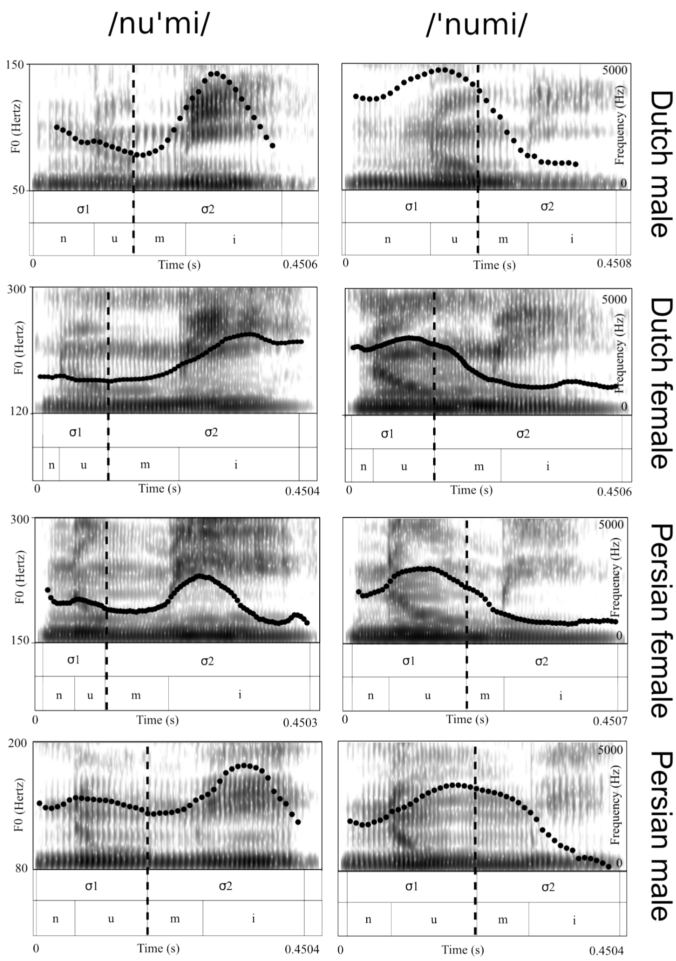

Figure 8.6: Spectrogram and F0 contour for /nuˈmi/ (left) and /ˈnumi/ (right). Examples show one out of three phonetic variants for each word by each of the four speakers.

The Dutch female speaker's initial syllable is on average 131 ms. long when it is unstressed, and 139 ms. long when stressed, a negligible difference of 8 ms. which means that for her stimuli, syllable duration differences cannot serve as a cue that helps listeners reliably distinguish between /ˈnumi/ and /nuˈmi/. The Dutch male speaker, in contrast, produces a stress-induced difference of 30 ms.[12] In fact, not only is duration missing as a reliable cue to stress position for the Dutch female speaker, it is potentially even a misleading cue, since initial stress /ˈnumi/ displays relative durational properties more appropriate for final stress.

[12] Rahmani, Rietveld & Gussenhoven (2015) report mean syllable duration per stimulus language and per stress contrast, but their values differ considerably from my own measurements. According to them, the Dutch initial syllable in /ˈnumi/ would be 215 ms. long (as opposed to my measured average of 167 ms.).

Token	/nuˈmi/		/ˈnumi/	
	DutchF	**DutchM**	**DutchF**	**DutchM**
1	104	157	130	208
2	147	191	156	218
3	142	148	131	160
speaker mean	131	165	139	195
language mean	148		167	

Table 8.1: Initial syllable duration (ms.) of target words

In addition to the durational differences, the Dutch female speaker also produces somewhat diverging F0 patterns for /nuˈmi/, also visible in Figure 8.6. All her three phonetic variants of /nuˈmi/ are like the one pictured, terminating high as opposed to falling, as is the case for all other speakers. It is not obvious that this F0 pattern contributes to participants' difficulties with these stimuli. While the non-falling terminal intonation certainly differs from the other speakers' realisations of word-final stress, it is still a pattern exclusively used for /nuˈmi/ (not /ˈnumi/) and therefore could presumably serve as a cue to final stress.

The question remains why the potentially conflicting durational cues (and/or final non-falling terminal intonation for tokens with final stress) would negatively influence TB and MA speakers' scores, while other listener groups do not appear to be influenced by it. Stimulus language, after all, was not a significant effect in Rahmani, Rietveld & Gussenhoven (2015). Part of the explanation may relate to the fact that Rahmani, Rietveld & Gussenhoven took STIMULUS LANGUAGE, not STIMULUS SPEAKER as a predictor, thereby collapsing scores across trials with stimuli from the Dutch male and female speaker. It is possible that the male speaker's accuracy scores were high enough to cancel out any potential effect relating to the female speaker.

In order to explore the possibility that there is no effect of STIMULUS LANGUAGE on the prosodic contrast when collapsing across speakers, yet another GLMM was run on the prosodic SRT results from all seven languages combined.[13]

Figure 8.7 visualises the effect of STIMULUS LANGUAGE (Persian/Dutch) for each language, on the prosodic contrast (note that the predictions for TB/MA are marginally different from the ones in Figure 8.4, which were based on a different GLMM with STIMULUS SPEAKER as a predictor). The only participant groups for which there was no effect of STIMULUS LANGUAGE were Japanese ($z = -0.18$, $p = 0.85$) and Persian ($z = 1.231$, $p = 0.21$). The Dutch group performed marginally worse on the Dutch stimuli than the Persian stimuli ($z = -1.79$, $p = 0.07$). Indonesian and French groups performed better on the Dutch contrast, whereas the Moroccan Arabic and Tashlhiyt Berber groups performed better on the Persian contrast.

[13] score ~ ParticipantLanguage + Contrast + StimulusLanguage + StimulusLength + ParticipantLanguage:Contrast + StimulusLanguage:Contrast + StimulusLanguage:ParticipantLanguage + (0 + Contrast|Participant) + (0 + StimulusLanguage|Participant)

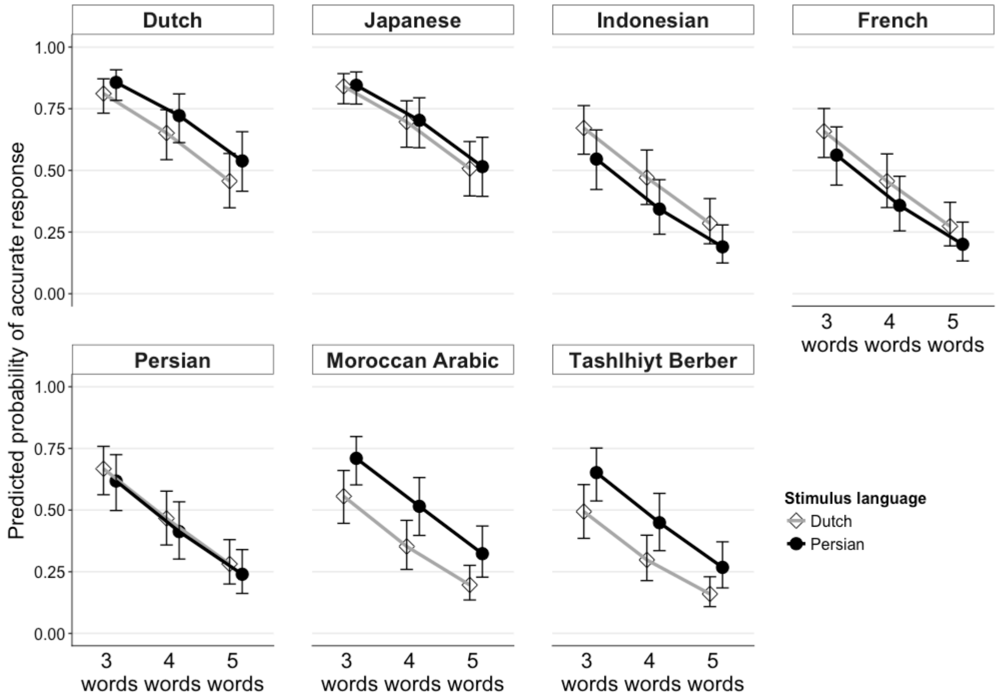

Figure 8.7: Predicted probability of an accurate response, per participant group and per stimulus language, on the prosodic SRT.

These results are not necessarily at odds with the lack of an effect of STIMULUS LANGUAGE reported in the original study, but they do highlight the importance of differences in statistical tools. For their five original participant groups the present model finds a clear effect of STIMULUS LANGUAGE only for the French and Indonesian groups (better performance on the Dutch stimuli), which will have been counteracted in the original model by tendencies in the opposite direction for Dutch and to a lesser extent, Japanese.

The present results indicate that there is no straightforward effect of the language in which stimuli are spoken *per se*. Rather, there are differences between participant groups in the relative difficulty of specific stimuli. This in turn suggests that the acoustic details of the prominence contrast, and speaker- or voice-related differences play an important role in determining response accuracy, at least across participant groups. It follows that, in general, it is crucial to ensure variability in test stimuli before making claims about language-specific effects (cf. Roettger et al. 2014 for different speaker voices).

At this point it is unclear if the female Dutch speaker's prosodic contrast causes a similar effect in different participant groups alike: The effect was observed for TB and MA groups because STIMULUS SPEAKER was logged, but it is possible that this specific speaker's effect also applied to (some of) the participant groups in Rahmani, Rietveld & Gussenhoven (2015). In the absence of this information, it can only be conjectured why the Dutch female speaker's realisation of prominence asymmetries should be so difficult for the TB and MA groups. Apart from the Dutch female speaker's stimuli, all stimuli exhibit relative acoustic enhancement of the stressed syllable in terms of the combined cues of duration, F0 and intensity, as well as some vowel quality differences (cf. Table 2 in Rahmani, Rietveld & Gussenhoven 2015). For the Dutch speaker, the durational cue is removed from this set, making the acoustic realisation of the contrast less salient. Under the assumption that native speakers of TB and MA cannot rely on a categorical–phonological representation of the contrast, but rather use the acoustics only (which is why they succeed in the training phase), it is likely that a reduction of the salience of the acoustic contrast significantly influences their performance. It should be pointed out that the prominence deafness effect is observed for the other three speakers as well, it is simply stronger for the Dutch female speaker.

Following this hypothesis about cue impoverishment, it would be worth finding out whether, in the presence of all other original cues, the absence of F0 cues to stress position would pose the same degree of challenge to TB and MA participants as the absence of durational cues. Given the results on highly variable F0 peak alignment in both languages in Chapters 6 and 7, the expectation would be that TB and MA speakers are not very sensitive to acoustic differences in terms of F0 movements. Perhaps, then, they would not find the task that much harder without F0 cues. Such a result would further strengthen the argument that the details of the acoustic cues involved in the contrast tested in the prosodic SRT are not trivial, especially in the case of speakers whose native languages do not provide them with the tools for categorical representation of lexical prominence.

8.5 Summary and conclusion

The results from the 'prominence deafness' experiment discussed in this chapter show first and foremost that native Tashlhiyt Berber and Moroccan Arabic speakers have low scores on a prosodic SRT, a short-term memory task involving a prosodic prominence contrast. Direct comparison of these results with results from an earlier study (Rahmani, Rietveld & Gussenhoven 2015) served to confirm that the scores of both TB and MA groups are compatible with an interpretation in terms of prominence deafness ('stress' deafness). Specifically, TB and MA participants behave no differently from French, Persian and Indonesian participants, who are considered to exhibit stress deafness by virtue of their native languages lacking lexically specified prominence asymmetries. Along the same lines, it was argued that prominence deafness in TB- and MA-speaking participants is compatible with a lack of lexical stress in their native languages.

In addition to these findings, a previously undiscussed effect of the acoustic correlates of the prominence contrast was observed. Among phonetically different tokens of the prosodic contrast, those in which the contrast was not cued in terms of duration were more difficult for TB and MA participants. This effect adds further weight to the argument that prominence deafness follows from participants' reliance on acoustic salience rather than on a categorical representation of prominence.

In conclusion, the results from the present experiment lend further credibility to claims that were brought forward in earlier parts of this thesis, namely that there is a lack of lexical prominence asymmetries in the form of lexical stress in both Tashlhiyt Berber and Moroccan Arabic.

Part V

Conclusion

9 Summary and general discussion

9.1 Summary of results

In this thesis I have reported results from various experiments that bear on the question of phonological prominence structure in Tashlhiyt Berber and Moroccan Arabic.

Part II was concerned with correlates of lexical prominence in both languages. The specific question asked here was whether, based on arguments relating to acoustic enhancement of syllables, TB and MA can be considered to lack lexical stress. For Tashlhiyt Berber (Chapter 3), the final syllable was not consistently enhanced relative to the penultimate syllable in disyllabic words. This result could not confirm an earlier claim that there is fixed final-position stress in Tashlhiyt. For Moroccan Arabic (Chapter 4), no consistent enhancement was found of syllables that are presumed stressed under a weight-sensitive view of stress in this language. This result provided no support for the claim that stress in MA targets either the penultimate or final syllable. In conclusion, the results of these experiments are compatible with the absence of stress in both languages, a possibility that has also been raised independently by several authors for each language.

Part III was concerned with the postlexical prosodic prominence structure of both languages, with specific focus on intonational prominence in interrogative phrases with a question word (qword). Qwords in both TB and MA were shown to attract the main prosodic prominence-marking event in the phrase: A rise–fall realised entirely or partly on the qword. In neither language did the rise–fall exhibit phonetic properties compatible with an interpretation along the lines of a pitch accent associating with a predetermined, lexically stressed syllable. Based on considerations of semantic-pragmatic prominence (focus) in qword interrogatives, and on the phonetic properties of the intonational event, it was argued that the rise–fall functions as a postlexical prominence-marking event in both cases. Unlike in languages that exhibit lexical prominence asymmetries, in which stressed syllables are the typical TBUs, the pitch event was interpreted as associating with a higher structural unit: The phonological domain of the qword. This aspect of postlexical prominence structure is compatible with, if not a direct corollary of, the absence of lexical prominence structure.

Part IV served to investigate the perception of prominence by native TB and MA speakers. Chapter 8 reported on a 'stress deafness' experiment in which the perceptual sensitivity of participants was tested with respect to a postlexical prominence contrast parasitic on a lexical prominence contrast. Native speakers of both TB and MA were shown to perform poorly on this task. Specifically, they performed significantly worse than speakers of languages that do have lexical prominence structure. At the same time their behaviour was similar to that of speakers of (other) languages that are known lack lexical prominence. Participants' perceptual behaviour therefore is one more result that provides support for the idea that lexical stress is absent in both TB and MA.

In the following, I will consider these converging results in light of the two secondary goals of this thesis, as mentioned in Chapter 1.1, in shedding further light on:

- The result of language contact between Tashlhiyt Berber and Moroccan Arabic in the (prosodic)-phonological domain
- The possible mappings between lexical and postlexical prominence structure and the theoretical implications of the present findings

9.2 Language contact: Prominence structure in TB and MA

As mentioned in Chapter 1, more than 1200 years of contact between Berber and Arabic in Morocco has resulted in linguistic convergence in several subdomains of phonology, including syllable structure and phoneme inventories. For Tashlhiyt Berber and Moroccan Arabic specifically, the results in this thesis suggest further similarities in terms of lexical and postlexical prominence structure, with both languages lacking stress and both exhibiting postlexical prominence events whose realisation exhibits a similar kind of variability.

A question of specific interest is whether the lack of stress in TB and MA can plausibly be explained by diachronic developments that favoured the loss of stress following some earlier stage of the languages in which stress did exist. If so, did the loss of stress arise independently in both languages or did it arise in one language first, spreading to the other language through contact? To answer these questions, I will briefly review what reference has been made to lexical prominence structure in other languages in the Arabic- and Berber-speaking world.

In the literature on Berber, firstly, there has been limited mention of stress. The main exception is Gussenhoven (2017) which includes an experimental investigation of Zwara Berber, spoken in Libya, which does have lexical stress. Like Tashlhiyt, Zwara belongs to Northern Berber, but is subclassified as Eastern Zenati, while Tashlhiyt is subclassified as Atlas (Simons & Fennig 2017). In contrast, other varieties of Northern Berber, which are all spoken in Morocco and Algeria, are "generally assumed" to lack stress (Kossmann 2012: 33).

Stress in varieties of Arabic other than MA has been reasonably well researched. It is therefore easy to observe that MA differs considerably in matters of stress from most other synchronic varieties of Arabic, which typically exhibit unequivocal lexical stress. Phonetic investigations have been conducted on several varieties, supporting native speakers' intuitions about stress positions, including Egyptian and Jordanian (Almbark, Bouchhioua & Hellmuth 2014), Lebanese (Chahal 2003), Tunisian (Bouchhioua 2008), and several more varieties based on ongoing work (Hellmuth & Almbark 2017). The literature on Metrical Stress Theory adds a number of varieties to this list, with at least nine synchronic varieties reported to have stress in Hayes (1995) alone. An overview of stress in Arabic is given in Watson (2011), including the observation that Modern Standard Arabic is considered to have stress, like its precursor Classical Arabic.[1]

Based on these facts about Berber and Arabic it seems plausible that the lack of stress in Moroccan Arabic is the product of language contact with (Tashlhiyt) Berber. Similar reasoning is found in Zellou (2010), which is concerned with the origin of consonant harmony in Moroccan Arabic, a feature not found in other varieties of Arabic but attested in Berber. Unfortunately, little is known about the phonological prominence structure of Proto-Berber, which was presumably spoken at some time after 1000 BCE (Kossmann 1999, 2012). This makes it unclear whether, and if so at what point, (some varieties of) Berber lost stress. If Proto-Berber was stressless, some varieties of Berber including Zwara must have developed it at some point. If Proto-Berber had stress, some varieties including Tashlhiyt must have lost it.

These open issues moreover raise the general question as to whether and how languages may 'lose' or 'develop' lexical stress, which has in fact rarely if ever been addressed.[2] French is uncontroversially considered a language that has 'lost' stress, as it is known to have evolved from Latin, the lexical phonology of which did include stress (cf. Jun & Fougeron 2002). In contrast, there are no specific reports of stress coming into existence from an originally stressless word prosodic system, although a general hypothesis about language change is brought forward in Hyman (1977), taken up again by Gordon (2014). According to this view of word stress, particularly in systems that have fixed penultimate stress, stress might arise as the result of a generalisation of phrase-level prosodic patterns. Neither author gives a very detailed description of such a process, but presumably refer to a process like the following: The starting point is a situation in which a language exhibits a consistent location of postlexical *pitch* (whether this should be interpreted as prominence- or edge-marking is not clear). It most likely takes the form of some sort of high pitch within a basic rising-falling intonational pattern, which seems to be observed in some form or another in every language.[3] When an intonational pattern like a rise–fall is realised on a short one-word IP, some syllable near the right edge, for example the penultimate, will be characterised by the pitch peak and/or considerable pitch movement. This is schematised in Figure 9.1:

[1]This claim should nevertheless be evaluated carefully due the fact that both MSA and Classical Arabic are not native languages and are to a large extent literary. When spoken they should be considered second language varieties that most likely exhibit prosodic features of the first language of the speaker, typically a 'colloquial' variety of Arabic.

[2]The changes in lexical phonological systems that lead to the development of lexical tone and pitch accent, as opposed to stress, are much better researched. See among many others: Coetzee, Beddor & Wissing (2014) and Coetzee et al. (2018) for tonogenesis in Afrikaans; Kingston (2005) for Athabaskan; Kirby (2014) for Khmer; Kang & Han (2013) for (renewed) tonogenesis in Korean.

[3]The privileged status of high pitch crosslinguistically has been observed in many places (e.g. Gussenhoven 2004; Ladd 2008) and is also reflected in the title of a special issue of *Journal of Phonetics*: "What's so special about H(igh)?" (Evans 2015).

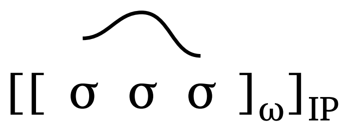

Figure 9.1: Schematised rising-falling intonational contour as typically realised on isolated (phonological) words that form an IP.

The resulting pitch pattern could then be generalised to words that are not produced in isolation, so that pitch prominence occurs in the penultimate position on each lexical word. Presumably, the pitch-prominent syllable becomes reinterpreted as reflecting fixed word-level prominence in that position, after which the pitch correlate might give way to other correlates of prominence and other lexical–phonological interactions that single this syllable out as prominent.

Some additional speculation can shed light on how exactly high pitch might diachronically lose its primacy in singling out the syllable that is becoming stressed. A plausible scenario is one in which high pitch involves increased articulatory effort, involving greater jaw and lip opening, which may in turn also yield greater loudness (cf. Żygis, Fuchs & Stoltmann 2017). It is also known that the presence of pitch movement tends to cause durational expansion of the syllable on which it is realised (e.g. 'accentual lengthening', see Cambier-Langeveld 2000). Once the point is reached at which this prominence is reanalysed as being positionally determined or due primarily to non-pitch properties, one could call it lexical stress. This last part of the process, in which stress becomes phonologised, is reminiscent of the idea that differences in the quality of vowels that may occur in stressed and unstressed syllables in Germanic languages are the phonologised result of stress-induced enhancement (e.g. stressed [iː, aː] versus unstressed [ɪ, ɑ] in Dutch, Gussenhoven 2004).

The above hypothesis about the phonologisation of stress involves a top-down explanation, with higher-level prosodic structure percolating down to lexical-level structure. The reverse, a bottom-up explanation, could presumably be invoked to explain the lack of stress in French. The fact that it is considered a language with fixed prominence in *phrase*-final position suggests that the lack of stress might have arisen as a result from a generalisation of what at some stage would have been fixed *word*-final stress to fixed position for prominence at the phrasal level.

At this point, not much is understood about how and why languages may come to lose word stress, how they might develop it, and to what extent the location of postlexical prosodic events, either in the form or prominence-marking or edge-marking, is involved in bringing about lexical prominence structure. It is clear, however, that the interaction between levels of prosodic structure plays some role, if not a crucial one, in determining the presence or absence of stress. The next section will discuss the synchronic correspondence between lexical and postlexical prominence in further detail.

For the two languages investigated here, the directionality of change is not entirely certain at this point. Did TB and MA lose stress or did they never develop it in the first place? With respect to the relationship between the two languages, however, it does seem likely that a prior absence of stress in TB has impacted on the lexical prosodic structure of MA, causing it to develop away from some earlier state in which it did have lexical stress, like other varieties of Arabic. Given the relative rarity of stressless languages and given the well-documented influence of Berber on Arabic, it seems likely that these similarities reflect linguistic convergence due to language contact, with Moroccan Arabic being influenced by Tashlhiyt Berber, rather than independent language-internal pathways resulting in comparable structures (see also Zellou 2010).

9.3 Phonological theory: Lexical and postlexical prominence structure

9.3.1 Different types of tonal association

Under the assumption that TB and MA indeed lack lexical prominence structure in the form of lexical stress, the results presented in Chapters 6 and 7 on the intonational properties of qwords raise the general question as to how the location of postlexical prominence is determined in the absence of lexical stress. In both languages, the prominence-marking pitch event on the qword in qword interrogatives was shown not to seek out a specific syllable to associate with, in contrast to what a prominence-marking pitch accent in a language with stress would do. I have so far refrained from using the attested tonal behaviour alone as either a clear *diagnostic* for the existence of stress or a clear *consequence* of the absence of stress, as there are reasons to consider it as being both. On the one hand, the variable alignment of the pitch event in question can be taken as providing support for the absence of stress (since the prominence-marking pitch event on the qword does not seek out a specific syllable to associate with). On the other hand, the variable alignment can also be considered to follow from the absence of stress (since there is no stress, there is no predetermined syllabic anchor for the F0 event). In the following, I will consider the tonal alignment results from this thesis relative to findings on yes–no interrogative intonation in TB (Grice, Ridouane & Roettger 2015) as well as MA (Hellmuth et al. 2015) in order to arrive at a better understanding of variable tonal alignment in both languages.

In Part III of this thesis I argued for both TB and MA that the intonational prominence co-occurring with the qword is best interpreted as a non-metrical prominence-marking event, or 'non-metrical pitch accent'. The use of this term serves to highlight the contrast with prominence-marking tonal events whose location *is* metrically governed, i.e. pitch accents associating to stressed syllables. Nevertheless, the absence of a lexical–metrical association point does not necessarily mean that the realisation of postlexical prominence is entirely unconstrained. In the present work, tonal realisation in qwords was found to be constrained by the edges of the (phonological) domain of the qword, a domain characterised by information structural prominence (in the context of the TB experiment the qword was under narrow focus, and there are several arguments that support an interpretation of qwords as inherently focused, cf. Chapter 5). However, as discussed previously, in earlier work on intonation in both TB and MA, different patterns of alignment were reported, with yet other constraints that governed tonal association. For Tashlhiyt Berber, Grice, Ridouane & Roettger (2015) showed that syllable sonority and weight govern the alignment of phrase-final peaks as found at the right phrasal edge of yes–no interrogatives, implying that to some extent low-level phonetic factors influence the realisation of intonational movement. Perhaps this finding is not that unexpected: Many languages assign lexical stress based on stress-to-weight principles, which might be a result from the phonologisation of similar constraints. For Moroccan Arabic, Hellmuth et al. (2015) proposed that there is a role for the foot in the realisation of a phrase-final rise–fall in yes–no interrogatives. The rise–fall in this case was interpreted to associate with both an edge (the right phrasal edge) and a metrical position (the start of the phrase-final foot). While Hellmuth et al.'s (2015) interpretation referred to lexical stress, the patterns could just as well be interpreted with reference to syllable weight alone, since stress position according to Benkirane (1998) is determined based on syllable weight. Specifically, the start of the final foot would then simply coincide with the start of the final two moras (i.e. the start of a heavy final syllable, or the start of a penultimate syllable in the absence of a heavy final syllable). This would make syllable weight a factor constraining tonal association in both TB and MA.

One of the implications of this interpretation would be that metrical structure does indeed play a role in tonal association in languages lacking stress, since tonal association that makes reference to syllable weight entails reference to metrical structure. It is nevertheless clear that there is a difference between metrically-constrained association that makes reference to metrical structure in terms of *weight* and metrically-constrained association that makes reference to *stress* (which may in turn refer to weight in languages with stress-by-weight systems).

A more fundamental difference between findings on TB and MA intonation concerns the nature of the intonational events. In the present thesis, the intonational event associated with qwords is clearly prominence-marking, whereas in the aforementioned studies, rising or high pitch at the right edge of yes–no questions functions as a modality-marking edge tone. In contrast, the nature of the intonational movements that co-occur with contrastively focused words in phrase-final position in TB appear less clear (cf. Grice, Ridouane & Roettger 2015). While these pitch events are likely candidates to serve prominence-marking purposes, the fact that they occurred in phrase-final position makes it difficult to determine whether they in part (also) served right-edge IP-marking. Grice, Ridouane & Roettger's (2015) interpretation leaves this question open, and the intonational marking in both yes–no questions and contrastive statements is interpreted in terms of an H boundary tone that has secondary association to either the final or the penultimate syllable of the contrastively focused word (there was in fact clear evidence for alignment with a specific TBU, unlike in the present qword data, see also Chapter 6).

The postlexical prominence associated with qwords in the experiments presented in this thesis did not exhibit clear distributional patterns below the word level, but the materials were not designed to test non-stress related metrical influence. This leaves the question for future work whether prominence-marking and edge-marking postlexical tonal events in TB and MA can be reliably distinguished in general. The expectation, however, is that they can be, at least in cases where intonational events co-occur with contrastively focused words that do not occur near a phrasal edge (see Section 9.4 below).

In conclusion, a synthesis of the present results with results from earlier studies leads to the conclusion that tonal events in languages without stress *may* on the one hand be characterised by non-metrical association (as in qwords) and on the other hand they may still exhibit association that refers to metrical structure other than stress (as in yes–no questions in both languages, and contrastive statements in TB). In this context, the absence (thus far) of intonational events that align with reference to a specific syllable is very much compatible with the absence of lexically stressed syllables that would serve as such anchor points. The reverse, i.e. the existence of intonational events that seek association to a specific syllable, does not by definition imply that the relevant syllable is stressed (as in the case of the aforementioned studies that found weight-influenced alignment patterns). Assuming this reasoning is correct, variable alignment of pitch events can serve as a (possible) *diagnostic* for the absence of stress rather than as a necessary *consequence* of the absence of stress.

9.3.2 Typology

Finally, the present discussion is also relevant to the analysis of postlexical prominence across languages. Intonational events in stressless languages can still be prominence-marking even if they do not associate to lexically stressed syllables. Using the term 'non-metrical pitch accent' to describe such events makes this explicit, and precludes the need to resort to edge tones and phrase accents as the remaining options among AM categories. Yet another alternative term could be 'phrasal pitch accent' in order to describe cases like French and possibly Mongolian, in which postlexical prominence seeks a culminative position of strength at the phrasal level (as opposed to at the lexical level, which would be the case for a standard pitch accent). The present view stands in contrast to the implicit idea in e.g. Jun (2014a) that if, for a given language, the category of pitch accent is not available by virtue of the absence of lexical stress, all intonational events must be interpreted as boundary tones. While the analysis of intonation in terms of boundary tones only may be appropriate in the case of specific languages (e.g. Korean and Ambonese Malay, see Section 2.5.2), the results from qword interrogative intonation in TB and MA clearly indicate that postlexical prominence in these languages exists. Analysing the relevant intonational events in terms of edge association does not do justice to the facts and intuitions of several authors about prominence, including Dell & Elmedlaoui (2002), and would wrongly suggest that these events are functionally different from pitch accents. As argued here, the main difference between the intonational events discussed here and a standard pitch accent is one of form, rather of function: The present cases are simply characterised by the absence of a metrical anchor in the form of a lexically stressed syllable.

The view on lexical and postlexical prominence brought forward here has consequences for a typology that includes both levels of prominence, as shown in the following table (this is an updated version of Table 2.1 in Section 2.5.2, where a more detailed discussion of the relevant languages can be found).

Postlexical prominence	Lexical stress	No lexical stress
metrical	many, including: most European languages, most varieties of Arabic	French, Mongolian
non-metrical	Kuot, Chickasaw	Tashlhiyt Berber, Moroccan Arabic
absent	Wolof?	Korean, Ambonese Malay, West Greenlandic

Table 9.1: Proposed typology of languages as a function of lexical and postlexical prominence structure.

The left column represents languages with lexical stress, the right column languages that lack it. The top left cell represents many languages whose intonation is well-documented. These are languages with lexical stress that have pitch accents associating to stressed syllables. The top right cell represents languages without stress in which prominence-marking seeks out positions of (metrical) prominence above the word level. In French this is AP- or PP-final position and in Mongolian it seems to be AP-initial position. As previously mentioned, postlexical prominence in these cases is sometimes termed 'phrasal pitch accent'. The left cell in the middle row represents those languages in which stress presumably exists, but in which intonational prominence-marking does not seek association to these same syllables. The existence of this type of language was called into question in Chapter 2.5.1. The right cell in the middle row contains Tashlhiyt Berber and Moroccan Arabic as languages that lack lexical stress and exhibit postlexical prominence that does not take account of metrical prominence structure neither above nor below the word level. It was argued that the intonational prominence associated with qwords in these languages could be characterised as 'non-metrical pitch accent'. Future work, especially the comparison with intonational marking of other types of prominence (such as contrastive focus) might nevertheless shed further light on whether the languages really lack postlexical–metrical prominence structure altogether. The bottom row contains languages that are considered to lack any form of postlexical phonological prominence. On the one hand, Wolof has been argued to have lexical stress but no marking of intonational prominence at all. On the other hand, there are languages that seemingly lack any phonological prominence, at both the lexical and postlexical level: Korean and Ambonese Malay.

A final notes concerns the classification of a language as lacking both lexical and postlexical prominence, as this does not entail that the language lacks prosodic means altogether to mark differences in information structure, including, in particular, focus.

Korean, for example, is known to mark focus by means of phrasing. The crucial point here is that it would lack a culminative position at either the lexical, AP or other phrasal level, that serves as the predetermined location for a localised pitch event that contributes to the percept of prominence.

9.4 Directions for future work

The above discussion of the potential of lexical stress to determine the realisational details of postlexical prominence raises an important question about the advantages of the presence of lexical stress. Given the relative rarity of languages lacking lexical prominence structure altogether, one might raise the hypothesis that the presence of stress or lexical pitch accent is an 'optimal' linguistic property, favoured over its absence, which might even be linked to language users' cognitive abilities (this kind of reasoning follows the train of thought in Evolutionary Phonology, as in Blevins 2004). The advantages of the presence of lexical prominence asymmetries, specifically lexical stress, are in fact well-documented and include most notably the idea that word recognition and identification in continuous speech is facilitated by the known position of lexical stress for all words in the lexicon (e.g. McClelland & Elman 1986; Cutler & Norris 1988).

The advantages of lacking a representation of lexical prominence asymmetries are less clear. It could be hypothesised that a lack of prominence representation would mean a relative reduction of detail, which in turn might make available mental storage for other purposes. This argument would hold more strongly for a comparison between stressless languages and languages with variable, unpredictable stress, where stress position for each lexical item should be memorized, than for a comparison with languages with fixed stress or stress derived by rule. Results from the SRT in Chapter 8 in any case strongly suggest that native speakers of TB and MA do not have lexical representations for prominence structure within individual words.

Additionally, the absence of lexical stress as a predetermined lexical anchor for intonational pitch events might be considered a favourable feature of a phonological system, if it is assumed that intonational tonal realisation is the result of active effort on the part of the speaker to realise a specific timing. This idea is voiced in a stronger form in Bolinger (1978: 483): "Of all the various kinds of stress systems, the most advantageous of all for intonation, is the rarest: that in which … no particular syllable of a word is marked to carry the accent." As was discussed in the previous section, tonal alignment and association are not entirely unconstrained, even in TB and MA, languages that lack positions of culminative prominence at both lexical and postlexical levels. Therefore, the idea that intonational prominence is "free to move" in languages without stress is perhaps a little too strong (cf. Van Heuven & Faust 2009).

Further work on the realisation of prominence-marking intonational events will be able to shed more light on the factors that constrain the realisation of intonational tones in both Tashlhiyt Berber and Moroccan Arabic. Specifically, the marking of contrastive focus in positions not adjacent to phrasal edges (as has been the case in most studies to date, including Chapters 6 and 7) will be informative. Edge adjacency might impose production constraints on the realisation of tonal targets that obscure any potential metrical (e.g. weight-related) factors that would otherwise play a more important role. At this point it is known that contrastive focus in phrase-medial position, at least in MA, is cued by local F0 protrusions (Yeou, Embarki & Al-Maqtari 2007). This experiment yielded the insight that the rising–falling F0 contour in MA was less localised in its location than a comparable event that occurred in the same context in two other varieties of Arabic that do have stress.

Similarly, the results on intonational prominence in TB reported in Chapter 6 raised many questions about the language's intonation system. What are the intonational characteristics of contrastively focused words that occur phrase-medially? Do factors like syllable weight and sonority play a role in determining tonal alignment (as in Grice, Ridouane & Roettger 2015) or do intonational movements rather exhibit gradient alignment (as in the case study on qwords in Chapter 6)? More detailed studies of tonal alignment in both TB and MA will be required to gain a better insight into how both intonational systems deal with the realisation of postlexical prominence in the absence of lexical stress.

More generally, it is hoped that future work on other languages that lack lexical prominence structure will uncover which possible correspondences there are between metrically prominent positions at lexical and postlexical levels of phonological structure.

Appendices

A Tashlhiyt Berber scripted telephone conversation

The below gives the English translation of the scripted telephone dialogue between two speakers, M(other) and S(aid). The four target sentences (printed in bold) are glossed as well, in order, the lines represent i) (in italics) the Latin script version as presented to participants, ii) a phonemic transcription, and iii) a morphological gloss.

M. Hello Said! What are you doing now?
I want to know when you are coming home.

riR	*ad*	*issangh*	*managu*	*rad*	*tackt*	*s*	*tgmmi.*
riʁ	ad	is:anʁ	managu	rad	taʃkt	s	tgm:i.
want.AOR.1SG	COMP	know.AOR.1SG	when	FUT	2SG.come.AOR.2SG	to	EA.house

S. Sorry, mother, that I've kept you waiting. If I tell you what it is, you will know why I am so happy. I've found work in Marrakech! I will stay here until I have acquired some papers which they asked of me at my work.
When I get home, I know [it's that] you want to ask.

managu	*a*	*rann*	*ackR*	*ssnx*	*is*	*rad*	*t*
managu	a	ran:	aʃkʁ	s:nʁ	is	rad	t
when	COMP	FUT	come.AOR.1SG	know.AOR.1SG	COMP	FUT	CL.ACC.3SG.M

tinit
tinit
2SG.say.AOR.2SG

S. So you then, mother, everything okay with you?

M. Fine, my son, since you just found work in Marrakech. I'm very happy for you, my son.
So when will you start work?

imma	*managu*	*rad*	*tbdut*	*lxdmt?*
im:a	managu	rad	tbdut	lχdmt?
then	when	FUT	2SG.start.AOR.2SG	work?

[… The line is breaking up … As for Said, his mother cannot hear him very well.]

M. I am not hearing you very well.
When will you start work?!

managu	*rad*	*tbdut*	*lxdmt?!*
managu	rad	tbdut	lχdmt?!
when	FUT	2SG.start.AOR.2SG	work

S. Possibly this month or the one that will come after that, I don't know.

M. Ah! But your sister will marry this month!

S. **Ah, you had not yet said to me when.**

Ah,	*ur*	*iyi*	*yadlli*	*tnnim*	*managu.*
Ah,	ur	ij	jadl:i	tn:im	managu.
Ah	NEG	DAT.1SG.M	which	2PL.M.say.AOR.2PL.M	when

[… Said is cut off (from the network) …]

M. There's a big problem, how come?

S. Alright, so when in this month will the wedding take place? Because I have many things to do in Marrakech.

M. We said that we would have the wedding on the 12th of this month. Well, now you know the date, when will you come?

S. Alright I will come there three days before the wedding, because they said to me that I will start work on the 20th of this month. Well then, bye, stay well, mother.

M. Wait! Wait! Don't hang up!
You will start work when?

rad	*tbdut*	*lxdmt*	*managu?*
rad	tbdut	lχdmt	managu?
FUT	2SG.start.AOR.2SG	work	when

S. On the 20th of this month, mister, they said, you will start work with us.

M. Ah alright, stay well!

S. Bye mother, stay well, say hi to the family.

B Moroccan Arabic scripted dialogue

The dialogue is given in Arabic on the right and the English translation is given on the left. The two speakers in the dialogue are referred to as A and B (if the participant was male, they were given a version with initial call to Ziyaad, if female to Manaal). Target qword questions as reported on in Chapter 7 are printed in bold and identified by their IVAr code.

A	Ziyaad / Manaal!	زياد / منال!
B	Yes, who is it?	شكون؟
A	Good evening!	مسا لخير
B	Welcome, good evening, please come in.	أهلاً مسا لخير تفضلي
A	How are you?	كِداَيرَة؟ شخبارك؟
B	I've been tired for two days.	عيّانة هادي يومين
A	Why? What's up?	علاش؟ مالك؟
B	These days, we are inviting for the wedding of my uncle Maazin's daughter.	هادي يومين و احنا كنعرضو على الناس لعرس ديال بنت عمي مازن
A	Oh, is your cousin getting married? Congratulations.	بنت عمك غاتجوّج مبروك
B	Thank you and I wish the same for you.	الله يبارك فيك و العقباليكم
A	Who is getting married, Dina or Mayyaada?	شكون للّي غاتجوّج دينا ولّا ميّادة؟
B	Dina is the one who is getting married. And how are you?	للّي غاتجوّج هي دينا. و كِداَيرَة نتيا؟
A	Thank god, everything is fine. What have you been up to today?	لحمد لله كل شي لا باس. شنو درتي ليوم؟

B		Early evening, we played sport with Layla. And we got countryside flowers on our way from the shop Manaal.	فلعشية مشينا لكلوب مع ليلى. مشينا لحانوت منال نجيبو الورد البِلْدي
A		Did you go to the Yamani sport centre?	واش مشيتي لسّوق اليمني؟
B		No, I went to the Japanese sport centre.	لا مشيت لسّوق الياباني
A		Did you go to the sport center with Layla or Lina?	معا من مشيتي لسّوق مع ليلى ولّا لينا؟
B		I went to the sport center with Lina.	مشيت لسّوق معا لينا
A	whq3	**When is the wedding of your cousin Dina?**	إيمتا عرس بنت عمك دينا؟
A		Dina's wedding is in two days.	عرس دينا من هنا يومين
B		Is Dina's groom Lebanese or Yemeni?	واش عريس دينا لبناني ولّا يمني ؟
A		Dina's groom is Yemeni.	عريس دينا يمني
B	whq1	**What is the name of the Yemeni man?**	شنو سميت ذا الراجل اليمني؟
A		The Yemeni man's name is Nabil Al-Badawi.	هذا الراجل اليمني سميتو نبيل البدوي
B		Did they meet through Zeena?	واش لاقاتهم زينة؟
A		Yes, they met each other through Zeena.	أه لاقاتهم زينة
B	whq5	**Where did Dina and Nabil meet?**	فين شافت دينا نبيل؟
A		She met him at the Alghali cafe.	شافتو فقهوة الغالي
B		Does that mean that she met him in the cafe which is in the Morocco centre?	زعما شافتو فهاديك لقهوة للّي كاينة فالمول المغربي؟
A		Yes, she met him at the Al-Maghribi shopping centre.	أه تعرفت عليه في المول المغربي
B		Is it going to be a religious or civil marriage?	واش العرس بلدي ولّا رومي؟
A		The marriage is mostly civil.	العرس غالباً رومي

B	whq2	**Who is going to witness the civil marriage?**	شكون للّي غيشهد على العرس الرومي؟
A		Noor and Zeen are going to witness the civil marriage.	نور و زين غيشِهْدو على العرس الرومي
B	whq4	**In which city will the wedding be held?**	العرس غيكون فأش من مدينة؟
A		The wedding is most likely in Dubai. Will you be able to go to Dubai?	العرس غالباً في دبي. تقدري تمشي لدبي؟
B		I will try because Dina is very dear to me.	و الله تنحاول .دينا عزيزة علية بزّاف
A		This wedding will be a mixture of urban and bedouin styles.	هاذ العرس مخلط بين البلدي و الرومي
A		Is the civil wedding going to be held at the Al-Baladi building?	العرس غيكون فالقصر البلدي؟
B		Yes, Dina's wedding is at the Al-Baladi building.	أه عرس دينا غيكون فالقصر البلدي
A		Is the party going to be in Layalina or Bayaan hall?	واش الحفلة غتكون فقاعة ليالينا ولّا بيان؟
B		They will most likely book Al-Bayaan hall.	غالباً غيريزرڤيو فبيان
A		Are they going to Dubai or to Lebanon after the wedding?	واش غيمشيو من بعد العرس لدبي ولّا لبنان؟
B		They are going to spend the days after marriage in Lebanon.	غيدوزو شي يامات من بعد العرس فلبنان
A		Does it mean that she will visit her sister Layali?	زعما غتمشي تشوف ختها ليالي؟
B		Yes, she will visit her sister Layali for a few days.	أه غتمشي تشوف ختها ليالي شي يامات
A		Is Nabil's father going to be there?	بّات نبيل غيكون موجود؟
B		Yes, everybody will be there.	أه كلشي غيكون موجود
A	whq6	**What are you going to get for Dina from Dalaal's shop?**	شنو غتجيبي لدينا من حانوت دلال؟
B		I am not going to Dalaal's, I am thinking of getting her a teddy bear from the White Flower. Today you'll get the invitation with Walid.	أنا ما غاداش عند دلال. فكِّرنجيب ليها نونوس من عند الوردة البيضا. اليوم غادي يعرض عليك وليد

A	Yes sure, anyway, congratulations. I will call her and congratulate her now.	أه بالصح؟ المهم مبروك. غنعيّط ليها و نبارك ليها دابا

Bibliography

Abercrombie, David. 1976. Stress and some other terms. *Work in Progress* 9. 51–53.

Aboh, Enoch O. 2007. Focused versus non-focused wh-phrases. In Enoch O. Aboh, Katharina Hartmann & Malte Zimmermann (eds.), *Focus strategies in African languages: The interaction of focus and grammar in Niger-Congo and Afro-Asiatic*, vol. 191 (Trends in Linguistics, Studies and Monographs), 287–314. Berlin: Mouton de Gruyter.

Aguadé, Jordi. 2008. Moroccan Arabic. In Kees Versteegh (ed.), *Encyclopedia of Arabic language and linguistics*, vol. III, 273–297. Leiden: Brill.

Akrofi Ansah, Mercy. 2010. Focused constituent interrogatives in Lɛtɛ (Larteh). *Nordic Journal of African studies* 19(2). 98–107.

Al-Ani, Salman H. 1970. *Arabic phonology: An acoustical and physiological investigation*. Indiana University PhD dissertation.

Almbark, Rana, Nadia Bouchhioua & Sam Hellmuth. 2014. Acquiring the phonetics and phonology of English word stress: Comparing learners from different L1 backgrounds. *Concordia Working Papers in Applied Linguistics* 5. 19–35.

Al-Tamimi, Jalal. 2009. Effect of pharyngealisation on vowels revisited: Static and dynamic analyses of vowels in Moroccan and Jordanian Arabic. In *Workshop on Pharyngeals and Pharyngealisation*. Newcastle.

Al-Tamimi, Jalal. 2017. Revisiting acoustic correlates of pharyngealization in Jordanian and Moroccan Arabic: Implications for formal representations. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 8(1). 1–40. https://doi.org/10.5334/labphon.19.

Antônio de Moraes, João. 1998. Intonation in Brazilian Portuguese. In Daniel Hirst & Albert Di Cristo (eds.), *Intonation systems: A survey of twenty languages*, 179–194. Cambridge: Cambridge University Press.

Applegate, Joseph R. 1958. *An outline of the structure of Shilha*. New York: American Council of Learned Societies.

Arnhold, Anja. 2014. Prosodic structure and focus realization in West Greenlandic. In Sun-Ah Jun (ed.), *Prosodic Typology II: The phonology of intonation and phrasing*, 216–251. Oxford: Oxford University Press.

Arvaniti, Amalia & D. Robert Ladd. 2009. Greek wh-questions and the phonology of intonation. *Phonology* 26(1). 43–74. https://doi.org/10.1017/S0952675709001717.

Arvaniti, Amalia, D. Robert Ladd & Ineke Mennen. 1998. Stability of tonal alignment: the case of Greek prenuclear accents. *Journal of Phonetics* 26(1). 3–25. https://doi.org/doi:10.1006/jpho.1997.0063.

Arvaniti, Amalia, D. Robert Ladd & Ineke Mennen. 2000. What is a starred tone? Evidence from Greek. In Michael J. Broe & Janet B. Pierrehumbert (eds.), *Papers in Labor-*

atory Phonology V: Acquisition and the lexicon, 119–131. Cambridge: Cambridge University Press.

Aspinion, Robert. 1953. *Apprenons le berbère: Initiation aux dialectes chleuhs*. Rabat: Moncho.

Athanasopoulou, Angeliki & Irene Vogel. 2016. The acoustic manifestation of prominence in stressless languages. In *Proceedings of Interspeech 2016*, 82–86. San Francisco. https://doi.org/10.21437/Interspeech.2016-1424.

Atterer, Michaela & D. Robert Ladd. 2004. On the phonetics and phonology of "segmental anchoring" of F0: Evidence from German. *Journal of Phonetics* 32(2). 177–197. https://doi.org/10.1016/S0095-4470(03)00039-1.

Barnes, Jonathan, Alejna Brugos, Stefanie Shattuck-Hufnagel & Nanette Veilleux. 2012a. On the nature of perceptual differences between accentual peaks and plateaux. In Oliver Niebuhr (ed.), *Understanding prosody: The role of context, function and communication* (Language, Context, and Cognition), 93–118. Berlin: De Gruyter.

Barnes, Jonathan, Nanette Veilleux, Alejna Brugos & Stefanie Shattuck-Hufnagel. 2012b. Tonal Center of Gravity. *Journal of Laboratory Phonology* 3(2). 337–383. https://doi.org/10.1515/lp-2012-0017.

Bartels, Christine. 1997. The pragmatics of wh-question intonation in English. *University of Pennsylvania Working Papers in Linguistics* 4(2). http://repository.upenn.edu/pwpl/vol4/iss2/1.

Bates, Douglas, Martin Maechler, Ben Bolker & Steven Walker. 2015. *lme4: Linear mixed-effects models using Eigen and S4*. R package. https://doi.org/https://github.com/lme4/lme4/.

Bateson, Mary Catherine. 1967. *Arabic language handbook*. Vol. 3. Washington, D.C.: Georgetown University Press.

Baunaz, Lena & Cédric Patin. 2009. Prosody refers to semantic factors: Evidence from French wh-words. In *Proceedings of IDP 09*, 93–107.

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

Beckman, Mary E., Julia Hirschberg & Stefanie Shattuck-Hufnagel. 2005. The original ToBI system and the evolution of the ToBI framework. In Sun-Ah Jun (ed.), *Prosodic Typology: The phonology of intonation and phrasing*, 9–54. Oxford: Oxford University Press.

Beckman, Mary E. & Janet B. Pierrehumbert. 1986. Intonational structure in Japanese and English. *Phonology Yearbook* 3. 255–309.

Benhallam, Abderrafi. 1989. Aspects de la recherche en phonologie de l'arabic marocain. In *Langue et société au Maghreb*, 13–23. Rabat: Publications de la faculté des Lettres et des Sciences Humaines.

Benhallam, Abderrafi. 1990. Native speaker intuitions about Moroccan Arabic stress. In Jochen Pleines (ed.), *Maghreb linguistics*, 91–109. Rabat: Éditions Okad.

Benkirane, Thami. 1982. *Étude phonétique et fonctions de la syllabe en arabe marocain*. Université de Provence Aix-Marseille I PhD dissertation.

Benkirane, Thami. 1998. Intonation in Western Arabic (Morocco). In Daniel Hirst & Albert Di Cristo (eds.), *Intonation systems: A survey of twenty languages*, 345–359. Cambridge: Cambridge University Press.

Bishop, Judith & Janet Fletcher. 2005. Intonation in six dialects of Bininj Gun-Wok. In Sun-Ah Jun (ed.), *Prosodic Typology: The phonology of intonation and phrasing*, 331–361. Oxford: Oxford University Press.

Blevins, Juliette. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.

Bocci, Giuliano & Cinzia Avesani. 2011. Phrasal prominences do not need pitch movements: Postfocal phrasal heads in Italian. In *Proceedings of Interspeech 2011*. Florence.

Boersma, Paul & David Weenink. 2015. *Praat: Doing phonetics by computer*. Software. Version 6.0.22.

Bolinger, Dwight (ed.). 1972. *Intonation: Selected readings*. Harmondsworth: Penguin Education.

Bolinger, Dwight. 1978. Intonation across languages. In Joseph H. Greenberg (ed.), *Universals of human language*, 471–524. Stanford: Stanford University Press.

Bolinger, Dwight. 1986. *Intonation and its parts: Melody in spoken English*. London: Edward Arnold.

Bolinger, Dwight L. 1958. A theory of pitch accent in English. *Word* 14(2–3). 109–149. `https://doi.org/10.1080/00437956.1958.11659660`.

Bouchhioua, Nadia. 2008. *The acoustic correlates of stress and accent in Tunisian Arabic: A comparative study with English*. Université de 7 Novembre: l'Institut Supérieur des Langues de Tunis PhD dissertation.

Boudlal, Abdelaziz. 2001. *Constraint interaction in the phonology and morphology of Casablanca Moroccan Arabic*. Université Mohammed V PhD dissertation.

Brockelmann, Carl. 1908. *Grundriss der vergleichenden Grammatik der semitischen Sprachen*. Berlin: Reuther & Reichard.

Bruce, Gösta. 1977. *Swedish word accents in sentence perspective*. Malmo: LiberLäromedel/Gleerup.

Bruggeman, Anna, Nabila Louriz, Rana Almbark & Sam Hellmuth. in press. Acoustic correlates of lexical stress in Moroccan Arabic. *Journal of the International Phonetic Association*.

Bruggeman, Anna & Timo Benjamin Roettger. 2017. *CoTaSS: Corpus of Tashlhiyt Semi-spontaneous Speech*. Online speech database. `http://cotass.uni-koeln.de`.

Bruggeman, Anna, Timo Benjamin Roettger & Martine Grice. 2017. Question word intonation in Tashlhiyt: Is 'high' good enough? *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 8(1), 5. `https://doi.org/10.5334/labphon.79`.

Brunelle, Marc. 2017. Stress and phrasal prominence in tone languages: The case of Southern Vietnamese. *Journal of the International Phonetic Association* 47(3). 283–320. `https://doi.org/10.1017/S0025100316000402`.

Burdin, Rachel Steindel, Sara Phillips-Bourass, Rory Turnbull, Murat Yasavul, Cynthia. G. Clopper & Judith Tonhauser. 2015. Variation in the prosody of focus in head- and head/edge-prominence languages. *Lingua* 165(Part B). 254–276. https://doi.org/10.1016/j.lingua.2014.10.001.

Büring, Daniel. 2007. Semantics, intonation and information structure. In Gillian Ramchand & Charles Reiss (eds.), *The Oxford handbook of linguistic interfaces*, 445–474. Oxford: Oxford University Press.

Büring, Daniel. 2012. Focus and Intonation. In Delia Graff Fara & Gillian Russell (eds.), *The Routledge Companion to the Philosophy of Language*, 103–115. New York: Routledge.

Cambier-Langeveld, Tina. 2000. *Temporal marking of accents and boundaries* (LOT series). The Hague: Holland Academic Graphics.

Cangemi, Francesco. 2015. *mausmooth*. Praat script. http://phonetik.phil-fak.uni-koeln.de/fileadmin/Phonetik_Files/Mitarbeiterdateien/Francesco_Cangemi/cangemi2015mausmooth.pdf (26 October, 2016).

Cantineau, Jean. 1960. *Cours de phonétique arabe*. Paris: Klincksieck.

Chafe, Wallace L. 1970. *A semantically based sketch of Onondaga* (Indiana University Publications in Anthropology and Linguistics). Baltimore: Waverley Press (Indiana University).

Chahal, Dana. 2003. Phonetic cues to prominence in Lebanese Arabic. In *Proceedings of the International Congress of Phonetic Sciences XV*, 2067–2070. Barcelona.

Chahal, Dana & Sam Hellmuth. 2014. The intonation of Lebanese and Egyptian Arabic. In Sun-Ah Jun (ed.), *Prosodic Typology II: The phonology of intonation and phrasing*, 365–404. Oxford: Oxford University Press.

Chapallaz, Marguerite. 1972. Notes on the intonation of Italian. In Dwight Bolinger (ed.), *Intonation: Selected readings*, 359–364. Harmondsworth: Penguin Education.

Chen, Aoju. 2012. Shaping the intonation of wh-questions: Information structure and beyond. In Jan P. de Ruiter (ed.), *Questions: Formal, functional and interactive perspectives*. Cambridge: Cambridge University Press.

Cheng, Lisa Lai-Shen. 2003. Wh-in-situ. *Glot International* 7(4). 103–109.

Choe, Hyon Sook. 1995. Focus and topic movement in Korean and licensing. In Katalin É. Kiss (ed.), *Discourse configurational languages*, 269–334. Oxford: Oxford University Press.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.

Coetzee, Andries W., Patrice S. Beddor, Kerby Shedden, Will Styler & Daan Wissing. 2018. Plosive voicing in Afrikaans: Differential cue weighting and tonogenesis. *Journal of Phonetics* 66(1). 185–216. https://doi.org/10.1016/j.wocn.2017.09.009.

Coetzee, Andries W., Patrice S. Beddor & Daan P. Wissing. 2014. Emergent tonogenesis in Afrikaans. *The Journal of the Acoustical Society of America* 135(4). 2421–2422. https://doi.org/10.1121/1.4878043.

Cohen, Antonie & Johan ’t Hart. 1968. On the anatomy of intonation. *Lingua* 19(1–2). 177–192. https://doi.org/10.1016/0024-3841(69)90118-1.

Coleman, John. 1999. The nature of vocoids associated with syllabic consonants in Tashlhiyt Berber. In *Proceedings of the International Congress of Phonetic Sciences XIV*, 735–738. San Francisco.

Coleman, John. 2001. The phonetics and phonology of Tashlhiyt Berber syllabic consonants. *Transactions of the Philological Society* 99(1). 29–64. https://doi.org/10.1111/1467-968X.00073.

Correia, Susana, Joseph Butler, Marina Vigário & Sónia Frota. 2015. A stress "deafness" effect in European Portuguese. *Language and Speech* 58(1). 48–67. https://doi.org/10.1177/0023830914565193.

Crosswhite, Katherine. 2001. *Vowel reduction in Optimality Theory* (Outstanding dissertations in linguistics). New York: Routledge.

Cruttenden, Alan. 1986. *Intonation*. Cambridge: Cambridge University Press.

Crystal, David. 1972. The intonation system of English. In Dwight Bolinger (ed.), *Intonation: Selected readings*, 110–137. Harmondsworth: Penguin Education.

Culicover, Peter W. & Michael Rochemont. 1983. Stress and focus in English. *Language* 59(1). 123–165. https://doi.org/10.2307/414063.

Cutler, Anne & Dennis Norris. 1988. The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human perception and performance* 14(1). 113–121. https://doi.org/10.1037/0096-1523.14.1.113.

d'Alessandro, Christophe & Piet Mertens. 1995. Automatic pitch contour stylization using a model of tonal perception. *Computer Speech and Language* 9. 257–288. https://doi.org/10.1006/csla.1995.0013.

Dascalu-Jinga, Laurentia. 1998. Intonation in Romanian. In Daniel Hirst & Albert Di Cristo (eds.), *Intonation systems: A survey of twenty languages*, 239–260. Cambridge: Cambridge University Press.

Dell, François & Mohamed Elmedlaoui. 1985. Syllabic consonants and syllabification in Imdlawn Tashlhiyt Berber. *Journal of African Languages and Linguistics* 7(2). 105–130. https://doi.org/10.1515/jall.1985.7.2.105.

Dell, François & Mohamed Elmedlaoui. 2002. *Syllables in Tashlhiyt Berber and in Moroccan Arabic*. Dordrecht: Kluwer.

Dell, François & Mohamed Elmedlaoui. 2008. *Poetic meter and musical form in Tashlhiyt Berber songs*. Köln: Rüdiger Köppe Verlag.

Dik, Simon C. 1997a. *The theory of Functional Grammar: Part 1, The structure of the clause* (Functional Grammar Series). Berlin: De Gruyter.

Dik, Simon C. 1997b. *The theory of Functional Grammar: Part 2, Complex and derived constructions* (Functional Grammar Series). Berlin: De Gruyter.

Dilley, Laura C., D. Robert Ladd & Astrid Schepman. 2005. Alignment of L and H in bitonal pitch accents: Testing two hypotheses. *Journal of Phonetics* 33(1). 115–119. https://doi.org/10.1016/j.wocn.2004.02.003.

D'Imperio, Mariapaola. 2006. Preface. *Rivista di Linguistica* 18(1). 3–18.

Dupoux, Emmanuel, Christophe Pallier, Nuria Sebastián & Jacques Mehler. 1997. A destressing "deafness" in French? *Journal of Memory and Language* 36(3). 406–421. https://doi.org/10.1006/jmla.1996.2500.

Dupoux, Emmanuel & Sharon Peperkamp. 2002. Fossil markers of language development: phonological 'deafnesses' in adult speech processing. In Jacques Durand & Bernard Laks (eds.), *Phonetics, phonology, and cognition*, 168–190. Oxford: Oxford University Press.

Dupoux, Emmanuel, Sharon Peperkamp & Núria Sebastián-Gallés. 2001. A robust method to study stress "deafness". *The Journal of the Acoustical Society of America* 110(3). 1606–1618. https://doi.org/10.1121/1.1380437.

Dupoux, Emmanuel, Sharon Peperkamp & Núria Sebastián-Gallés. 2010. Limits on bilingualism revisited: Stress 'deafness' in simultaneous French-Spanish bilinguals. *Cognition* 114(2). 266–275. https://doi.org/10.1016/j.cognition.2009.10.001.

Dupoux, Emmanuel, Núria Sebastián-Gallés, Eduardo Navarrete & Sharon Peperkamp. 2008. Persistent stress 'deafness': The case of French learners of Spanish. *Cognition* 106(2). 682–706. https://doi.org/{10.1016/j.cognition.2007.04.001}.

Durand, Olivier. 1994. *Profilo di arabo marocchino*. Roma: Universitá degli studie 'La sapienza'.

É. Kiss, Katalin. 1995. Discourse configurational languages: Introduction. In Katalin É. Kiss (ed.), *Discourse configurational languages*, 3–27. Oxford: Oxford University Press.

El Aissati, Abderrahman. 2005. A socio-historical perspective on the Amazigh (Berber) cultural movement in North America. *Afrika Focus* 18(1/2). 59–72.

El Zarka, Dina. 2011. Leading, linking, and closing tones and tunes in Egyptian Arabic: What a simple intonation system tells us about the nature of intonation. In Ellen Broselow & Hamid Ouali (eds.), *Perspectives on Arabic linguistics: Papers from the annual symposia on Arabic linguistics volume XXII-XXIII (College Park, Maryland, 2008 and Milwaukee, Wisconsin, 2009)*, 57–74. Amsterdam: John Benjamins.

Elordieta, Gorka. 1998. Intonation in a pitch accent variety of Basque. *International Journal of Basque Linguistics and Philology* 32. 511–569.

Elordieta, Gorka & José Ignacio Hualde. 2003. Tonal and durational correlates of accent in contexts of downstep in Lekeitio Basque. *Journal of the International Phonetic Association* 33(2). 195–209. https://doi.org/10.1017/S0025100303001294.

Elordieta, Gorka & José Ignacio Hualde. 2014. Intonation in Basque. In Sun-Ah Jun (ed.), *Prosodic Typology II: The phonology of intonation and phrasing*, 405–463. Oxford: Oxford University Press.

E-Prime 3.0. 2017. Software. Pittsburgh, PA: Psychology Software Tools, Inc. http://www.pstnet.com.

Erteschik-Shir, Nomi. 1986. Wh-questions and focus. *Linguistics and Philosophy* 9(2). 117–149. https://doi.org/10.1007/BF00635608.

Erteschik-Shir, Nomi. 2007. *Information structure: The syntax-discourse interface*. Oxford: Oxford University Press.

Evans, Jonathan (ed.). 2015. *What's so special about H(igh)? Multi-disciplinary perspectives on the linguistic functions of raised pitch*. Special issue of *Journal of Phonetics*.

Fischer, August. 1917. *Zur Lautlehre des Marokkanisch-Arabischen*. Leipzig: Edelmann.

Fletcher, Janet. 2014. Intonation and prosody in Dalabon. In Sun-Ah Jun (ed.), *Prosodic Typology II: The phonology of intonation and phrasing*, 252–272. Oxford: Oxford University Press.

Foster, Michael K. 1982. Alternating weak and strong syllables in Cayuga words. *International Journal of American Linguistics* 48(1). 59–72.

Frota, Sónia. 2002. Nuclear falls and rises in European Portuguese: A phonological analysis of declarative and question intonation. *Probus* 14(1). 113–146. https://doi.org/10.1515/prbs.2002.001.

Frota, Sónia & Marina Vigário. 2000. Aspectos de prosódia comparada: Ritmo e entoação no PE e no PB. In *Actas do XV encontro da Associação Portuguesa de Linguística*, 533–555. Braga: Associação Portuguesa de Linguística.

Fry, Dennis B. 1955. Duration and intensity as physical correlates of linguistic stress. *The Journal of the Acoustical Society of America* 27(4). 765–768. https://doi.org/10.1121/1.1908022.

Fry, Dennis B. 1958. Experiments in the perception of stress. *Language and Speech* 1(2). 126–152. https://doi.org/10.1177/002383095800100207.

Fujisaki, Hiroya & Keikichi Hirose. 1984. Analysis of voice fundamental frequency contours for declarative sentences of Japanese. *Journal of the Acoustical Society of Japan (E.)* 5(4). 233–242. https://doi.org/10.1250/ast.5.233.

Genzel, Susanne. 2013. *Lexical and post-lexical tones in Akan*. Universität Potsdam PhD dissertation.

Goedemans, Rob & Harry van der Hulst. 2013. Fixed stress locations. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/14.

Goedemans, Rob & Ellen Van Zanten. 2007. Stress and accent in Indonesian. *LOT Occasional series* 9. 35–62.

Goldsmith, John A. 1976. *Autosegmental phonology*. Bloomington, IN: Indiana University Linguistics Club Bloomington.

Gooden, Shelome. 2014. Aspects of the intonational phonology of Jamaican Creole. In Sun-Ah Jun (ed.), *Prosodic Typology II: The phonology of intonation and phrasing*, 273–301. Oxford: Oxford University Press.

Gordon, Matthew. 2005. Intonational phonology of Chickasaw. In Sun-Ah Jun (ed.), *Prosodic Typology: The phonology of intonation and phrasing*, 301–330. Oxford: Oxford University Press.

Gordon, Matthew. 2014. Disentangling stress and pitch accent: A typology of prominence at different prosodic levels. In Harry van der Hulst (ed.), *Word stress: Theoretical and typological issues*, 83–118. Cambridge: Cambridge University Press.

Gordon, Matthew & Latifa Nafi. 2012. Acoustic correlates of stress and pitch accent in Tashelhiyt Berber. *Journal of Phonetics* 40(5). 706–724. https://doi.org/10.1016/j.wocn.2012.04.003.

Gordon, Matthew & Timo Roettger. 2017. Acoustic correlates of word stress: A cross-linguistic survey. *Linguistics Vanguard* 3(1). https://doi.org/10.1515/lingvan-2017-0007.

Gorman, Kyle, Jonathan Howell & Michael Wagner. 2011. Prosodylab-Aligner: A tool for forced alignment of laboratory speech. *Canadian Acoustics* 39(3). 192–193. jcaa.caa-aca.ca/index.php/jcaa/article/view/2476.

Grice, Martine, Stefan Baumann & Ralf Benzmülller. 2005. German intonation in Autosegmental-Metrical phonology. In Sun-Ah Jun (ed.), *Prosodic Typology: The phonology of intonation and phrasing*, 55–83. Oxford: Oxford University Press.

Grice, Martine, D. Robert Ladd & Amalia Arvaniti. 2000. On the place of phrase accents in intonational phonology. *Phonology* 17(2). 143–185. https://doi.org/10.1017/S0952675700003924.

Grice, Martine, Rachid Ridouane & Timo Benjamin Roettger. 2015. Tonal association in Tashlhiyt Berber: Evidence from polar questions and contrastive statements. *Phonology* 32(2). 241–266. https://doi.org/10.1017/S0952675715000147.

Grice, Martine, Simon Ritter, Henrik Niemann & Timo B. Roettger. 2017. Integrating the discreteness and continuity of intonational categories. *Journal of Phonetics* 64(1). 90–107. https://doi.org/10.1016/j.wocn.2017.03.003.

Grice, Martine, Timo Benjamin Roettger, Rachid Ridouane & Cécile Fougeron. 2011. The association of tones in Tashlhiyt Berber. In *Proceedings of the International Congress of Phonetic Sciences XVII*. Hong Kong.

Grice, Martine, Alexandra Vella & Anna Bruggeman. 2019. Stress, pitch accent, and beyond: Intonation in Maltese questions. *Journal of Phonetics* 76. 100913. https://doi.org/10.1016/j.wocn.2019.100913.

Gussenhoven, Carlos. 1983. Focus, mode and the nucleus. *Journal of Linguistics* 19(2). 377–417. https://doi.org/10.1017/S0022226700007799.

Gussenhoven, Carlos. 2004. *The phonology of tone and intonation* (Research Surveys in Linguistics). Cambridge: Cambridge University Press.

Gussenhoven, Carlos. 2005. Transcription of Dutch intonation. In Sun-Ah Jun (ed.), *Prosodic Typology: The phonology of intonation and phrasing*, 118–145. Oxford: Oxford University Press.

Gussenhoven, Carlos. 2015. Does phonological prominence exist? *Linge e Linguaggio* XIV(1). 7–24.

Gussenhoven, Carlos. 2017. Zwara (Zuwārah) Berber. *Journal of the International Phonetic Association*. 1–17. https://doi.org/10.1017/S0025100317000135.

Haan, Judith. 2002. *Speaking of questions: An exploration of Dutch question intonation* (LOT series). The Hague: Holland Academic Graphics.

Hagstrom, Paul. 2003. What questions mean. *Glot International* 7. 188–201.

Halim, Amran. 1974. *Intonation in relation to syntax in bahasa Indonesia*. Jakarta: Djambatan.

Hamdi, Rachid. 1991. *Etude phonologique et expérimentale de l'emphase en arabe marocain de Casablanca*. Université Lyon 2 PhD dissertation.

Harrell, Richard Slade. 1962. *A short reference grammar of Moroccan Arabic: With an appendix of texts in Urban Moroccan Arabic by Louis Brunot.* Washington, D.C.: Georgetown University Press.

Hayes, Bruce. 1995. *Metrical stress theory: Principles and case studies*. Chicago: The University of Chicago Press.

Hayes, Bruce & Aditi Lahiri. 1991. Bengali intonational phonology. *Natural Language and Linguistic Theory* 9(1). 47–96. `https://doi.org/10.1007/BF00133326`.

Heath, Jeffrey. 1997. Moroccan Arabic Phonology. In Alan S. Kaye (ed.), *Phonologies of Asia and Africa (including the Caucasus)*, 205–217. Winona Lake (IN): Eisenbrauns.

Heath, Jeffrey. 2005. *A grammar of Tamashek (Tuareg of Mali)* (Mouton grammar library 35). Berlin: De Gruyter.

Hellmuth, Sam. 2006. *Intonational pitch accent distribution in Egyptian Arabic*. SOAS University of London PhD dissertation.

Hellmuth, Sam & Rana Almbark. 2017. *IVAr: Intonational Variation in Arabic Corpus*. Data collection. Colchester. `https://dx.doi.org/10.5255/UKDA-SN-852878`.

Hellmuth, Sam, Nabila Louriz, Basma Chlaihani & Rana Almbark. 2015. F0 peak alignment in Moroccan Arabic polar questions. In *Proceedings of the International Congress of Phonetic Sciences XVIII*. Glasgow, UK.

Hellmuth, Sam, Becky Muradás-Taylor & Bethany Karrinton. to appear. Cue dependent stress perception in english listeners: the effect of linguistic experience.

Hellmuth, Sam, Becky Muradás-Taylor & Bethany White. N.d. Non-persistent stress deafness in English listeners.

Henriksen, Nicholas. 2014. Initial peaks and final falls in the intonation of Manchego Spanish wh-questions. *Probus* 26(1). 83–133. `https://doi.org/10.1515/probus-2013-0003`.

Hirschberg, Julia. 2000. A corpus-based approach to the study of speaking style. In Merle Horne (ed.), *Prosody: Theory and experiment. Studies presented to Gösta Bruce*, 335–350. Dordrecht: Kluwer Academic Publishers.

Hirst, Daniel & Albert Di Cristo (eds.). 1998. *Intonation systems: A survey of twenty languages*. Cambridge: Cambridge University Press.

Holes, Clive. 2004. *Modern Arabic: Structures, functions, and varieties*. Revised edition. Washington, D.C.: Georgetown University Press.

Horvath, Julia. 1986. *FOCUS in the theory of grammar and the syntax of Hungarian*. Dordrecht: Foris Publications.

Hyman, Larry. 1977. On the nature of linguistic stress. In Larry Hyman (ed.), *Studies in stress and accent* (Southern California Occasional Papers in Linguistics 4), 37–82. Los Angeles: University of Southern California.

Hyman, Larry. 2014. Do all languages have word accent? In Harry van der Hulst (ed.), *Word stress: Theoretical and typological issues*, 56–82. Cambridge: Cambridge University Press.

Igarashi, Yosuke. 2014. Typology of intonational phrasing in Japanese dialects. In Sun-Ah Jun (ed.), *Prosodic Typology II: The phonology of intonation and phrasing*, 464–492. Oxford: Oxford University Press.

Jacobsen, Birgitte. 2000. The question of 'stress' in West Greenlandic. *Phonetica* 57(1). 40–67. https://doi.org/10.1159/000028458.

Jaggar, Philip J. 2006. More on in-situ wh- and focus constructions in Hausa. In Dymitr Ibriszimow, Henry Tourneux & Ekkehard H. Wolff (eds.), *Chadic Linguistics/Linguistique Tchadique/Tschadistik*, vol. 3, 49–73. Köln: Rüdiger Köppe Verlag.

Jun, Sun-Ah. 1993. *The phonetics and phonology of Korean prosody*. The Ohio State University PhD dissertation.

Jun, Sun-Ah. 2005a. Korean intonational phonology and prosodic transcription. In Sun-Ah Jun (ed.), *Prosodic Typology: The phonology of intonation and phrasing*, 201–229. Oxford: Oxford University Press.

Jun, Sun-Ah. 2005b. Prosodic Typology. In Sun-Ah Jun (ed.), *Prosodic Typology: The phonology of intonation and phrasing*, 430–458. Oxford: Oxford University Press.

Jun, Sun-Ah (ed.). 2005c. *Prosodic Typology: The phonology of intonation and phrasing*. Oxford: Oxford University Press.

Jun, Sun-Ah. 2014a. Prosodic typology: By prominence type, word prosody, and macro-rhythm. In Sun-Ah Jun (ed.), *Prosodic typology II: The phonology of intonation and phrasing*, 520–539. Oxford: Oxford University Press.

Jun, Sun-Ah (ed.). 2014b. *Prosodic Typology II: The phonology of intonation and phrasing*. Oxford: Oxford University Press.

Jun, Sun-Ah & Cécile Fougeron. 2002. Realisations of accentual phrase in French intonation. *Probus* 14(1). 147–172. https://doi.org/10.1515/prbs.2002.002.

Jun, Sun-Ah & Mira Oh. 1996. A prosodic analysis of three types of wh-phrases in Korean. *Language and Speech* 39(1). 37–61. https://doi.org/10.1177/002383099603900103.

Kabak, Barış & Irene Vogel. 2001. The phonological word and stress assignment in Turkish. *Phonology* 18(3). 315–360. https://doi.org/10.1017/S0952675701004201.

Kang, Yoonjung & Sungwoo Han. 2013. Tonogenesis in early Contemporary Seoul Korean: A longitudinal case study. *Lingua* 134(1). 62–74. https://doi.org/10.1016/j.lingua.2013.06.002.

Karlsson, Anastasia. 2014. The intonational phonology of Mongolian. In Sun-Ah Jun (ed.), *Prosodic Typology II: The phonology of intonation and phrasing*, 187–215. Oxford: Oxford University Press.

Karlsson, Anastasia & Jan-Olof Svantesson. 2004. Prominence and mora in Mongolian. In *Proceedings of Speech Prosody 2*. Nara, Japan.

Keane, Elinor. 2003. Word-level prominence distinctions in Tamil. In *Proceedings of the International Congress of Phonetic Sciences XV*, 1257–1260. Barcelona.

Keane, Elinor. 2006a. Phonetics vs. phonology in Tamil wh-questions. In *Proceedings of Speech Prosody 3*. Dresden.

Keane, Elinor. 2006b. Prominence in Tamil. *Journal of the International Phonetic Association* 36(1). 1–20. https://doi.org/10.1017/S0025100306002337.

Kendall, Tyler & Erik R. Thomas. 2014. *vowels*. R package. http://ncslaap.lib.ncsu.edu/tools/norm/.

Khan, Sameer ud Dowla. 2014. The intonational phonology of Bangladeshi Standard Bengali. In Sun-Ah Jun (ed.), *Prosodic Typology II: The phonology of intonation and phrasing*, 81–117. Oxford: Oxford University Press.

Kingston, John. 2005. The phonetics of Athabaskan tonogenesis. In Sharon Hargus & Keren Rice (eds.), *Athabaskan Prosody* (Current Issues in Linguistic Theory), 137–184. Amsterdam: John Benjamins.

Kirby, James. 2014. Incipient tonogenesis in Phnom Penh Khmer: Acoustic and perceptual studies. *Journal of Phonetics* 43(1). https://doi.org/10.1016/j.wocn.2014.02.001.

Knight, Rachael-Anne. 2008. The shape of nuclear falls and their effect on the perception of pitch and prominence: Peaks vs. plateaux. *Language and Speech* 51(3). 223–244. https://doi.org/10.1177/0023830908098541.

Knight, Rachael-Anne & Francis Nolan. 2006. The effect of pitch span on intonational plateaux. *Journal of the International Phonetic Association* 36(1). 21–38. https://doi.org/10.1017/S0025100306002349.

Kossmann, Maarten. 1996. *Grammaire du parler berbère de Figuig (Maroc oriental)*. Paris: Éditions Peeters.

Kossmann, Maarten. 1999. *Essai sur la phonologie du proto-berbère*. Köln: Rüdiger Köppe Verlag.

Kossmann, Maarten. 2012. Berber. In Zygmunt Frajzyngier & Erin Shay (eds.), *The Afroasiatic Languages*, 18–101. Cambridge: Cambridge University Press.

Krahmer, Emiel & Marc Swerts. 2001. On the alleged existence of contrastive accents. *Speech Communication* 34(4). 391–405. https://doi.org/10.1016/S0167-6393(00)00058-3.

Krifka, Manfred. 2011. Questions: An international handbook of natural language meaning. In Klaus von Heusinger, Claudia Maienborn & Paul Portner (eds.), *Semantics*, vol. 2, 1742–1785. Berlin: Mouton de Gruyter.

Ladd, D. Robert. 2006. Segmental anchoring of pitch movements: Autosegmental association or gestural coordination? *Italian Journal of Linguistics* 18(1). 19–38. http://www.italian-journal-linguistics.com/wp-content/uploads/2.Ladd_.pdf.

Ladd, D. Robert. 2008. *Intonational phonology*. Second edition [1996]. Cambridge: Cambridge University Press.

Ladd, D. Robert, Ineke Mennen & Astrid Schepman. 2000. Phonological conditioning of peak alignment in rising pitch accents in Dutch. *The Journal of the Acoustical Society of America* 107(5). 2685–2696. https://doi.org/10.1121/1.428654.

Ladd, D. Robert & Astrid Schepman. 2003. "Sagging transitions" between high pitch accents in English: Experimental evidence. *Journal of Phonetics* 31(1). 81–112. https://doi.org/10.1016/S0095-4470(02)00073-6.

Ladefoged, Peter. 1972. Three areas of experimental phonetics: Stress and respiratory activity the nature of vowel quality. *Perception and Production of Speech* 15.

Lafkioui, Mena. 2010. La topicalisation en berbère: Formes et structures. In Harry Stroomer, Maarten Kossmann, Dymitr Ibriszimow & Rainer Vossen (eds.), *Etudes berbères V:*

Essais sur des variations dialectales et autres articles (Berber Studies 28). Köln: Rüdiger Köppe Verlag.

Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents* (Cambridge Studies in Linguistics 71). Cambridge: Cambridge University Press.

Laoust, Emile. 2012. *Cours de berbère marocain*. Deuxième Édition [1936, 1920]. Casablanca: Éditions Frontispice.

Lehiste, Ilse. 1970. Suprasegmental features of speech. In Norman J. Lass (ed.), *Contemporary issues in experimental phonetics*, 225–239. New York: Academic Press.

Lenth, Russell V. & Maxime Hervé. 2015. *lsmeans: Least-Squares Means*. R package.

Levi, Susannah V. 2005. Acoustic correlates of lexical accent in Turkish. *Journal of the International Phonetic Association* 35(1). 73–97. https://doi.org/10.1017/S0025100305001921.

Liberman, Mark & Alan Prince. 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8(2). 249–336. https://www.jstor.org/stable/4177987.

Lindström, Eva & Bert Remijsen. 2005. Aspects of the prosody of Kuot, a language where intonation ignores stress. *Linguistics* 43(4). 839–870. https://doi.org/10.1515/ling.2005.43.4.839.

Lux, Cécile. 2014. Focalization process and intonation in Meridional Berber: The case of Tamasheq and Tetserret. *STUF – Language Typology and Universals* 67(1). 113–126. https://doi.org/10.1515/stuf-2014-0009.

Maas, Utz. 2013. Die marokkanische Akzentuierung: Beiträge zur semitischen Dialektologie: Festschrift für Werner Arnold zum 60. Geburtstag. In Renaud Kuty, Ulrich Seeger & Shabo Talay (eds.), *Nicht nur mit Engelszungen*, 223–234. Wiesbaden: Harrassowitz Verlag.

Maas, Utz. N.d. *Marokkanisches Arabisch: Zur Struktur einer Sprache im Werden*. Ms., Universität Graz. https://zentrum.virtuos.uni-osnabrueck.de/utz.maas/Main/MarokkanischesArabischZurStrukturEinerSpracheImWerden.

Maas, Utz & Stefan Procházka. 2012a. Introduction: Moroccan Arabic in typological perspective. *STUF – Language Typology and Universals* 65(4). 321–328. https://doi.org/10.1524/stuf.2012.0020.

Maas, Utz & Stefan Procházka. 2012b. Moroccan Arabic in its wider linguistic and social contexts. *STUF – Language Typology and Universals* 65(4). 329–357. https://doi.org/10.1524/stuf.2012.0021.

Mahjani, Behzad. 2003. *An instrumental study of prosodic features and intonation in modern Farsi (Persian)*. University of Edinburgh Master's thesis.

Maskikit-Essed, Raechel & Carlos Gussenhoven. 2016. No stress, no pitch accent, no prosodic focus: The case of Ambonese Malay. *Phonology* 33(2). 1–37. https://doi.org/10.1017/S0952675716000154.

McClelland, James L. & Jeffrey L. Elman. 1986. The TRACE model of speech perception. *Cognitive psychology* 18(1). 1–86. https://doi.org/10.1016/0010-0285(86)90015-0.

Mettouchi, Amina. 2008. Case-marking, syntactic domains and information structure in Kabyle (Berber). In Zygmunt Frajzyngier & Erin Shay (eds.), *Interaction of morphology and syntax: Case studies in Afroasiatic*, vol. 75 (Typological Studies in Language), 7. John Benjamins.

Michalsky, Jan. 2017. *Frageintonation im Deutschen: Zur intonatorischen Markierung von Interrogativität und Fragehaltigkeit*. Berlin: De Gruyter. https://doi.org/10.1515/9783110538564-001.

Misheva, Anastasia & Michel Nikov. 1998. Intonation in Bulgarian. In Daniel Hirst & Albert Di Cristo (eds.), *Intonation systems: A survey of twenty languages*, 278–287. Cambridge: Cambridge University Press.

Mitchell, Terence Frederick. 1993. *Pronouncing Arabic 2*. Oxford: Oxford University Press.

Möhler, Gregor. 1998. Describing intonation with a parametric model. In *Proceedings of the 5th International Conference on Spoken Language Processing*. Sydney.

Mourigh, Khalid. 2015. *A grammar of Ghomara Berber* (LOT series). The Hague: Holland Academic Graphics.

Naumann, Christfried. 2012. *Acoustically based phonemics of Siwi (Berber)*. Köln: Rüdiger Köppe Verlag.

Nejmi, Hassan. 1993. *Contribution à une étude des processus accentuels de l'arabe marocain de Casablanca: Approche phonétique et phonologique (accent de mot)*. Paris III, Université de la Sorbonne Nouvelle PhD dissertation.

Nespor, Marina & Irene Vogel. 2007. *Prosodic Phonology*. Second edition [1986]. Vol. 28 (Studies in Generative Grammar). Berlin: De Gruyter.

Niebuhr, Oliver & Jarich Hoekstra. 2015. Pointed and plateau-shaped pitch accents in North Frisian. *Laboratory Phonology* 6(3-4). 433–468. https://doi.org/10.1515/lp-2015-0013.

Nolan, Francis & Hae-Sung Jeon. 2014. Speech rhythm: a metaphor? *Philosophical Transactions of the Royal Society B* 369(20130396). https://doi.org/10.1098/rstb.2013.0396.

O'Connor, Joseph D. & Gordon F. Arnold. 1973. *Intonation of colloquial English*. Second edition. London: Longman.

Ouhalla, Jamal. 1999. Focus and Arabic clefts. In Georges Rebuschi & Laurice Tuller (eds.), *The grammar of focus*, 335–359. Amsterdam: John Benjamins.

Owens, Jonathan (ed.). 2013. *The Oxford handbook of Arabic linguistics*. Oxford University Press.

Penchoen, Thomas G. 1973. *Tamazight of the Ayt Ndhir*. Los Angeles: Undena Publications.

Peperkamp, Sharon & Emmanuel Dupoux. 2002. A typological study of stress 'deafness'. In Carlos Gussenhoven & Natasha Warner (eds.), *Laboratory Phonology 7*, 203–240. Berlin: De Gruyter.

Peperkamp, Sharon, Inga Vendelin & Emmanuel Dupoux. 2010. Perception of predictable stress: A cross-linguisic investigation. *Journal of Phonetics* 38(3). 422–430. https://doi.org/10.1016/j.wocn.2010.04.001.

Pierrehumbert, Janet. 1980. *The phonology and phonetics of English intonation*. Massachusetts Institute of Technology PhD dissertation.

Pierrehumbert, Janet & Mary Beckman. 1988. *Japanese tone structure* (Linguistic Inquiry Monographs). Cambridge, MA: MIT Press.

Post, Brechtje. 2000. *Tonal and phrasal structures in French intonation* (LOT series). The Hague: Holland Academic Graphics.

Prieto, Pilar. 2004. The search for phonological targets in the tonal space: H1 scaling and alignment in five sentence-types in Peninsular Spanish. In Timothy Face (ed.), *Laboratory approaches to Spanish phonology*, 29–59. Berlin: De Gruyter.

Prieto, Pilar. 2011. Tonal alignment. In Marc Van Oostendorp, Colin Ewen, Beth Hume & Keren Rice (eds.), *Companion to Phonology*, 1185–1203. Chichester: Wiley-Blackwell.

Prieto, Pilar. 2014. The intonational phonology of Catalan. In Sun-Ah Jun (ed.), *Prosodic Typology II: The phonology of intonation and phrasing*, 43–80. Oxford: Oxford University Press.

Prieto, Pilar, Mariapaola D'Imperio & Barbara Gili Fivela. 2005. Pitch accent alignment in Romance: Primary and secondary associations with metrical structure. *Language and Speech* 48(4). 359–396. https://doi.org/10.1177/00238309050480040301.

R Core Team. 2016. *R: A Language and Environment for Statistical Computing*. Software. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rahmani, Hamed, Toni Rietveld & Carlos Gussenhoven. 2015. Stress 'deafness' reveals lexical marking of stress and tone in the adult grammar. *PLoS ONE* 10(12). https://doi.org/10.1371/journal.pone.0143968.

Reichel, Uwe D. 2010. *Datenbasierte und linguistisch interpretierbare Intonationsmodellierung*. University of Munich PhD dissertation. https://edoc.ub.uni-muenchen.de/12650/1/Reichel_Uwe.pdf.

Remijsen, Bert, Fabienne Martis & Ronald Severing. 2014. The marked accentuation pattern of Curaçao Papiamentu. In Sun-Ah Jun (ed.), *Prosodic Typology II: The phonology of intonation and phrasing*, 302–323. Oxford: Oxford University Press.

Rhiati-Salih, Najib. 1984. *Etude de l'interrogation en arabe marocain*. Université de Paris III, Sorbonne-Nouvelle PhD dissertation.

Rialland, Annie & Stéphane Robert. 2001. The intonational system of Wolof. *Linguistics* 39(5). 893–940. https://doi.org/10.1515/ling.2001.038.

Ridouane, Rachid. 2008. Syllables without vowels: Phonetic and phonological evidence from Tashlhiyt Berber. *Phonology* 25(2). 321–359. https://doi.org/10.1017/S0952675708001498.

Ridouane, Rachid. 2014. Tashlhiyt Berber. *Journal of the International Phonetic Association* 44(2). 207–221. https://doi.org/10.1017/S0025100313000388.

Rietveld, Toni, Joop Kerkhoff & Carlos Gussenhoven. 2004. Word prosodic structure and vowel duration in Dutch. *Journal of Phonetics* 32(3). 349–371. https://doi.org/10.1016/j.wocn.2003.08.002.

Roettger, Timo B. 2017. *Tonal placement in Tashlhiyt: How an intonation system accommodates to adverse phonological environments* (Studies in Laboratory Phonology 3). Berlin: Language Science Press. https://doi.org/10.5281/zenodo.814472.

Roettger, Timo B., Bodo Winter, Sven Grawunder, James Kirby & Martine Grice. 2014. Assessing incomplete neutralization of final devoicing in German. *Journal of Phonetics* 43(1). 11–25. https://doi.org/10.1016/j.wocn.2014.01.002.

Roettger, Timo Benjamin, Anna Bruggeman & Martine Grice. 2015. Word stress in Tashlhiyt: Postlexical prominence in disguise? In *Proceedings of the International Congress of Phonetic Sciences XVIII*. Glasgow, UK.

Roettger, Timo Benjamin & Martine Grice. 2015. The role of high pitch in Tashlhiyt Tamazight (Berber): Evidence from production and perception. *Journal of Phonetics* 51(1). 36–49. https://doi.org/10.1016/j.wocn.2014.12.004.

Roettger, Timo Benjamin, Rachid Ridouane & Martine Grice. 2012. Sonority and syllable weight determine tonal association in Tashlhiyt. In *Proceedings of Speech Prosody 6*. Shanghai.

Rooth, Mats. 1985. *Association with focus*. University of Massachusetts PhD dissertation.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1(1). 75–116. https://doi.org/10.1007/BF02342617.

Sabel, Joachim & Jochen Zeller. 2006. Wh-question formation in Nguni. In *Selected proceedings of the 35th annual conference on African linguistics*, 271–283.

Sadeghi, Vahid. 2011. Acoustic correlates of lexical stress in Persian. In *Proceedings of the International Congress of Phonetic Sciences XVII*, 1738–1741. Hong Kong.

Sadiqi, Fatima. 1997. *Grammaire du berbère*. Paris: Editions l'Harmattan.

Selkirk, Elisabeth O. 1986. *Phonology and syntax: The relationship between sound and structure*. Cambridge, MA: MIT Press.

Shattuck-Hufnagel, Stefanie. 1994. "Stress shift" as early placement of pitch accents: a comment on Beckman and Edwards. In Patricia Keating (ed.), *Papers in Laboratory Phonology III: Phonological structure and phonetic form*, 34–43. Cambridge: Cambridge University Press.

Simons, Gary F. & Charles D. Fennig (eds.). 2017. *Ethnologue: Languages of the World*. Dallas, Texas. http://www.ethnologue.com.

Singmann, Henrik, Ben Bolker, Jake Westfall & F. Aust. 2015. *afex: Analysis of factorial experiments*. R package.

Skoruppa, Katrin, Ferran Pons, Anne Christophe, Laura Bosch, Emmanuel Dupoux, Núria Sebastián-Gallés, Rita Alves Limissuri & Sharon Peperkamp. 2009. Language-specific stress perception by 9-month old French and Spanish infants. *Developmental Science* 12(6). 914–919. https://doi.org/10.1111/j.1467-7687.2009.00835.x.

Sluijter, Agaath & Vincent van Heuven. 1996. Notes on the phonetics of word prosody. In Rob Goedemans, Harry van der Hulst & Ellis Visch (eds.), *Stress patterns of the world*, 233–269. The Hague: Holland Academic Graphics.

Sluijter, Agaath, Vincent van Heuven & Jos Pacilly. 1997. Spectral balance as a cue in the perception of linguistic stress. *Journal of the Acoustical Society of America* 101(1). 503–513. https://doi.org/10.1121/1.417994.

Song, Sanghoun. 2017. *Modelling information structure in a cross-linguistic perspective* (Topics at the Grammar-Discourse Interface 1). Berlin: Language Science Press. https://doi.org/10.5281/zenodo.818365.

Sosa, Juan Manuel. 2003. Wh-questions in Spanish: Meanings and configuration variability. *Catalan Journal of Linguistics* 2. 229–247. https://doi.org/10.5565/rev/catjl.51.

Stoyanova, Marina. 2004. The typology of multiple wh-questions and language variation. In Sylvia Blaho, Luis Vicente & Mark de Vos (eds.), *ConSOLE XII: Proceedings of the conference of the Student Organisation of Linguistics in Europe XII*.

Stumme, Hans. 1899. *Handbuch des Schilhischen von Tazerwalt*. Leipzig: J.C. Hinrichs'sche Buchhandlung.

Stumme, Hans & Albert Socin. 1894. *Der arabische Dialekt der Houwara des Wad Sus in Marokko*. Leipzig: Hirzel.

Svetozarova, Natalia. 1998. Intonation in Russian. In Daniel Hirst & Albert Di Cristo (eds.), *Intonation systems: A survey of twenty languages*, 261–274. Cambridge: Cambridge University Press.

Swerts, Marc. 2007. Contrast and accent in Dutch and Romanian. *Journal of Phonetics* 35(3). 380–297. https://doi.org/10.1016/j.wocn.2006.07.001.

Swerts, Marc, Emiel Krahmer & Cinzia Avesani. 2002. Prosodic marking of information status in Dutch and Italian: A comparative analysis. *Journal of Phonetics* 30(4). 629–654. https://doi.org/10.1006/jpho.2002.0178.

Szalontai, Ádám, Petra Wagner, Katalin Mády & Andreas Windmann. 2016. Teasing apart lexical stress and sentence accent in Hungarian and German. In Christoph Draxler & Felicitas Kleber (eds.), *Tagungsband der 12. Tagung Phonetik und Phonologie im Deutschsprachigen Raum*, 215–218. Munich University.

't Hart, Johan. 1976. Psychoacoustic backgrounds of pitch contour stylization. In *IPO annual progress report*, vol. 11. Eindhoven: IPO.

't Hart, Johan. 1981. Differential sensitivity to pitch distance, particularly in speech. *Journal of the Acoustical Society of America* 69(3). 811–821. https://doi.org/10.1121/1.385592.

't Hart, Johan & Antonie Cohen. 1973. Intonation by rule: A perceptual quest. *Journal of Phonetics* 1. 309–327. https://doi.org/10.1016/S0095-4470(19)31400-7.

't Hart, Johan & René Collier. 1975. Integrating different levels of intonation analysis. *Journal of Phonetics* 3(4). 235–255.

't Hart, Johan, René Collier & Antonie Cohen. 1990. *A perceptual study of intonation: An experimental-phonetic approach to speech melody* (Cambridge Studies in Speech Science and Communication). Cambridge: Cambridge University Press.

Tench, Paul. 1996. *The intonation systems of English*. London: Cassell Academic.

Thomas, Erik R. & Tyler Kendall. 2007. *NORM: The vowel normalization and plotting suite*. Online resource. http://ncslaap.lib.ncsu.edu/tools/norm/.

Trager, George L. & Henry Lee Smith. 1951. An outline of English structure. In *Studies in linguistics: Occasional papers*, vol. 3. Norman, OK: Battenberg Press.

Trubetzkoy, Nikolai Sergeevich. 1969. *Principles of phonology*. [1958, 1939]. Berkeley: University of California Press.

Ultan, Russell. 1978. Some general characteristics of interrogative systems. In *Universals of human language*, vol. 4: Syntax, 211–248. Stanford: Stanford University Press.

Van den Boogert, Nico. 1997. *The Berber literary tradition of the Sous* (Publications of the "De Goeie Fonds" 27). Leiden: Nederlands Instituut voor het Nabije Oosten.

van der Hulst, Harry. 2014a. The study of word accent and stress: Past, present and future. In Harry van der Hulst (ed.), *Word stress: Theoretical and typological issues*, 3–55. Cambridge: Cambridge University Press.

van der Hulst, Harry (ed.). 2014b. *Word stress: Theoretical and typological issues*. Cambridge: Cambridge University Press.

van der Wal, Jenneke. 2016. Diagnosing focus. *Studies in Language* 40(2). 259–301. `https://doi.org/10.1075/sl.40.2.01van`.

Van Heuven, Vincent J. & Vera Faust. 2009. Are Indonesians sensitive to contrastive accentuation below the word level? *Wacana* 11(2). 226–240.

Vanrell, Maria del Mar & Olga Fernández Soriano. 2013. Variation at the interfaces in Ibero-Romance. Catalan and Spanish prosody and word order. *Catalan Journal of Linguistics* 12. 253–282. `https://doi.org/10.5565/rev/catjl.63`.

Varga, László. 2002. *Intonation and stress: Evidence from Hungarian*. Houndmills/New York: Palgrave Macmillan.

Vella, Alexandra. 2007. The phonetics and phonology of wh-question intonation in Maltese. In *Proceedings of the International Congress of Phonetic Sciences XVI*, 1285–1289. Saarbrücken.

Versteegh, Kees. 2014. *The Arabic language*. Second edition. Edinburgh: Edinburgh University Press.

Vicenik, Chad & Sun-Ah Jun. 2014. An Autosegmental Metrical analysis of Georgian intonation. In Sun-Ah Jun (ed.), *Prosodic Typology II: The phonology of intonation and phrasing*, 154–186. Oxford: Oxford University Press.

Vigário, Marina & Sónia Frota. 2003. The intonation of Standard and Northern European Portuguese: A comparative intonational phonology approach. *Journal of Portuguese Linguistics* 2(2). 115–137. `https://doi.org/10.5334/jpl.31`.

Vilkuna, Maria. 1995. Discourse configurationality in Finnish. In Katalin É. Kiss (ed.), *Discourse configurational languages*, 244–268. Oxford: Oxford University Press.

Vycichl, Werner. 2005. *A sketch of Siwi Berber (Egypt)*. Köln: Rüdiger Köppe Verlag.

Warren, Paul. 2016. *Uptalk: The phenomenon of rising intonation*. Cambridge: Cambridge University Press.

Watson, Janet C. E. 2011. Word stress in Arabic. In Marc van Oostendorp, Colin J. Ewen, Elizabeth V. Hume & Keren Rice (eds.), *The Blackwell Companion to Phonology*, 2990–3019. Oxford: Wiley-Blackwell.

Welby, Pauline. 2006. French intonational structure: Evidence from tonal alignment. *Journal of Phonetics* 34(3). 343–371. https://doi.org/10.1016/j.wocn.2005.09.001.

Willis, Erik. 2007. Utterance signaling and tonal levels in Dominican Spanish declaratives and interrogatives. *Journal of Portuguese Linguistics* 5/6. 179–202. https://doi.org/10.5334/jpl.149.

Xu, Yi. 2004a. The PENTA model of speech melody: Transmitting multiple communicative functions in parallel. *Proceedings of From sound to sense* 50. 91–96.

Xu, Yi. 2004b. Transmitting tone and intonation simultaneously — the parallel encoding and target approximation (PENTA) model: With emphasis on tone languages. In *International symposium on Tonal Aspects of Languages*.

Yeou, Mohamed. 2005. Variability of F0 peak alignment in Moroccan Arabic accentual focus. In *Proceedings of Interspeech 2005*, 1433–1436. Lisbon.

Yeou, Mohamed, Mohamed Embarki, Sallal Al Maqtari & Christelle Dodane. 2007. F0 alignment patterns in Arabic dialects. In *Proceedings of International Congress of Phonetic Sciences XVI*, 1493–1496. Saarbrücken.

Yeou, Mohamed, Mohamed Embarki & Sallal Al-Maqtari. 2007. Contrastive focus and F0 patterns in three Arabic dialects. *Nouveaux cahiers de linguistique française* 28. 317–326.

Yip, Moira. 2002. *Tone*. Cambridge: Cambridge University Press.

Zellou, Georgia. 2010. Moroccan Arabic consonant harmony: A multiple causation hypothesis. *Toronto Working Papers in Linguistics* 33(1).

Zimmermann, Malte & Edgar Onea. 2011. Focus marking and focus interpretation. *Lingua* 121. 1651–2670. https://doi.org/10.1016/j.lingua.2011.06.002.

Zuraw, Kie, Kristine M. Yu & Robyn Orfitelli. 2014. The word-level prosody of Samoan. *Phonology* 31(2). 271–327. https://doi.org/10.1017/S095267571400013X.

Żygis, Marzena, Susanne Fuchs & Katarzyna Stoltmann. 2017. Orofacial expressions in German questions and statements in voiced and whispered speech. *Journal of Multimodal Communication Studies* 4(1–2). 87–92.

Curriculum Vitae

Name	Anna Maria Bruggeman
Date of Birth	29.09.1990

2019–	Senior Speech Scientist, Cerence Inc., Aachen
2018–2019	Associate, Department of Language & Linguistic Science, University of York
2014–2018	Wissenschaftliche Hilfskraft, I*f*L Phonetik, Universität zu Köln
2014–2018	DAAD scholarship for PhD research at the Universität zu Köln Title of dissertation: *Lexical and postlexical prominence in Tashlhiyt Berber and Moroccan Arabic* Supervisors: Prof. Dr. Martine Grice and Dr. Sam Hellmuth
2012–2013	M.Phil. Theoretical and Applied Linguistics, University of Cambridge
2008–2011	B.A. Languages and Linguistics, University of York
2002–2008	VWO, Christelijk Gymnasium Utrecht, the Netherlands