Journal of Phonetics ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/phonetics

The role of high pitch in Tashlhiyt Tamazight (Berber): Evidence from production and perception

Timo B. Roettger, Martine Grice *

University of Cologne, Germany

ARTICLE INFO

Article history:
Received 7 April 2014
Received in revised form
9 December 2014
Accepted 11 December 2014

Keywords:
Tashlhiyt Berber
Tamazight
Intonation
Tonal alignment
Pitch register
Echo questions
Contrast

ABSTRACT

The aim of this study was to determine the functional relevance of high pitch in Tashlhiyt Tamazight (Berber) echo questions and contrastive statements in both production and perception. Production data revealed that compared to statements: (a) questions have a higher pitch register, manifested as an overall higher pitch level and a wider pitch span, (b) questions are more often realized with a pitch peak on the final syllable, and that (c) this pitch peak in questions tends to be realised later in the syllable. Despite these tendencies, pitch peak placement involved considerable variability, both within and across speakers. A subsequent perception study, designed to test to what extent these prosodic cues to sentence modality are used in perception, found that (a) pitch register and (b) the placement of peak on a particular syllable both affect listeners' judgments. The more subtle alignment of the peak within the syllable, (c), was not reliable as a cue.

1. Introduction

Cross-linguistically, high pitch is commonly reported to signal two specific functions, interrogativity and contrastivity (Bolinger, 1978; Cruttenden, 1997; Gussenhoven, 2002; Hirst & Di Cristo, 1998; Ohala, 1983, 1984; Ultan, 1978). The present study investigates the intonation of Tashlhiyt, a Tamzight (Berber) language that has been shown to mark both functions using high pitch, but with a great deal of variability in both the frequency and time domains (Grice, Roettger, Ridouane, & Fougeron, 2011; Grice, Roettger, & Ridouane, accepted for publication; Roettger, Ridouane, & Grice, 2013). The present paper reports on two studies, a production study and a perception study, both of which investigate the use of high pitch to express interrogativity and contrastivity.

1.1. Two dimensions of high pitch

High pitch can be described in terms of two dimensions, *scaling* and *timing*. Scaling can be *global* or *local*. Global scaling, also referred to as *pitch register* (after Ladd, 2008), involves phrase-length contours. If both the lowest and highest points of a contour are raised, there is an increase in *pitch level*, if the lowest point is lowered and the highest point is raised, there is an increase in *pitch span*. Local scaling involves a tonal event, such as a rise or a fall, and typically involves an increase in pitch span (also referred to as *pitch excursion*).

High pitch usually involves a rise, culminating in a high pitch *target* that is timed relative to prosodically defined units. In autosegmental phonology, this target is analysed as a H(igh) tone.

The timing of such tones has been described in terms of *phonological association* and *phonetic alignment*. We take association to be discrete, referring to a phonological linking of a tone with a structural element, a tone bearing unit. By contrast, alignment is continuous, referring to the exact position of a *F*0 peak in relation to a landmark in the speech signal, such as an acoustically defined segment boundary (see e.g., Arvaniti, Ladd, & Mennen, 1998; Prieto, van Santen, & Hirschberg, 1995; Silverman & Pierrehumbert,

* Correspondence to: University of Cologne, Herbert-Levin-Str. 6, D-50931 Köln, Germany. Tel.: +49 221 470 7047; fax: +49 221 470 5938.
E-mail address: Martine.grice@uni-koeln.de (M. Grice).

0095-4470/$ - see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.wocn.2014.12.004

1990). Depending on the language, H tones can be lexical or post-lexical. Lexical tones are specified in the lexicon and serve to distinguish word meanings. Postlexical tones are assigned to words at the level of the utterance and serve to express meanings relating to discourse and pragmatics.

Post-lexical tones are typically either pitch accents, associated with tone bearing units (such as the metrically strong syllable of a word, or, in the case of nuclear pitch accents, the phrase), or edge tones. Edge tones are associated with the phrase edge itself and can also be secondarily associated with a tone bearing unit in the phrase (edge tones with a secondary association are referred to as phrase accents, Grice, Ladd, & Arvaniti, 2000; Pierrehumbert & Beckman, 1988). Generally, postlexical tones are used to express a range of functions, including, crucially, signalling sentence modalities, such as declarative vs. interrogative, and highlighting important elements in the discourse in relation to alternatives (contrastivity).

1.2. High pitch marking interrogativity and contrastivity

Both dimensions of high pitch, high pitch register and a high pitch target, have been shown to mark interrogativity. Many languages distinguish questions from corresponding statements by means of scaling, e.g., a higher pitch level. In such cases the entire contour, including all of its low and high tones, exhibits raised *F*0 values (among many others: American English: Hirst & Di Cristo, 1998; Hausa: Inkelas & Leben, 1990; Mandarin Chinese: Shen, 1990; Moroccan Arabic: Benkirane, 1998; and Vietnamese: Brunelle, Ha, & Grice, 2012; cf. Haan, 2002, for an excellent overview). In addition to a raised overall pitch level, questions tend to have a wider pitch span than statements, either realized globally or locally (e.g., Ladd, 2008). For instance, in Moroccan Arabic, a language spoken in Morocco alongside Tashlhiyt, questions have both a higher pitch level and a greater local pitch excursion than statements (Benkirane, 1998).

Apart from scaling differences, a cross-linguistically common pattern found in questions is a sharp final *rise* to a high tone at the end of questions (e.g., information-seeking questions, especially polar questions and counter-expectational echo questions; e.g., Bolinger, 1978). This rise is analysable as a final H boundary tone. However, another pattern has also been found to be relatively common across languages, the *rise fall*. What is important in such contours is the timing of the high target (and, crucially, the rise up to this target). In Palermo Italian, for example, the H tone occurs at a position in the intonation phrase that is structurally salient: the head of the intonation phrase, i.e., it is part of the nuclear pitch accent (Grice, 1995). Thus, the rise is on the syllable with the highest metrical strength, rather than at the edge of the phrase. This rise fall pattern has been observed in other varieties of Italian (for an overview see Grice, D'Imperio, Savino, & Avesani, 2005; Savino, 2012) as well as in other languages such as Bengali (Hayes & Lahiri, 1991), Bulgarian (Grice, Benzemüller, Savino, & Andreeva, 1995), Greek (Arvaniti, 2001; Arvaniti & Ladd, 2009), Hungarian (Gosy & Terken, 1994; Ladd, 1983; Varga, 2002), Moroccan Arabic (Benkirane, 1998), and Russian (e.g., Makarova, 2007). In other languages, the rise does not occur on a fully-fledged accent, but the H tone may have a secondary association to a metrically strong syllable that does not bear an accent (Pierrehumbert & Beckman, 1988). For example, in a number of languages discussed in Grice et al. (2000), e.g., Standard Greek and Romanian, the H tone is analysed as an edge tone of an intonation phrase with a secondary association to a metrically strong syllable following the nuclear accent.

In many languages, rise fall contours similar to those found in questions might also be used to mark contrastivity (in statements, especially in contrastive corrective focus contexts), achieved by means of a pitch accent on the contrasted word (e.g., Bolinger, 1972; Van Heuven, 1994; Ladd, 2008). This pitch accent usually involves a high pitch peak, represented as a high tone.

Thus, both interrogativity and contrastivity can be expressed by means of high pitch, the sole difference between questions and contrastive statements often being one of timing, with later peaks for questions (Gosy & Terken, 1994; D'Imperio & House, 1997; Makarova, 2007). For example, in Neapolitan Italian, polar questions are characterised by a rise to a H tone on the accented vowel, followed by a fall marking the end of the phrase (possibly with an intervening plateau). The contour of a contrastive statement, marked by a narrow focus, is also characterised by a rise to a H tone on the accented vowel followed by a fall. Even though these contours appear to be very similar in certain contexts, they have been described as differing in the alignment of the high tone, which reaches its target later in the accented vowel in questions than in contrastive statements.

Both of the dimensions of high pitch discussed in relation to production above (scaling and timing) have also been found to be relevant for perceptually disambiguating statements from questions. In their seminal study, Hadding-Koch and Studdert-Kennedy (1964) manipulated *F*0 values of Swedish utterances with regard to three different positions: the accent peak, the post accentual low and the end of the phrase. Participants judged whether the sentence presented was a question or a statement. Results indicated that listeners were more likely to rate the sentence as a questions where there was higher *F*0 at all three reference points. This has been taken as evidence that listeners use scaling to distinguish sentence modality. Local scaling effects on the interpretation of sentence modality have been shown for a variety of languages, such as Bari Italian (Savino & Grice, 2007), Hungarian (Gosy & Terken, 1994), Russian (Makarova, 2007), and Swedish (House, 2003). These studies consistently showed that greater pitch excursions of the rise fall are more frequently perceived as questions than statements.

In addition to scaling, most of the above cited studies have also shown that listeners are able to use the timing of the high pitch target as a cue to sentence modality. Gosy and Terken (1994) showed that in Hungarian later peaks are more often perceived as questions. Similar results were found for Neapolitan Italian, where this timing difference has been shown to be crucial for distinguishing sentence modality when the accented word is final in the phrase (D'Imperio & House, 1997, see also House, 2003, for Swedish and Makarova, 2007, for Russian).

Please cite this article as: Roettger, T. B., & Grice, M. The role of high pitch in Tashlhiyt Tamazight (Berber): Evidence from production and perception. *Journal of Phonetics* (2015), http://dx.doi.org/10.1016/j.wocn.2014.12.004

To sum up, high pitch, manifested as high pitch register or a high tone, is a common acoustic cue to sentence modality across languages, and can be exploited by listeners to disambiguate questions from statements.

1.3. Interrogativity and contrastivity in Tashlhiyt

This paper investigates the functional relevance of high pitch in Tashlhiyt, one of three main Tamazight (Berber) dialects spoken in Morocco. The language is renowned for its particularly long consonantal sequences (Clements, 1997; Hyman, 1985; Zec, 1995) and typologically rare syllable structures (Dell & Elmedlaoui, 1985, 1988, 2002; Ridouane, 2008). Recently, aspects of higher prosodic structure in the language have become the focus of experimental research. This includes studies on metrical structure (Gordon & Nafi, 2012) and intonation (Grice et al., 2011, accepted for publication; Roettger et al., 2013).

Early descriptions of Tashlhiyt have claimed that stress is likely to be a property of constituents larger than words (Applegate, 1958; Dell & Elmedlaoui, 2002). However, Gordon and Nafi (2012) argue for word level prominence on the basis of a study investigating three acoustic correlates of potential phonological prominence (*F*0, intensity, and duration). Word-final syllables were longer and had a greater intensity than penultimate syllables, especially in phrase final position. Interestingly, Gordon and Nafi also found that phrase-final syllables were associated with higher *F*0 than other syllables, and interpreted this as a reflex of a high pitch accent associated with a metrically strong syllable. Grice et al., accepted for publication argue that, given the variation in *F*0 peak location within the words in their data set, discussed in more detail below, metrical strength could only be at a postlexical level.

Grice, Ridouane, and Roettger (Grice et al., accepted for publication; Roettger et al., 2013) have shown that polar questions (marked morphosyntactically with a question preverb) and statements (contrastive, corrective focus) show very similar intonation contours, characterised by a rise to a *F*0 peak followed by a fall. Which syllable the *F*0 peak was located on was determined by a number of competing factors, resulting in probabilistic distributions (Grice et al., accepted for publication): There was a general preference for the *F*0 peak to be located on the final syllable, but at the same time there was a preference for the *F*0 peak to be located on more sonorous elements and heavy syllables. Importantly, questions revealed a stronger preference to be realized with an H tone on the final syllable. Moreover, if the final word did not contain sonorous nuclei, as in /inːa tkʃf/ 'Did he say 'it dried'?', the *F*0 peak was frequently realized on the preceding word (here: /inːa/). Despite probabilistic tendencies, tonal placement was to some degree freely alternating, since the same speaker often produced both penultimate and final *F*0 peaks in different repetitions of either of the sentence modalities, even with the same target word. This is consistent with an account in which the H tone is associated with a constituent higher than the prosodic word, since it has been demonstrated that, if the target word contains no sonorants, the tone can be placed on a syllable outside the target word (Grice et al., accepted for publication).

Generally, we analysed the H tone in both questions and statements as an edge tone of a prosodically defined constituent with a secondary association to a tone bearing unit, a syllable with a sonorant nucleus. Such an association might lead to increases in intensity and duration (Katsika, Krivokapić, Mooshammer, Tiede, & Goldstein, 2014), explaining Gordon and Nafi's observations on prominence asymmetries between penultimate and final syllables. However, the differential representation of the H tone in questions and statements, as well as the nature of the constituent with which it is associated, is still unclear.

In addition to differences in the syllable on which the H tone was located, interpreted as phonological association, questions and statements differed in the position of the *F*0 peak within the syllable. *F*0 peaks in questions were aligned later in the syllable than in statements (Roettger et al., 2013). With regard to scaling, questions were produced with a higher overall pitch level and a wider pitch excursion to the *F*0 peak than statements (Grice et al., accepted for publication). The latter observations were impressionistic in nature and thus require quantitative confirmation.

1.4. The present study

Some of the variability in the intonation of questions and statements in Tashlhiyt might reflect the redundant nature of *F*0 peak location for distinguishing between the two sentence modalities. Polar questions are marked by an initial preverb /is/. This morphosyntactic marker could already be sufficient for indicating sentence modality. However, the language also has an echo question used in counter-expectational contexts (/inːa tndm/ 'He said 'she regretted'?'). This type of question exhibits the same intonation contour as polar questions. Since there is no preverb, it differs from corresponding contrastive statements in its prosody alone. The present study compares the tonal manifestation of this contrast.

The objectives of the present paper involve both production and perception: Starting with production, we investigate whether the difference between echo questions and contrastive statements is reflected in scaling differences and the timing of the H tone (Experiment 1). Based on previous findings on Tashlhiyt (Grice et al., accepted for publication; Roettger et al., 2013), as well as cross linguistic evidence, we advance the following hypotheses.

H1a. Questions exhibit a higher pitch register than corresponding statements, manifested as a higher pitch level and/or wider pitch span.

H1b. Questions exhibit a rise to a *F*0 peak on the final syllable more often than corresponding statements.

H1c. The rise to a *F*0 peak in questions reaches its high pitch target later in the syllable than in statements.

For perception, we test how far speakers of the language use these acoustic cues to distinguish between the two sentence modalities (Experiment 2). We advance the following hypotheses:

Please cite this article as: Roettger, T. B., & Grice, M. The role of high pitch in Tashlhiyt Tamazight (Berber): Evidence from production and perception. *Journal of Phonetics* (2015), http://dx.doi.org/10.1016/j.wocn.2014.12.004

H2a. Contours exhibiting a higher pitch register are more often perceived as questions than statements.

H2b. Contours exhibiting a rise to a *F*0 peak on the final syllable are more often perceived as questions than statements.

H2c. Rises that reach the *F*0 peak later in the syllable are more often perceived as questions than statements.

2. Production data

In the following section, we analyse production data from a read speech corpus, collected on a field trip in Agadir, Morocco. We compare echo questions to contrastive statements with regard to global pitch register (Section 2.2.1) and timing of the rise to a high pitch target. The latter is analysed in terms of both discrete co-occurrence of *F*0 peaks with specific syllables (Section 2.2.2), and details of alignment of the *F*0 peak within the syllable (Section 2.2.3).

2.1. Method

2.1.1. Participants

Ten native speakers of Tashlhiyt (five males, five females, mean age = 22) were recorded. All live in Agadir, Morocco, are fluent in Moroccan Arabic and have basic commands of French. All of them had normal or corrected-to-normal vision. None reported on any hearing impairments. Subjects were paid for their participation.

2.1.2. Speech materials and procedure

The present production data are part of a large scale corpus of read speech. Here we shall focus on a subset of the data which is relevant for addressing the present research question, including six disyllabic and two trisyllabic target words. All stimuli were entirely voiced. We aimed to reduce some of the variability found in previous experiments, by selecting target words with predominantly vowels as the syllable nucleus and light syllables in word final position (cf. Table 1).

Participants were seated in front of a computer screen and read out orthographically presented materials containing the target words as presented in (1)–(4) in mini monologues. The materials were presented in one of the scripts used for the language, the Latin script, commonly used alongside the traditional Tamazight script, Tifinagh.

(1) ⟨is inna TARGET A?⟩ 'Did he say TARGET A?'
(2) ⟨ur inna TARGET A.⟩ 'He did not say TARGET A.'
(3) ⟨inna TARGET B.⟩ 'He said TARGET B.'
(4) ⟨manik? inna TARGET B? Irwas.⟩ 'How? He said TARGET B? It looks like it.'

Recordings were made in a quiet room at the Ibn Zohr University in Agadir. The production data was recorded using a Marantz PMD 670 solid-state recorder at a sampling rate of 44.1 kHz, and an AKG C420 III head-mounted microphone. Before recording began, participants were asked to read a word list containing all of the target words to ensure that they were familiar with the words and their meaning. Monologues were presented in random order. In (1) the target word is in a polar question. In (2) the same target word is in a negative assertion. Because of the preceding negation, a different target word is explicitly contrasted in (3). (This type of contrast can also be analysed as a corrective focus.) Finally, in (4), the proposition in the contrastive statement is called into question in the counter-expectational echo question. The polar question and the negative assertion are not analysed further here due to space limitations. In the present study we concentrate on the contrastive statement in (3) and the echo question in (4), both of which have the same segmental structure, and are thus only distinguishable – if at all – by prosodic means. In both contexts the target word is located at the right edge of the Intonational Phrase to ensure comparison to Grice et al. (accepted for publication) and Roettger et al. (2013). Each target word appeared in both sentence modalities at least once. Two target words appeared twice in each sentence modality, resulting in 10 tokens per sentence modality per speaker (overall $n = 200$).

Table 1
List of stimuli in phonemic form and respective translations. Dots mark syllable boundaries.

Target word	Translation
ba.ba	My father
da.ri	In my house
di.ma	Always
il.di	He pulls
i.min.nun	Your mouths
in.da	He shakes (milk)
ir.ba	He carries one's back
ma.na.gu	When
Ri.la	Now
u.dm	Face

Please cite this article as: Roettger, T. B., & Grice, M. The role of high pitch in Tashlhiyt Tamazight (Berber): Evidence from production and perception. *Journal of Phonetics* (2015), http://dx.doi.org/10.1016/j.wocn.2014.12.004

2.1.3. Analyses

All acoustic material was manually annotated employing the following labelling criteria: We identified segment boundaries in the target word in the acoustic waveform. To do this, we displayed an oscillogram and a wide-band spectrogram simultaneously. All segmental boundaries of vowels and consonant were labelled at abrupt changes in the spectra at the time at which the closure was formed or released: this was the case for the nasals, the laterals (especially in the spectra for the intensity of higher formants) and the fricatives (at random noise patterns in the higher frequency regions).

For analyses of the *F*0 contour, the minimum and maximum *F*0 values and the mean *F*0 values were extracted automatically for each word in each utterance via Praat (Boersma & Weenink, 2013). Pitch level was operationalized as *F*0 *mean* of the utterance; pitch span was operationalized as *F0 range*, i.e., the difference between the maximum pitch value (corresponding to the *F*0 peak) and the minimum pitch value on /inːa/. We did not take the lowest pitch value of the whole utterance into account here because there is an asymmetry of phrase final truncation patterns across sentence modalities, there tending to be more pronounced falls in statements. Since /inːa/ is characterised by relatively flat *F*0 over the course of the word in both sentence modalities, it was better suited for comparison.[1] Taken together, pitch level and pitch span account for overall pitch register differences.

Additionally, the location of the rise to the *F*0 peaks on specific syllables was judged by ear and subsequently confirmed by visual inspection of the *F*0 contour and spectrogram provided by Praat. Since the syllable with which the rise co-occurred was more prominent, i.e., impressionistically louder and longer, the decision as to the location of the rise was straightforward. A subset of the data (20%) was annotated by a second transcriber blindly (without visual inspection of the *F*0 contour and without knowing the sentence modality). Agreement was 98%. In what follows, *peak position* refers to the syllable on which the *F*0 peak occurs.

Around the rise contour area we identified a local maximum in the *F*0 contour by hand. In line with Roettger et al. (2013), the relative position of the *F*0 peak in the syllable on which it is located is calculated as a time normalised proportional value, i.e., the interval between the *F*0 peak and the acoustic onset of the syllable divided by the duration of the syllable. In other words, a value of 0 indicates that the *F*0 peak is realized at the onset of the syllable, a value of 1 indicates that it is realized at the offset of the syllable. In the rest of the paper *peak alignment* refers to this value.

2.1.4. Statistics

We excluded 23 data points due to dysfluent productions, hesitations or reading errors.[2] All remaining data were analysed with generalised linear mixed models, using R (R Core Team, 2012) and the package lme4 (Bates, Maechler, Bolker, & Walker, 2014). To analyse continuous variables, we used models with Gaussian error distributions. To analyse categorical variables, we used mixed logistic regression models. We included a term for random intercepts as well as random slopes for speakers for the fixed effects under investigation (Barr, Levy, Scheepers, & Tily, 2013). To determine *p*-values for the main effects of factors/interactions between factors, a model including the fixed factor/interaction of interest was compared to the same model with no fixed factor/no interaction via Likelihood Ratio Tests (LRT).

2.2. Results and discussion

Before we look at contexts in which the target words are in final position and a potential tonal contrast between questions and statements could be neutralised, we will briefly provide evidence that sheds light on the nature of the intonation contours of questions and contrastive statements. Data collected along with the above described corpus demonstrate that contrastive statements are characterised by a rise fall at the right edge of the focused constituent, while questions (both polar and echo questions) are characterised by a rise fall on the right edge of the intonation phrase. Fig. 1 illustrates this contrast: In the question (1a) the rise is located at the end of the phrase, i.e., on the final syllable of /abadan/ ‘always’. In the contrastive statement (1b), on the other hand, the rise is located at the end of the focused constituent/corrective element, i.e., here on the final syllable of /iminːun/ ‘your mouths’.

Thus, the rise fall appears to be intonation phrase final in questions and focused constituent final in contrastive statements. We will return to the analysis of this difference in the general discussion.

Now we turn to contexts in which this contrast is in intonation phrase final position. We first look at the contours descriptively. Figs. 2 and 3 show the aggregated *F*0 contours for statements and questions produced with either a *F*0 peak on the final syllable (Fig. 2) or a *F*0 peak on the penultimate syllable (Fig. 3). In line with Grice et al., accepted for publication and Roettger et al. (2013), both sentence types were characterised by relatively flat *F*0 over the course of /inːa/ and a rise to a *F*0 peak somewhere on the target words followed by a fall in *F*0 that might or might not be truncated. This truncation was speaker dependent and appeared to be related to the position of the *F*0 peak, i.e., the fall is heavily truncated in the case of a final *F*0 peak in questions (Fig. 2).

In line with previous impressionistic observations, we observe a strong difference between statements and questions in terms of pitch register. In general, questions had an overall higher pitch level and a greater pitch span with a steeper rise to the *F*0 peak. In fact, speakers frequently changed phonation type towards the end of the phrase producing a falsetto-like phonation in questions. This was acoustically manifested by strikingly high *F*0 values up to 700 Hz and a sudden shift into low vibrational amplitude (Laver, 1994).

[1] Even if we define *F*0 *range* as the difference between maximum pitch value (corresponding to the *F*0 peak) and minimum pitch value of the utterance, the reported statistical results hold.

[2] Note that Tashlhiyt speakers are not used to reading in their language. Thus, some speakers had problems with reading the Latin script during the familiarisation phase.

Please cite this article as: Roettger, T. B., & Grice, M. The role of high pitch in Tashlhiyt Tamazight (Berber): Evidence from production and perception. *Journal of Phonetics* (2015), http://dx.doi.org/10.1016/j.wocn.2014.12.004

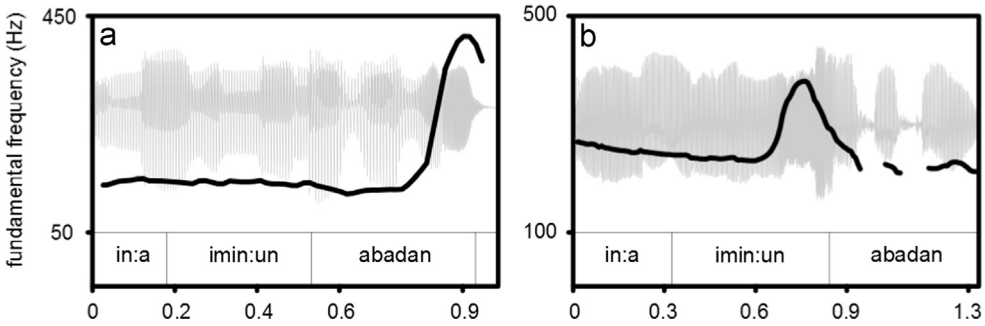

Fig. 1. Representative waveforms and *F*0 contours (a) of the echo question /in:a imin:un abadan/ 'he said 'your mouths' always' with the rise fall on the final syllable of /abadan/ produced by M1; and (b) of the corresponding contrastive statement with a rise fall on the final syllable of the target /imin:un/ produced by F1.

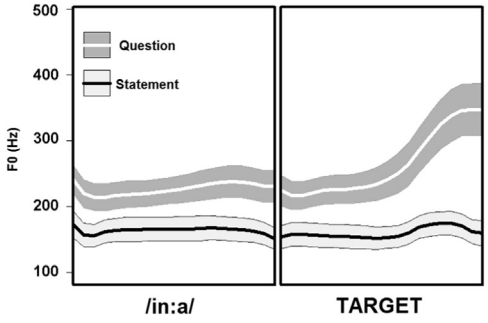

Fig. 2. Mean *F*0 contours (bold line) and standard error (semi-transparent shell) for contrastive statements ($n=25$) and echo questions ($n=78$) produced with a high tone on the final syllable (pooled across speakers and target words; word durations are time normalised).

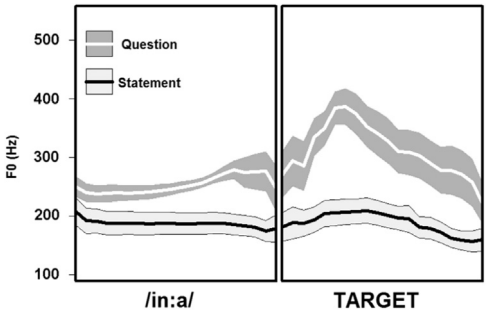

Fig. 3. Mean *F*0 contours (bold line) and standard error (semi-transparent shell) for contrastive statements ($n=58$) and echo questions ($n=8$) produced with a high tone on the penultimate syllable (pooled across speakers and target words; word durations are time normalised).

2.2.1. Pitch register

For statistical analysis we submitted *F*0 *mean* and *F*0 *range* to linear mixed effects models. We included the test variable SENTENCE MODALITY (statement vs. question) and the control variable GENDER as fixed effects and a term for random intercepts as well as random slopes for speakers for SENTENCE MODALITY. There was a significant effect of SENTENCE MODALITY on both *F*0 *mean* and *F*0 *range*, such that statements have a significant lower *F*0 mean (180 Hz) and smaller *F*0 range (89 Hz) than questions (*F*0 mean = 254 Hz; *F*0 range = 201 Hz) (*F*0 mean: SE = 13 Hz, $\chi^2(1)=16.9$, $p<0.0001$; *F*0 range: SE = 23 Hz, $\chi^2(1)=13.7$, $p<0.001$) (cf. Figs. 4 and 5). In sum, questions have a raised pitch level and wider pitch span than statements.

2.2.2. Peak position

The *F*0 peak corresponding to the H tone is placed on specific syllables rather than being evenly distributed across the syllables of the word. In other words, the distribution of the *F*0 peak is clearly bimodal, as shown in the histogram in Fig. 6, indicating that it was *either* on the penult *or* the final syllable, but rarely in between. Similar to previous observations (Grice et al., accepted for publication; Roettger et al., 2013), there is a certain degree of variation as to which syllable the *F*0 peak is located on. Fig. 7 shows two realisations of /in:a imin:un/ 'he said "your mouths"'. The left panel (7a) shows a realization with the rise to *F*0 peak on the penultimate syllable and the right panel (7b) shows a realization with the rise to *F*0 peak on the final syllable. In line with earlier work this is

Please cite this article as: Roettger, T. B., & Grice, M. The role of high pitch in Tashlhiyt Tamazight (Berber): Evidence from production and perception. *Journal of Phonetics* (2015), http://dx.doi.org/10.1016/j.wocn.2014.12.004

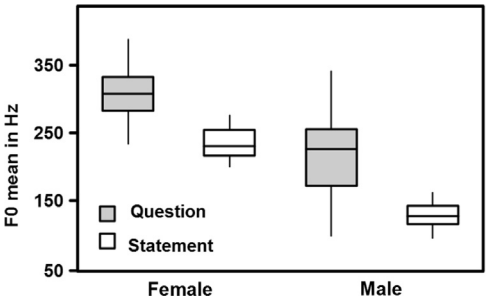

Fig. 4. Mean fundamental frequency as a function of sentence modality and gender. Inside each plot, the boxes indicate the inter-quartile range (IQR), the range between the first and third quartile. The solid line indicates the median. The whiskers indicate the range, up to 1.5 times the IQR away from the median.

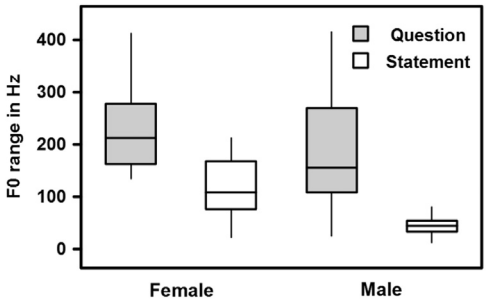

Fig. 5. Mean fundamental frequency range as a function of sentence modality and gender. *F*0 range was defined as the maximum pitch value (corresponding to the *F*0 peak) – the lowest pitch value on /inːa/. Inside each plot, the boxes indicate the inter-quartile range (IQR), the range between the first and third quartile. The solid line indicates the median. The whiskers indicate the range, up to 1.5 times the IQR away from the median.

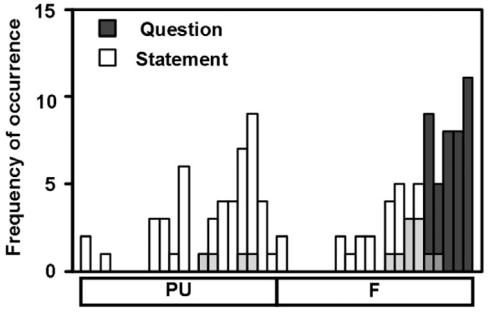

Fig. 6. Raw counts of occurrence of *F*0 peak relative to the penult (PU) and final syllables (F) (normalised duration for penult and final syllables – divided into 20 equal bins for each syllable). Grey bars indicate overlap between questions and statements.

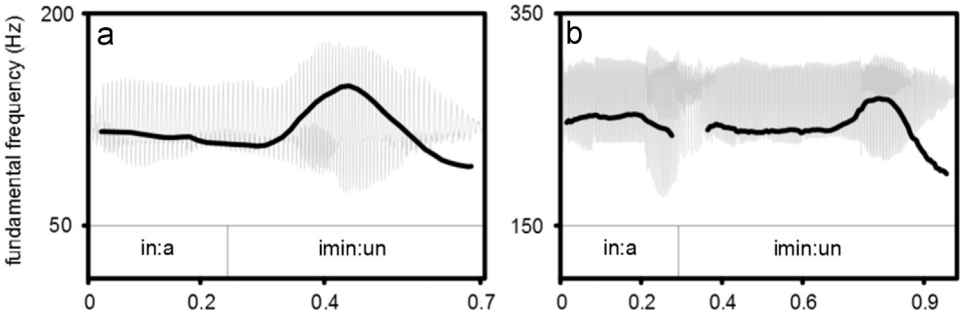

Fig. 7. Representative waveforms and *F*0 contours of the contrastive statements /inːa iminːun/ "he said 'your mouths'" (a) with the rise on the penultimate syllable produced by speaker M1; and (b) the rise on the final syllable produced by speaker F2.

interpreted as free alternation of tonal association. Interestingly, however, the rise is not found on the antepenultimate syllable in trisyllabic words (e.g., /i̲minːun/ 'your mouths').[3]

Despite this alternation, tonal association is largely determined by sentence modality, with 30% *F*0 peaks on the final syllable in statements and 90% on the final syllable in questions (cf. Fig. 6, cf. Appendices A and B). To test the inferential nature of this difference, we

[3] In the following analysis, we excluded the target word /mɑnɑgu/ 'when', because it exhibited unusual behaviour with regard to tonal placement throughout the entire corpus. We assume that its status as a wh-question word might have elicited categorically different contours.

Please cite this article as: Roettger, T. B., & Grice, M. The role of high pitch in Tashlhiyt Tamazight (Berber): Evidence from production and perception. *Journal of Phonetics* (2015), http://dx.doi.org/10.1016/j.wocn.2014.12.004

performed a mixed logistic regression model on PEAK POSITION (penultimate (PU) vs. final (F)). The test variable SENTENCE MODALITY was included as a fixed effect and both speakers and words were included as terms for random intercepts. Additionally, we included a random slope for speakers for the factor SENTENCE MODALITY. There was a significant effect of SENTENCE MODALITY on *F*0 peak placement, such that statements are estimated to have a significantly lower likelihood of bearing a final *F*0 peak than questions ($\chi^2(1)=16.9$, $p<0.0001$).

2.2.3. Peak alignment

Having explored *F*0 peak position in terms of discrete association to a specific syllable, we now turn to a more gradual alignment of the *F*0 peak with the segmental material. In line with Roettger et al. (2013), the relative position of the *F*0 peak in the syllable on which it is located is stated as a relative value. On average, *F*0 peaks occurred in the second half of the syllable (0.74). However, the *F*0 peak appears to be realized slightly later in questions (0.82) than in statements (0.68). We ran a linear mixed effects model on the relative *F*0 peak position in the syllable. The test variables SENTENCE MODALITY and PEAK POSITION (penultimate (PU) vs. final (F)) were included as fixed effects and both speakers and words were included as terms for random intercepts. Additionally, we included a random slope for speakers for the factor SENTENCE MODALITY. The effect of SENTENCE MODALITY on *F*0 peak alignment in the syllable was significant (SE = 0.05, $\chi^2(1)=16.1$, $p<0.0001$). There was also a significant effect of PEAK POSITION on *F*0 peak alignment (SE = 0.04, $\chi^2(1)=6.5$, $p<0.05$), such that *F*0 peaks are estimated to be realized later on penultimate syllables than on final syllables. Note that any of these differences cannot be driven by any potential durational differences between statements and questions because we use time normalised values.

There was no interaction between SENTENCE MODALITY and PEAK POSITION, in other words the effect of sentence modality holds regardless of whether the *F*0 peak is realized on the final or penultimate syllable. However, looking at the distribution of the *F*0 peaks in Fig. 6 descriptively, *F*0 peaks in questions appear to be strongly clustered at the right edge of the (final) syllable, while *F*0 peaks in statements are more spread out within the syllable, although in the penult they tend to occur near the end.

To sum up, this production study demonstrated that echo questions and contrastive statements in Tashlhiyt mainly differ according to the following three tonal cues confirming hypotheses H1a–H1c: Questions exhibit a higher pitch register than corresponding statements, manifested as a higher pitch level and wider pitch span. Moreover, questions exhibit a rise fall on the final syllable more often than corresponding statements. In fact, statements are realized predominantly with the rise fall on the penultimate syllable. And this rise fall reaches its high pitch target later in the syllable in questions than the rise fall in statements.

These results extend earlier work on Tashlhiyt (Grice et al., accepted for publication; Roettger et al., 2013) by demonstrating that the intonational form of counter-expectational echo questions resembles that of unbiased polar questions which, in turn, differs from contrastive statements in the tonal parameters discussed. The present study extends previous findings by demonstrating quantitatively that there is a very consistent pitch register difference between echo questions and statements, manifested by both pitch level and pitch span. Moreover, we confirm that the *F*0 peak, analysed as a H tone, does not have a fixed association to a particular syllable. For both questions and contrastive statements, the H tone can be realized on the penultimate or final syllable of the target word. This highly variable tonal association is particularly striking, since the functional load to distinguish the two sentence modalities rests on the intonation there being no morphological marking.

Despite this variability, we observed more consistent marking of sentence modality than Grice et al. (accepted for publication). They reported on a strong general preference to place the *F*0 peak on the final syllable regardless of sentence modality. This pattern was not observed here. *F*0 peak placement was highly determined by sentence modality, with penultimate *F*0 peaks preferred in statements and final *F*0 peaks preferred in questions. In fact there are only 11 instances of a penultimate *F*0 peak in questions, of which 10 are on the target word /u.dṃ/ 'face'. As reported in Grice et al. (accepted for publication), sonority of the syllable nucleus is a strong factor determining *F*0 peak placement, so /u.dṃ/ attracts a penultimate *F*0 peak because the penultimate nucleus is more sonorous (vowel) than the final nucleus (nasal). So the question arises as to why the production behaviour in this study is less variable than that reported by Grice et al. (accepted for publication). One important difference between the present study and earlier studies is the nature of the stimulus material. While Grice et al. used target words with different phonotactic properties that caused some of the variability (due to syllable weight and sonority), the present study focused on investigating the impact of sentence modality only, thus keeping syllable weight and sonority relatively constant (for most target words there were two vowels in syllable nucleus position and most target words had a light final syllable).

We now turn to the contribution of the identified tonal cues to the discriminability of the contrast between statements and echo questions in a perception experiment.

3. Perception experiment

We have shown that statements and questions differ according to pitch register, peak placement, and peak alignment (cf. H1a–c). The following perception study investigates the impact of these three factors on perception.

3.1. Method

3.1.1. Participants

Nine native speakers of Tashlhiyt (four males, five females, mean age = 21) participated in our experiment. None of them participated in the production experiment. All live in Agadir, Morocco, and also speak Moroccan Arabic and, to a lesser extent, French.

Please cite this article as: Roettger, T. B., & Grice, M. The role of high pitch in Tashlhiyt Tamazight (Berber): Evidence from production and perception. *Journal of Phonetics* (2015), http://dx.doi.org/10.1016/j.wocn.2014.12.004

3.1.2. Speech materials and procedure

We used stimuli with a resynthesized *F*0 contour in order to control for pitch register and *F*0 peak placement. As base stimuli we used four short phrases /inːa baba/, /inːa bibi/, /inːa dima/, and /inːa ːila/ '3*ms-say* ('*father, turkey, always, now*')' produced by a phonetically trained native speaker of Tashlhiyt. The produced phrases highly resembled the produced phrases in the production phrase. For each phrase, the speaker produced two contours corresponding to two different peak placements. This resulted in two sets of stimuli. One set contained the rise and *F*0 peak on the final syllable (F) and the other set contained a rise to an *F*0 peak on the penultimate syllable (PU). The speaker was instructed to produce the sets in the same register.

Both sets were resynthesized using PSOLA in Praat (Boersma & Weenink, 2013). *F*0 was manipulated resulting in two different pitch register conditions: The *low* register condition started with a baseline of 130 Hz the *high* register condition started 4 semitones higher (~164 Hz).

Generally, *F*0 was manipulated to start rising at the offset of /inːa/ towards two different *F*0 maximum locations for each set: In the *early* peak condition *F*0 reached its maximum at 1/3 of the way into vowel (penult in set PU and final in set F), in the late peak condition *F*0 reached its maximum at 2/3 of the vowel. Note that the alignment differences exceed those typically found in actual production to maximise a potential effect. The maximum *F*0 value was 4 semitones higher than the baseline (164 Hz and 206 Hz, respectively). These values resembled typical rise excursions of the speaker. After reaching its maximum, *F*0 fell towards the baseline located at the end of the target word.

These manipulations resulted in 32 stimuli (4 target words∗2 pitch registers (low vs. high)∗2 peak placements (penultima vs. final) ∗2 peak alignments (early vs. late)) (cf. Fig. 8).

Participants were seated in front of a computer screen in a quiet room at the Ibn Zohr University in Agadir. They were told that they were going to listen to a robot which does well in speaking Tashlhiyt, however, struggles with producing the difference between statements and questions. They were asked to decide whether they would consider the produced sentences as a statement or a question by pressing one of two buttons.

The experiment was controlled using Superlab (Abboud, 1991). At the beginning of each trial, a fixation stimulus consisting of a '+' was presented in the centre of the screen for 1500 ms. Following this, two sentences appeared on the right and left side of the screen. Participants had to press a left or right button on the computer keyboard. On one side the statement was displayed in blue (e.g., ⟨inna baba !⟩), on the other side the question was displayed in red (e.g., ⟨inna baba ?⟩). Both were presented in the Latin alphabet. The position of question vs. statement was kept constant within participants, but was counterbalanced across participants. After response delivery, a blank screen appeared for 500 ms.

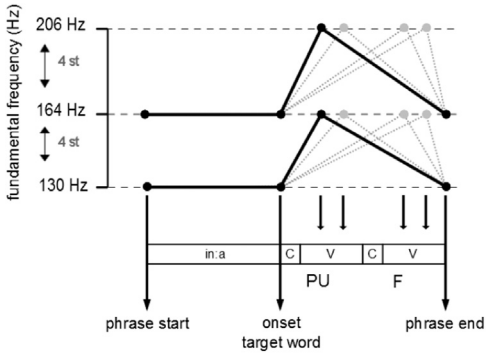

Fig. 8. Schematized manipulation condition displaying the differences of pitch register, peak placement (penultimate = PU, final = F) and peak alignment. Small arrows indicate early and late alignment within the syllable.

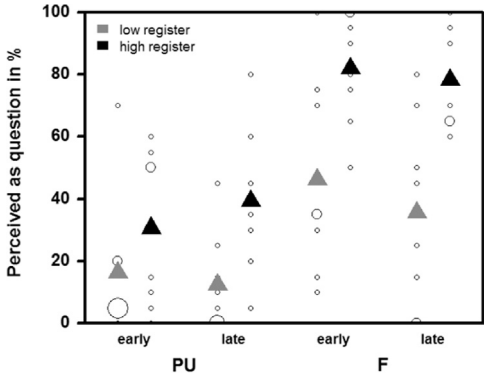

Fig. 9. Ratings as a function of pitch register (low vs. high), peak placement (PU vs. F) and peak alignment (early vs. late). Filled triangles give the grand mean, solid circles give means for participants; size of solid circles indicate number of overlapping data points (same values for multiple participants).

Please cite this article as: Roettger, T. B., & Grice, M. The role of high pitch in Tashlhiyt Tamazight (Berber): Evidence from production and perception. *Journal of Phonetics* (2015), http://dx.doi.org/10.1016/j.wocn.2014.12.004

Each participant started with a training session, in which all combinations of pitch registers, peak placements and peak alignments were presented once realized on the four target words randomly. In the test blocks, each target word in each of the manipulation condition was repeated five times and presented in randomized order resulting in a total of 160 data points per speaker.

3.1.3. Statistics

All data were analysed with generalised linear mixed models. To analyse responses categorically, we used mixed logistic regression models with "RATING" (question or statement) as the dependent measure. As fixed effects we included PITCH REGISTER (low vs. high), PEAK PLACEMENT (PU vs. F), PEAK ALIGNMENT (early vs. late), TARGET WORD and mean-centered REPETITION. We included a term for random intercepts for participants, which quantifies by-participant variability as well as random slopes for the fixed effects PITCH REGISTER, PEAK PLACEMENT AND PEAK ALIGNMENT for participants. To analyse reaction times (RTs, measured from the offset of the audio stimulus), we used models with Gaussian error distribution with response latency as dependent variable. As fixed effects we included the two-way interactions of RATING and PITCH REGISTER (low vs. high), RATING and PEAK PLACEMENT (PU vs. F) and RATING and PEAK ALIGNMENT (early vs. late), and the fixed effect TARGET WORD in addition to mean-centered REPETITION. We included a term for random intercepts for participants as well as random slopes for the fixed effects RATING, PITCH REGISTER, PEAK PLACEMENT and PEAK ALIGNMENT for participants, respectively. As described in Section 2.1.4, models including the fixed factor/interaction of interest were compared to the same model with no fixed factor/no interaction via Likelihood Ratio Tests (LRT) to determine *p*-values.

3.2. Results and discussion

Rating: Fig. 9 depicts the rating results according to the factors PITCH REGISTER, PEAK PLACEMENT and PEAK ALIGNMENT. Overall, participants rated the stimuli as corresponding to questions in 43% of the cases. There was a significant effect of pitch register ($\chi^2(1)=7.5$, $p<0.01$), such that items with a high pitch register were significantly more often rated as a question (58% vs. 28%). There was a significant effect of PEAK PLACEMENT ($\chi^2(1)=8.9$, $p<0.005$) as well, such that items with the *F*0 peak on the final syllable were significantly more often rated as a question than statements (61% vs. 25%). PEAK ALIGNMENT did not have a significant effect on ratings. *F*0 peaks early in the respective syllables were rated as corresponding to questions comparably as often as *F*0 peaks late in the respective syllable (44% vs. 41%) ($\chi^2(1)=1.16$, $p=0.28$). As can be seen in Fig. 9, there was no apparent interaction of TONAL ASSOCIATION AND PITCH REGISTER. Rather, the effects are additive, with a final *F*0 peak in a high register being the preferred question type, and a penultimate *F*0 peak in a low register being the least preferred question type. However, there appears to be no clear cut distinction between questions and statements. Even the least preferred intonational pattern for questions (low register and *F*0 peak on PU) shows a considerable amount of question ratings (14%).

Reaction time: Fig. 10 depicts the RT results according to the factors PITCH REGISTER, PEAK PLACEMENT and PEAK ALIGNMENT. It becomes apparent, that the patterns in rating are reflected in the observed response latencies. There was a significant interaction of RATING and PITCH REGISTER (SE=102 ms, $\chi^2(1)=13.5$, $p<0.0005$), such that there was a greater response latency advantage for questions when stimuli were in a high register (305 ms) as opposed to low register (47 ms). There was a significant interaction of RATING and PEAK PLACEMENT (SE=105 ms, $\chi^2(1)=19.7$, $p<0.0001$) such that items with the *F*0 peak on the final syllable were responded to faster when rated as a question (307 ms), while items with the *F*0 peak on the penultimate syllable were responded to faster when rated as a statement (123 ms). Thus the RT analyses reflect the general finding that both a higher pitch register and a *F*0 peak on the final syllable are preferred properties of questions. Confirming the absence of an effect of PEAK ALIGNMENT on the rating results, we find no significant interaction of RATING and PEAK ALIGNMENT for RTs ($\chi^2(1)=2.86$, $p=0.09$). Descriptively there was a greater response latency advantage for questions when stimuli had late *F*0 peaks (191 ms) as opposed to early *F*0 peaks (91 ms) (cf. Fig. 10).

To sum up, we were able to identify two main factors influencing the perception of the distinction between statements and questions. First, contours in a high pitch register are perceived more frequently and faster as questions than contours in a low register

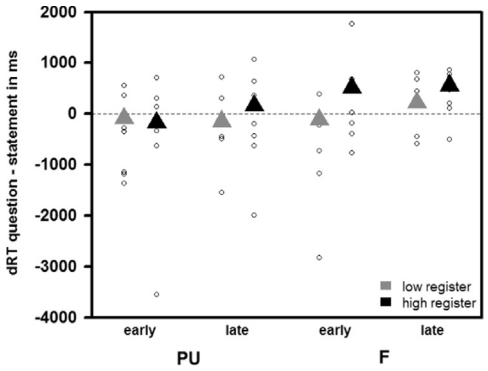

Fig. 10. Differences in response latency between statement and question as a function of pitch register (low vs. high), peak placement (PU vs. F) and peak alignment (early vs. late). Filled triangles give the grand mean, solid circles give means for individual participants. The dashed line indicates no RT difference between statement and question rating (dRT=0 ms). Positive values indicate faster responses for questions; negative values indicate faster responses for statements.

Please cite this article as: Roettger, T. B., & Grice, M. The role of high pitch in Tashlhiyt Tamazight (Berber): Evidence from production and perception. *Journal of Phonetics* (2015), http://dx.doi.org/10.1016/j.wocn.2014.12.004

confirming H2a. This matches the strong pitch register differences found in production (H1a). Second, contours with *F*0 peaks on the final syllable are perceived more frequently and faster as questions than *F*0 peaks on the penultimate syllable confirming H2b. However, even contours with a *F*0 peak on the penultimate syllable appear to be acceptable for questions reflecting the variation of *F*0 peak placement in our production experiment. Third, more gradient differences in terms of tonal alignment did not appear to have much of an effect on ratings or reaction times proving H2c false. Any observable trends failed to reach significance. The *F*0 peak alignment results could be due to the nature of the experimental design. Since both pitch register and tonal association are perceptually very prominent and sufficient to solve the task, listeners might not pay attention to subtle alignment differences.

4. General discussion

The present study has demonstrated the functional relevance of high pitch for questions and statements in Tashlhiyt. Production data revealed that compared to statements (a) questions had a higher pitch level and a greater pitch span, and (b) were more often realized with the *F*0 peak on the final syllable. In statements, the *F*0 peak occurred more often on the penult. Furthermore, there was a tendency for (c) the *F*0 peak in questions to be realised later in the syllable than in statements.

The pitch register differences were consistent within and across speakers and appeared to be a robust cue for disambiguating questions from statements, both in production and perception. In terms of *F*0 peak placement, although there was a significantly different distribution across the two sentence modalities in production, there was considerable variation, both within and across speakers. Recall that 90% of *F*0 peaks were on the final syllable in questions, leaving the remaining 10% to occur on the penult. More strikingly, only 70% of *F*0 peaks in statements were on the penult, i.e., 30% of them were on the final syllable. Despite this variability in production, the perception results showed that listeners did indeed use *F*0 peak placement to guide their perception of sentence modality, although both *F*0 peak positions (on penult or final syllable) were acceptable tonal patterns for either questions or statements. This variability in *F*0 peak placement can be seen as an alternation that is to some degree free, i.e., it does not necessarily have an impact on the pragmatic interpretation of the utterance. As discussed in Section 1, however, this alternation is constrained by different factors in a probabilistic way. Moreover, *F*0 peak alignment differences within the syllable in (c) proved to be only marginal cues in perception, even if they were fairly robust in production.

We have shown that echo questions differ from contrastive statements along the two following dimensions:

Global pitch register: The language exploits high pitch to mark interrogativity by means of a raised pitch level and wider pitch span, as in many other languages such as Mandarin Chinese, Hausa and Moroccan Arabic. One way to account for such register differences phonologically is to incorporate tonal features that capture such syntagmatic differences (Inkelas & Leben, 1990; Snider, 1999). Building on autosegmental theory, these approaches assume a register tier (in addition to a tone tier, such that H and L tones can, for instance, be in an upper or lower register). Alternatively, register can be seen as a syntagmatic property of intonation phrases (cf. Ladd, 1990).

Local target: The language also exploits high pitch in terms of a rise to a local high pitch target represented as a H tone, similarly to many other languages. Differences in the location of this tone have been analysed in different ways for different languages. We review here how the patterns in Tashlhiyt compare with the reported location of H tones in these other languages, and explore the options for a phonological representation. In autosegmental phonology there are different representations for the rise fall, depending on the regularities of the tonal contour across the different texts it co-occurs with. In varieties of Italian that have a rise fall question contour (cf. Grice et al., 2005; Savino & Grice, 2011), the H is taken to be part of the focal pitch accent, and thus associated with the head of the intonation phrase. In autosegmental terms, this kind of association relies on a metrical structure that specifies the strong nodes down which the tone percolates. The metrical structure is usually determined by lexically specified prominence relations, with the H tone being associated with a lexically stressed syllable.

In other languages, such as Standard Hungarian (Ladd, 1983; Varga, 2002) and Cypriot Greek (Arvaniti, 1998), the H is an edge tone, and is taken to be associated with the edge of the intonation phrase. Its placement can be represented in two ways, either it has a secondary association to the penult or final syllable (Grice et al., 2000) or it is specified as simply aligning with prosodic constituents without an association to a tone bearing unit (e.g., with the right edge of the phrase, following Pierrehumbert & Beckman, 1988, but also with the left edge of the following L tone, following Gussenhoven, 2000). In a third type of language, such as Standard Greek and Romanian, the H tone is placed further away from the edge, typically on a post-focal lexically strong syllable (Grice et al., 2000). Here the H is neither on the strongest syllable in the phrase, nor close to the edge of the phrase. Such tones are referred to as phrase accents. They mark the head of a post-focal constituent but not of the entire phrase.

How, then, can H tones in Tashlhiyt be accounted for? We could tentatively propose the following analysis for echo questions: The rise to a H tone always appears close to the right edge of the intonational phrase (cf. Figs. 1–3), analysed as a complex right edge boundary tone structure. Since this tonal complex can be located on either the penult or the final syllable, we shall analyse it as an IP boundary tone with a secondary association to the penult or final syllable, similar to Standard Hungarian and Cypriot Greek (Grice et al., 2000). Which of these two syllables it is associated with is, unlike in other languages, not dependent on lexically determined prominent heads (word stress). Previous work has argued that metrical structure in the lexicon is illusive, although the placement of a H tone on a syllable appears to be accompanied by concomitants of stress, such as duration and intensity (Gordon & Nafi, 2012). As Grice et al. (accepted for publication) have argued, the variation in *F*0 peak placement within words leads us to interpret metrical strength as something that is determined at a postlexical rather than a lexical level. Even though this postlexical prominence appears to be assigned to some degree freely, it is determined probabilistically by certain factors such as sonority, syllable weight and sentence modality (Grice et al., accepted for publication).

Similarly, in contrastive statements the rise to a H tone always appears close to the right edge of the contrasted constituent (cf. Figs. 1 and 7). This might at first sight be represented as a complex right edge boundary tone structure, too. However, closer

Please cite this article as: Roettger, T. B., & Grice, M. The role of high pitch in Tashlhiyt Tamazight (Berber): Evidence from production and perception. *Journal of Phonetics* (2015), http://dx.doi.org/10.1016/j.wocn.2014.12.004

inspection reveals that, unlike questions, this tonal complex is associated to a prosodic constituent lower in the prosodic hierarchy, in this case a phrase defined by the focal structure of the utterance. Again, this boundary tone is secondarily associated with either the penult or final syllable of that constituent with a preference for the penult.

Alternatively, one could consider the H tone in statements as something like a focal pitch accent associated to the head of the phrase. However, evidence against this view comes from data presented in previous work (Grice et al., 2011, accepted for publication; Roettger et al., 2013). These studies present cases in which the contrasted word contains no sonorous nuclei. In those cases, the *F*0 peak was frequently realized on the preceding word, thus outside the focused constituent. This finding can be captured, however, by an edge tone analysis: In absence of a sonorous tone bearing unit at the right edge of the phrase, the H tone aligns with the rightmost available sonorous element in the preceding word.

So H tones in both questions and contrastive statements can be accounted for as secondarily associated tones. These tones are at once boundary tones, being close to the phrase boundary, and prominence lending tones, resembling pitch accents. This analysis accounts for both Gordon and Nafi's findings as to the phonetic prominence of syllables bearing these tones and our own observations that these tones tend to occur near the right edge of phrases.

The above analysis would account for the general occurrence of H tones in terms of their position within the utterance. However, it fails to explain the apparent alternation between penultimate and final H tones. While there have been reports on free alternation of word prominence for Indonesian (e.g., Goedemans & van Zanten, 2007), free alternation in higher prosodic domains is so far unattested. It appears that Tashlhiyt does have some degree of free alternation in *F*0 peak placement, making it difficult to provide a phonological account in discrete terms. Our results illustrate rather the need to account for more gradient probabilistic aspects of tonal placement in the grammar. Ladd (2014) explicitly distinguished between *physical gradience*, for instance the gradient phonetic variability in tonal scaling, and *statistical gradience*, the variance in frequency of occurrence of categorically definable events. However, as Ladd acknowledges, these two types of gradience are often difficult to keep apart. In case of tonal placement in Tashlhiyt, there appears to be statistical gradience in terms of what syllable the H tone is associated with. Categorically definable events, here an H tone on the penultimate or final syllable, reveals statistical variation in location associated probabilistically with categorically different meanings. This statistical gradience is further reflected in physical gradience, i.e., gradient variability in the timing of the tone within the syllable. While physical gradience is well acknowledged in intonation research, as it is naturally found in the measurement of alignment and scaling, statistical gradience has been widely neglected. However, as more studies on different languages are carried out, it becomes apparent that statistical gradience is an inherent property of many intonational systems.

Acknowledgements

We would like to thank Rachid Ridouane for his generous support at both the conceptual stage and during the actual data collection. We also thank all members of the faculty of Amazigh studies at the Ibn Zohr University in Agadir. We are grateful for all subjects that patiently participated in these experiments. This research was funded by the Volkswagen Stiftung (Project: Tonal Placement – the Interaction of Qualitative and Quantitative Factors: ToPIQQ).

Appendix A

Mean proportion (in %) of peak location on the final syllable (F) and penultimate syllable (PU) (target word /managu/ 'when' was excluded) for each speaker separately pooled over target words; number of tokens is given in brackets. Note that echo questions produced by speaker F5 were frequently excluded from the data analysis due to hesitations and a monotonous speaking style which made it impossible to identify a high target.

Speaker	Statement		Question	
	PU	*F*	PU	*F*
M1	25 (2)	75 (6)	12.5 (1)	87.5 (7)
M2	66.7 (6)	33.3 (3)	0 (0)	100 (9)
M3	100 (8)	0 (0)	22.2 (2)	77.8 (7)
M4	28.6 (2)	71.4(5)	0 (0)	100 (9)
M5	66.7 (6)	33.3 (3)	0 (0)	100 (9)
F1	88.9 (8)	11.1 (1)	22.2 (2)	77.8 (7)
F2	66.7 (6)	33.3 (3)	0 (0)	100 (6)
F3	83.3 (5)	16.7 (1)	0 (0)	100 (9)
F4	77.8 (7)	22.2 (2)	22.2 (2)	77.8 (7)
F5	88.9 (8)	11.1 (1)	100 (1)	0 (0)

Please cite this article as: Roettger, T. B., & Grice, M. The role of high pitch in Tashlhiyt Tamazight (Berber): Evidence from production and perception. *Journal of Phonetics* (2015), http://dx.doi.org/10.1016/j.wocn.2014.12.004

Appendix B

Mean proportion (in %) of peak location on the final syllable (F) and penultimate syllable (PU) for each target word separately pooled over speakers; number of tokens is given in brackets. Words are represented in Latin orthography as they were presented in the experiments. Note that ⟨baba⟩ and ⟨udm⟩ are the only target words displaying variance of the *F*0 peak location in questions.

Target word	Statement		Question	
	PU	*F*	PU	*F*
baba	77.8 (7)	22.2 (2)	11.1 (1)	88.9 (8)
dari	90 (9)	10 (1)	0 (0)	100 (9)
dima	88.9 (8)	11.1 (1)	0 (0)	100 (8)
ildi	50 (10)	50 (10)	0 (0)	100 (17)
iminnun	28.6 (2)	71.4 (5)	0 (0)	100 (9)
managu	14.3 (1)	85.7 (7)	0 (0)	100 (8)
Rila	85.7 (7)	14.3 (1)	0 (0)	100 (9)
udm	75 (15)	25 (5)	41.1 (7)	58.8 (10)

References

Abboud, H. (1991). *SuperLab*. Wheaton, MD: Cedrus.
Applegate, J. R. (1958). *An outline of the structure of Shilha*. New York: American Council of Learned Societies.
Arvaniti, A. (1998). Phrase accents revisited: Comparative evidence from Standard and Cypriot Greek. In *Proceedings of the 5th international conference on spoken language processing* (Vol. 7, pp. 2883–2886).
Arvaniti, A. (2001). The intonation of wh-questions in Greek. *Studies in Greek Linguistics*, *21*, 57–68.
Arvaniti, A., Ladd, D. R., & Mennen, I. (1998). Stability of tonal alignment: The case of Greek prenuclear accents. *Journal of Phonetics*, *26*, 3–25.
Arvaniti, A., & Ladd, D. R. (2009). Greek wh-questions and the phonology of intonation. *Phonology*, *26*, 43–74.
Barr, D. J., Levy, R., Scheepers, C., & Tily, H. (2013). Random-effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278.
Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*, *R package version 1.0-6*. ⟨http://CRAN.R-project.org/package=lme4⟩.
Benkirane, T. (1998). Intonation in Western Arabic (Morocco). In D. Hirst, & A. Di Cristo (Eds.), *Intonation systems: A survey of twenty languages* (pp. 345–359). Cambridge, UK: Cambridge University Press.
Boersma, P., & Weenink, D. (2013). Praat: *Doing phonetics by computer* [*Computer program*]. ⟨http://www.praat.org/⟩.
Bolinger, D. (1978). Intonation across languages. In J. Greenberg (Ed.), *Universals of human language, Vol. 2: Phonology* (pp. 471–524). Standford: Stanford University Press.
Bolinger, D. (1972). Accent is predictable (if you are a mind reader). *Language*, *48*, 633–644.
Brunelle, M., Ha, K. P., & Grice, M. (2012). Intonation in Northern Vietnamese. *The Linguistic Review*, *29*(1), 3–36.
Clements, G. N. (1997). Berber syllabification: Derivations or constraints?. In Iggy Roca (Ed.), *Derivations and constraints in phonology* (pp. 289–330). Oxford: Clarendon.
Cruttenden, A. (1997). *Intonation* (2nd ed.). New York: Cambridge University Press.
D'Imperio, M., & House, D. (1997). Perception of questions and statements in Neapolitan Italian. *Proc. Eurospeech*, *97*, 251–254.
Dell, F., & Elmedlaoui, M. (1985). Syllabic consonants and syllabification in Imdlawn Tashlhiyt Berber. *Journal of African Languages and Linguistics*, *7*, 105–130.
Dell, F., & Elmedlaoui, M. (1988). Syllabic consonants in Berber: Some new evidence. *Journal of African Languages and Linguistics*, *10*, 1–17.
Dell, F., & Elmedlaoui, M. (2002). *Syllables in Tashlhiyt Berber and in Moroccan Arabic*. Dordrecht: Kluwer.
Goedemans, R. W. N., & van Zanten, E. A. (2007). Stress and accent in Indonesian. In V. J. van Heuven, & E. A. van Zanten (Eds.), *Prosody in Indonesian languages* (pp. 35–62). LOT: Occasional Series (9).
Gordon, M., & Nafi, L. (2012). The acoustic correlates of stress and pitch accent in Tashlhiyt Berber. *Journal of Phonetics*, *40*, 706–724.
Gosy, M., & Terken, J. (1994). Question marking in Hungarian: Timing and height of pitch peaks. *Journal of Phonetics*, *22*, 269–281.
Grice, M. (1995). *The intonation of interrogation in Palermo Italian; implications for intonation theory*. Tübingen: Niemeyer.
Grice, M., Benzemüller, R., Savino, M., & Andreeva, B. (1995). The intonation of queries and checks across languages: Data from Map Task dialogues. In *Proceedings of the 13th international congress of phonetic sciences* (pp. 648–651).
Grice, M., D'Imperio, M., Savino, M., & Avesani, C. (2005). Towards a strategy for ToBI labelling varieties of Italian. In S. A. Jun (Ed.), *Prosodic typology and transcription: A unified approach* (pp. 55–83). Oxford: Cambridge University Press.
Grice, M., Ladd, D. R., & Arvaniti, A. (2000). On the place of phrase accents in intonational phonology. *Phonology*, *17*, 143–185.
Grice, M., Roettger, T.B., Ridouane, R., Tonal association in Tashlhiyt Berber: Evidence from polar questions and contrastive statements. *Phonology* (accepted for publication).
Grice, M., Roettger, T. B., Ridouane, R., & Fougeron, C. (2011). Tonal association in Tashlhiyt Berber. In *Proceedings of the 17th international congress of phonetic sciences*. Hong Kong (pp. 775–778).
Gussenhoven, C. (2000). The boundary tones are coming: On the nonperipheral realization of boundary tones. In Broe, & Pierrehumbert (Eds.), *Papers in laboratory phonology V: Acquisition and the lexicon* (pp. 132–151). Cambridge: Cambridge University Press.
Gussenhoven, C. (2002). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.
Haan, J. (2002). *Speaking of questions. An exploration of Dutch question intonation*. LOT dissertation.
Hadding-Koch, K., & Studdert-Kennedy, M. (1964). An experimental study of some intonation contours. *Phonetica*, *11*, 175–185.
Hayes, B., & Lahiri, A. (1991). Bengali intonational phonology. *Natural Language and Linguistic Theory*, *9*, 47–96.
Hirst, D., & Di Cristo, A. (1998). *Intonation systems: A survey of twenty languages*. Cambridge: Cambridge University Press.
House, D. (2003). Perceiving question intonation: The role of pre-focal pause and delayed focal peak. In *Proceedings of the 15th international congress of phonetic sciences*. Barcelona (pp. 755–758).
Hyman, Larry M. (1985). *A theory of phonological weight*. Dordrecht: Foris.
Inkelas, S., & Leben, W. R. (1990). Where phonology and phonetics intersect: The Case of Hausa intonation. In M. Beckman, & J. Kingston (Eds.), *Papers in laboratory phonology: Between the grammar and physics of speech* (pp. 17–34). Cambridge: Cambridge University Press.
Katsika, A., Krivokapić, J., Mooshammer, C., Tiede, M., & Goldstein, L. (2014). The coordination of boundary tones and their interaction with prominence. *Journal of Phonetics*, *44*, 62–82.
Ladd, D. R. (1983). Phonological features of intonational peaks. *Language*, *59*, 721–759.
Ladd, D. R. (1990). Metrical representation of pitch register. In M. Beckman, & J. Kingston (Eds.), *Papers in laboratory phonology: Between the grammar and physics of speech* (pp. 35–57). Cambridge: Cambridge University Press.
Ladd, D. R. (2008). *Intonational phonology* (2nd edition). Cambridge: Cambridge University Press.
Ladd, D. R. (2014). *Simultaneous structure in phonology*. Oxford: Oxford University Press.
Laver, J. (1994). *Principles of phonetics*. Cambrige: Cambridge University Press.
Makarova, V. (2007). The effect of Pitch Peak alignment on sentence type identification in Russian. *Language and Speech*, *50*, 385–422.
Ohala, J. J. (1983). Cross-Language use of pitch: An ethological perspective. *Phonetica*, *40*, 1–18.

Please cite this article as: Roettger, T. B., & Grice, M. The role of high pitch in Tashlhiyt Tamazight (Berber): Evidence from production and perception. *Journal of Phonetics* (2015), http://dx.doi.org/10.1016/j.wocn.2014.12.004

Ohala, J. J. (1984). An ethological perspective on common cross-language utilization of *F*0 of voice. *Phonetica*, *41*, 1–16.
Pierrehumbert, J., & Beckman, M. (1988). *Japanese tone structure*. Cambridge, MA: MIT Press.
Prieto, P., van Santen, J., & Hirschberg, J. (1995). Tonal alignment patterns in Spanish. *Journal of Phonetics*, *4*, 429–451.
R Core Team. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. ⟨http//www.R-project.org⟩.
Ridouane, R. (2008). Syllables without vowels: Phonetic and phonological evidence from Tashlhiyt Berber. *Phonology*, *25*, 321–359.
Roettger, T. B., Ridouane, R., & Grice, M. (2013). Phonetic alignment and phonological association in Tashlhiyt Berber. In *Proceedings of 21st ICA*. Montreal.
Savino, M. (2012). The intonation of polar questions in Italian: Where is the rise?. *Journal of the International Phonetic Association*, *42*, 23–48.
Savino, M., & Grice, M. (2007). The role of pitch range in realizing pragmatic contrasts. In *Proceedings of the 16th international congress of phonetic sciences*. Saarbrücken, Germany (pp. 1037–1040).
Savino, M., & Grice, M. (2011). The perception of negative bias in Bari Italian questions. In S. Frota, E. Gorka, & P. Prieto (Eds.), *Prosodic categories: Production perception and comprehension, Studies in natural language and linguistic theory* (Vol. 82, pp. 187–206).
Shen, X.-N. S. (1990). *The prosody of Mandarin Chinese*. Berkeley: University of California Press.
Silverman, K., & Pierrehumbert, J. (1990). The timing of prenuclear high accents in English. In J. Kingston, & M. Beckman (Eds.), *Papers in laboratory phonology I* (pp. 72–106). Cambridge, UK: Cambridge University Press.
Snider, K. (1999). *The geometry and features of tone*. Dallas: Summer Institute of Linguistics.
Ultan, R. (1978). Some general characteristics of interrogative systems. In J. Greenberg (Ed.), *Universals of human language. Vol. 4: Syntax* (pp. 211–248). Standford: Stanford University Press.
Van Heuven, V. J. (1994). What is the smallest prosodic domain?. In P. A. Keating (Ed.), *Phonological structure and phonetic form: Papers in laboratory phonology III* (pp. 76–98). Cambridge: Cambridge University Press.
Varga, L. (2002). *Intonation and stress: Evidence from Hungarian*. New York: Palgrave Macmillan.
Zec, D. (1995). Sonority constraints on syllable structure. *Phonology*, *12*, 85–129.

Please cite this article as: Roettger, T. B., & Grice, M. The role of high pitch in Tashlhiyt Tamazight (Berber): Evidence from production and perception. *Journal of Phonetics* (2015), http://dx.doi.org/10.1016/j.wocn.2014.12.004