Perception of Peak Placement in Tashlhiyt Berber

Timo B. Roettger[1], Rachid Ridouane[2] & Martine Grice[1]

[1]IfL Phonetik, University of Cologne;
[2]Laboratoire de Phonétique et Phonologie (UMR 7018) CNRS/Sorbonne Nouvelle

timo.roettger@uni-koeln.de; martine.grice@uni-koeln.de; rachid.ridouane@univ-paris3.fr

Abstract

Previous production studies on Tashlhiyt Berber have demonstrated that questions and statements have similar intonation contours, i.e., a final rise to a F0 peak and subsequent fall. The contours tended to differ in overall pitch register and peak location: questions (a) revealed a stronger tendency to be realized with the F0 peak on the final syllable than statements and (b) even within the same syllable, peaks were often aligned later in questions than in statements. The peak location, however, was reported to vary strongly both within and across speakers, interpreted as free alternation of tonal association. Given this high degree of variation, the question arises as to how relevant this variation is for communication. The present perception study shows that both pitch register (low vs. high) and tonal placement (peak on penultimate vs. final syllable) affect listeners' judgments on sentence modality as well as reaction times. Whereas peak alignment within the syllable (early vs. late) did not affect judgments, it did have a marginal effect on reaction times. By demonstrating their perceptual impact, this study confirms that the patterns found in production are communicatively relevant.

Index Terms: Tashlhiyt Berber, intonation, tonal association, tonal alignment, pitch register

1.　　Introduction

Berber is an Afro-Asiatic language spoken across large parts of North Africa. Tashlhiyt Berber, the variety investigated here, is one of three major Berber varieties in Morocco.

Recent studies on Tashlhiyt have shown that the intonation of this language involves a high degree of variation both within and across speakers [1,2,3,4]. In particular, it was shown that in polar questions and contrastive statements pitch peaks consistently co-occur with available sonorant nuclei [3,4]. The authors accounted for this placement in terms of a phonological association. More specifically, the tone is analysed as an edge tone, H, of a prosodically defined phrase with a secondary association to a tone bearing unit, a syllable with a sonorant nucleus. However, there was considerable variation with regard to which sonorant nucleus the peak co-occurs with. In fact, speakers showed a certain degree of free alternation: Consider the examples in Figure 1 which demonstrate this apparent free alternation. The same speaker produced the sentence /inːa tugl/ (*'He said 'she held''*), embedded in the same context, with an intonational peak on the final syllable (Top) or on the penultimate syllable (Bottom). Nonetheless, this alternation was not found to be entirely free. Even though there was a strong trend to realize the H on the final syllable in both types of utterances, there were more occurrences of H on the final syllable in questions than in statements. Impressionistically the peak placement goes hand in hand with some degree of segmental prominence, i.e., the syllable the peak co-occurs with is longer and louder than it would be without a peak (compare /u/ and /l/ in /tugl/ of Figure 1 with and without an intonational peak)
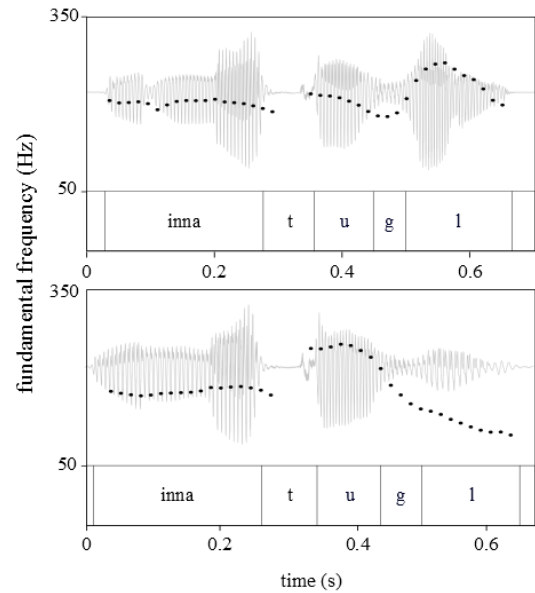

Figure 1: *Waveform and F0 contours of statements /inːa tugl/ (3ms-say 3fs-held) 'He said 'she held'' produced by the same speaker in the same context. Top: Peak placed on final syllable; Bottom: peak placed on penultimate syllable.*

Moreover, this discrete asymmetry, interpreted as phonological association of the H tone, was reflected in more gradual phonetic detail, i.e., the phonetic alignment within the syllable [cf. 4]. Peaks in questions were aligned later in the syllable than in statements. This tendency appeared to be highly speaker dependent. In addition to these differences in peak placement, questions differed from statements in terms of pitch register. Questions were produced with a higher baseline and a concomitant steeper rise to the pitch peak than statements.

The discussed different preferences in questions and statements have to be considered as functionally motivated to distinguish sentence modality. The remaining variability, however, might reflect the redundant nature of peak location in Tashlhiyt. Polar questions are marked by an initial preverb /is/. This morphosyntactic marker could already be sufficient for indicating sentence modality. Tonal cues might thus be redundant to some degree. However, there is also an echo question in this language (inːa tugl? *'He said 'she held'?'*). This type of question is morphosyntactically identical to statements, and differs from statements by intonation only. The present paper reports on a perception study which investigates the contribution of tonal cues to the discriminability of the contrast between statements and echo questions.

On an alternative account, the observed variability could be due to the sample tested in previous studies. The production data of [3,4] was based on Tashlhiyt speakers that have lived in

Paris for a significant amount of time, which might have had an influence on their productions. Thus, the present study attempts to validate the obtained patterns in production (based on speakers living in France) by testing these contours perceptually with Tashlhiyt speakers living in Morocco.

2.   The present study

To summarize, polar questions and statements in Tashlhiyt have been mainly shown to differ according to following dimensions:

(a) Questions are more likely to have the peak on the final syllable than statements.

(b) Peaks in questions are aligned later within the syllable than peaks in statements.

(c) Questions reveal a greater pitch register than statements.

The present study investigates the impact of those three factors on the perception of sentence modality.

2.1   Methodology

2.1.1   Speech material

We used stimuli with a resynthesized F0 contour in order to control for pitch register and peak placement. As base stimuli we used four short phrases /inːa baba/, /inːa bibi/, /inːa dima/, and /inːa ʁila/ '3ms-say ('father, turkey, always, now')' produced by a trained native speaker of Tashlhiyt. For each phrase, the speaker produced two contours corresponding to two different tonal associations of the high tone as displayed in Figure 2, resulting in two sets of stimuli. One set (PU) contained a rise to an F0 peak on the prefinal syllable (Figure 2 bottom) and the other set (F) contained the rise and peak on the final syllable (Figure 2 top). As mentioned above, tonal association was accompanied by segmental prominence, i.e., the syllable the peak was co-occurring with was longer and louder (cf. Figure 1-2).

Both sets were resynthesized using PSOLA in Praat [5]. F0 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, F0 was manipulated to start rising at the offset of /inːa/ towards two different F0 maximum locations for each set: In the *early* peak condition F0 reached its maximum at 1/3 of the vowel (penult in set PU and final in set F), in the late peak condition F0 reached its maximum at 2/3 of the vowel. The maximum F0 value was 4 semitones higher than the baseline (~164 Hz and ~206 Hz respectively). After reaching its maximum, F0 fell towards the baseline located at the end of the target word. These manipulations resulted in 32 stimuli (4 target words * 2 tonal associations (penultima vs. final) * 2 peak alignments (early vs. late) * 2 pitch registers (low vs. high)) (cf. Figure 3).
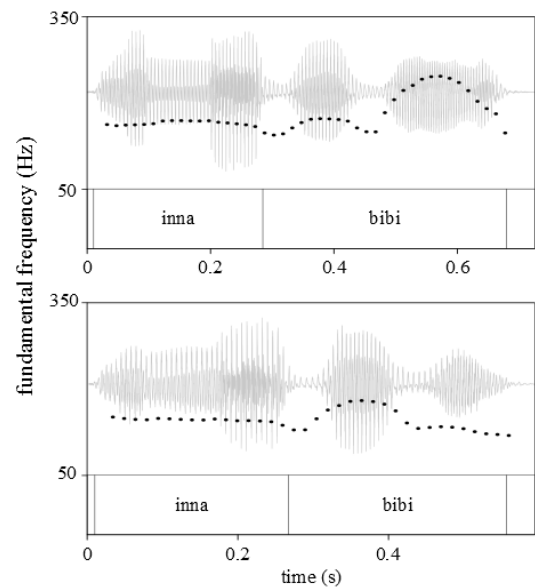

Figure 2: *Waveform and F0 contours of base stimulus productions /inːa bibi/ (3ms-say turkey) 'He said 'turkey'' as an echo question with a peak on the final syllable (Top) and as a statement with a peak on the penultimate syllable (Bottom).*
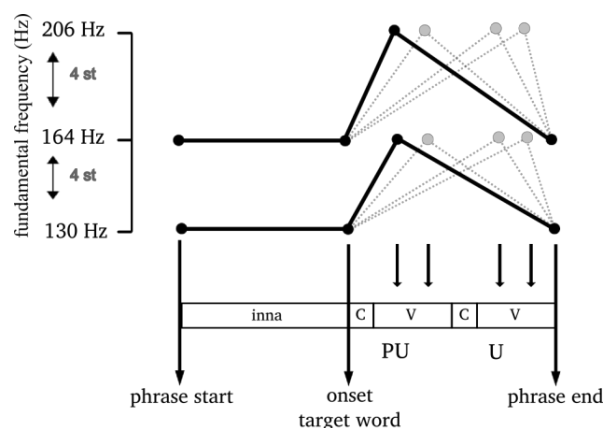

Figure 3: *Schematized manipulation condition displaying the differences of pitch register, tonal association and peak alignment.*

2.1.2   Participants and procedure

Eight native speakers of Tashlhiyt Berber (average age = 21; 4 men; 4 women) participated in our experiment. All live in Agadir, Morocco, and speak at least Moroccan Arabic and to a lesser extent French. Participants were seated in front of a computer screen in a quiet room at the lbn Zohr University in Agadir. Participants 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.

The experiment was controlled using Superlab [6]. 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, the stimulus was presented auditorily. Simultaneously 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 Latin script in blue (e.g. inna baba !), on the other side the question was displayed in red (e.g. inna baba ?). 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.

2.1.3   Analyses

All data were analyzed with generalized linear mixed models, using R [7] and the package *lme4* [8]. To analyze responses categorically, we used mixed logistic regression models with "RATING" (question or statement) as the dependent measure. As fixed effects we included TONAL ASSOCIATION (PU vs. F), PEAK ALIGNMENT (early vs. late), PITCH REGISTER (low vs. high), in addition to TARGET WORD and mean-centered REPETITION. To analyze reaction times (RTs, measured from the offset of the audio stimulus), we used models with Gaussian error distribution with RTs as dependent variable. As fixed effects we included the two-way interactions of RATING and TONAL ASSOCIATION, RATING and PEAK ALIGNMENT, and RATING and PITCH REGISTER, in addition to TARGET WORD and mean-centered REPETITION. For both analyses, we included a term for random intercepts for speakers as well as correlated random slopes for the fixed effects TONAL ASSOCIATION, PEAK ALIGNMENT and PITCH REGISTER (and RATING). For both dependent variables, we tested whether the inclusion of any of the fixed effects did improve the models prediction significantly via likelihood ratio tests.

3.   Results and Discussion

Rating: Figure 4 depicts the rating results according to the factors TONAL ASSOCIATION, PEAK ALIGNMENT and PITCH REGISTER. Overall, participants rated the stimuli to correspond to questions in 43% of the cases. There was a significant effect of TONAL ASSOCIATION ($\chi^2(1)$=8.66, p=0.003) such that items with the peak on the final syllable were significantly more often rated as a question than statements (61% vs. 25%). There was a significant effect of PITCH REGISTER ($\chi^2(1)$=9.14, p=0.003), as well, such that items with a high pitch register were significantly more often rated as a question (58% vs. 28%). PEAK ALIGNMENT did not have a significant effect on responses. Early peaks were rated to correspond to questions comparably as often as late peaks (44% vs. 41%) ($\chi^2(1)$= 1.18, p=0.28). As can be seen in Figure 4, there was no apparent interaction of TONAL ASSOCIATION and PITCH REGISTER. Those effects rather add up, with a final peak in a high register being the most preferred question type, and a prefinal peak in a low register being the least preferred question type. However, there appears to be no clear cut. Even the least preferred intonational pattern for questions (low register and peak on PU) shows a considerable amount of question ratings (14%).
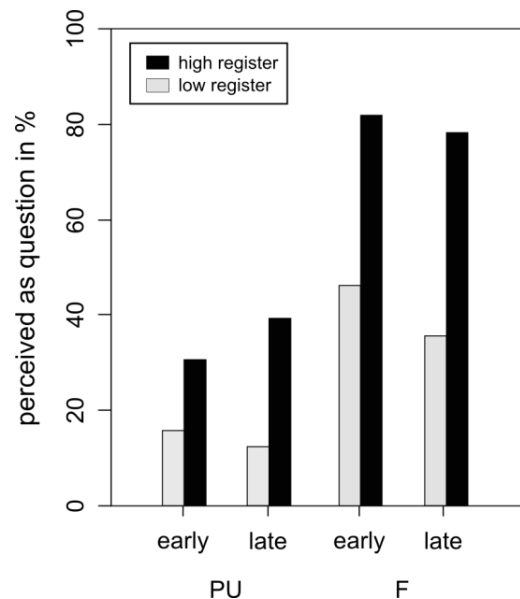

Figure 4: *Ratings as a function of tonal association (PU vs. F), peak alignment (early vs. late) and pitch register (low vs. high).*
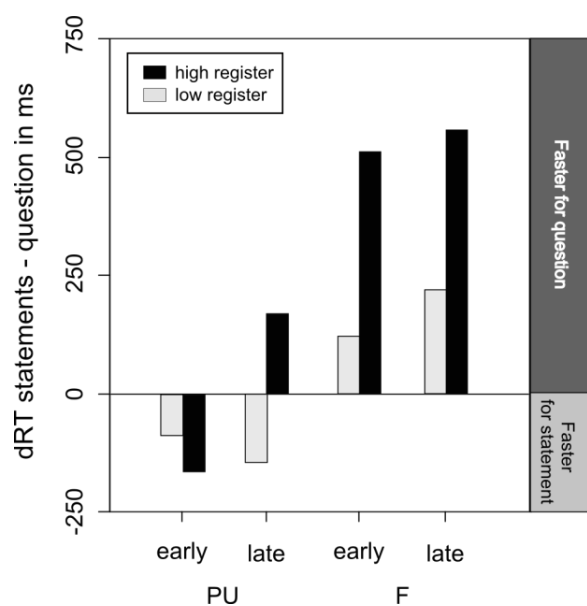

Figure 5: *Differences in response latency between statement and question as a function of tonal association (PU vs. F), peak alignment (early vs. late) and pitch register (low vs. high). Positive values indicate faster responses for questions; negative values indicate faster responses for statements. dRT = 0 indicates no RT difference between statement and question rating.*

RTs: Figure 5 depicts the RT results according to the factors TONAL ASSOCIATION, PEAK ALIGNMENT and PITCH REGISTER. There was a significant interaction of RATING and TONAL ASSOCIATION ($\chi^2(1)$=13.49, p=0.002) such that items with the peak on the final syllable were responded to faster when rated as a question (307 ms), while items with the peak on the

penultimate syllable were responded to faster when rated as a statement (123 ms). There was a significant interaction of rating and pitch register ($\chi^2(1)$=19.72, p<0.0001), as well, 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). Thus the RT analyses reflect the general finding that both a higher pitch register and a peak on the final syllable are preferred properties of questions. Despite the absence of an effect of PEAK ALIGNMENT on RATING, we find a marginally significant interaction of RATING and PEAK ALIGNMENT for RTs ($\chi^2(1)$=2.86, p=0.09), such that there was a greater response latency advantage for questions when stimuli had late peaks (191 ms) as opposed to early peaks (91 ms) (cf. Figure 5).

To sum up, we were able to identify two main factors influencing the perception of the distinction between statements and questions. First, contours with H tones associated to the final syllable are perceived more frequently and faster as questions than H tones associated to the penultimate syllable. However, even contours with a peak on the penultimate syllable appear to be acceptable for questions. Second, contours in a high pitch register are perceived more frequently and faster as questions than contours in a low register. The evidence for peak alignment within each syllable is somewhat weaker. While there was no significant rating asymmetry between early and late peaks, response latencies indicate that late peaks are processed faster when rated as a question (marginally significant). These findings are comparable to results from production, in which [3,4] found some degree of free alternation of tonal association of H tones in polar questions and statements, although questions had a stronger tendency to be realized with a high tone on the final syllable than statements. They further report on later peaks in questions in terms of where the peak is located within the syllable it co-occurs with. Finally, based on impressionistic observations, they report on higher pitch register in questions.

4.   General Discussion

The most important contribution of the present study is that Tashlhiyt speakers not only exhibit free alternation of H tone association in production, but are also tolerant with regard to peak placement in perception. For example in questions, there is a probabilistic trend for the peak to be placed on the final syllable, nonetheless, speakers not only produce peaks on the penultimate syllable but also accept those as questions in perception.

We conclude that this variability in locating the tone is free alternation of tonal association. While there have been reports on free alternation of word prominence for Indonesian [9], free alternation in higher prosodic domains are so far unattested. Even though such an alternation of tonal association has not been reported yet, there is some degree of alternation in a number of languages, depending on the properties of syllables at or near the phrase edge. For instance, in Standard Greek, there is secondary association to a lexical stressed syllable if there is one available, otherwise the edge tone is associated with the final syllable [10,11].

We have also observed a preference to place the peak as far to the right as possible. The general preference for late peaks in questions (in terms of both association and alignment), i.e., a rise in pitch, is common across languages [12]. Moreover, several studies showed that speakers use pitch peak alignment to disambiguate sentence types. For example, [13] showed that

peak alignment plays a role in disambiguating Hungarian statements and polar questions, whereby an early pitch peak (in the accented vowel) is associated with declaratives and a late peak with interrogatives (see also [14], for Neapolitan Italian; [15], for Swedish; and [16], for Russian). As mentioned above, pitch peaks in Tashlhiyt questions have been observed to exhibit a concomitant steeper rise than in statements. Thus, in production, peak alignment differences could be an artifact of reaching a higher peak. It is well known that a higher peak can perceptually have the same effect as peak delay [14,15,17]. The present study indicate that Tashlhiyt listeners might be sensitive to peak delay irrespective of peak height (although only for RTs).

The remaining variability found in both production and perception, might reflect the redundant nature of peak location in Tashlhiyt. Global pitch cues such as pitch register could already be sufficient for marking sentence modality. The actual peak location phrase finally is thus one of a number of cues and redundant to some degree.

It is important to stress, that while earlier reports on Tashlhiyt intonation [1,3,4] were based on Tashlhiyt speakers living in Paris. The present study recruited subjects that have lived in Morocco for all their lives. We were thus able to confirm the observed patterns in production making it unlikely that they are merely due to interference. Thus, this study not only reveals information about the perception of tonal cues in Tashlhiyt for the first time, it is also a first validation of obtained patterns in production.

5.   Acknowledgements

6.   References

[1]   Grice, M, Roettger, T. B., Ridouane, R. and Fougeron, C., "Tonal association in Tashlhiyt Berber", Proc. 17th ICPhS, 775–778, 2011.

[2]   Gordon, M. and Nafi, L., "The acoustic correlates of stress and pitch accent in Tashlhiyt Berber", Journal of Phonetics, 40: 706–724, 2012.

[3]   Roettger, T. B., Ridouane, R. and Grice, M., "Phonetic alignment and phonological association in Tashlhiyt Berber", Proc. 21st ICA, 2013.

[4]   Roettger, T. B., Ridouane, R. and Grice, M., "Sonority and syllable weight determine tonal association in Tashlhiyt", Proc. 6th Speech Prosody, 2012.

[5]   Boersma, P., and Weenink, D., "Praat: doing phonetics by computer", [Computer program], http://www.praat.org/, 2013.

[6]   Abboud, H., "SuperLab", Wheaton, MD: Cedrus, 1991

[7]   R Core Team, "R: A language and Environment for Statistical Computing", R foundation for statistical computing, Vienna, URL: http//www.R-project.org.

[8]   Bates, D., Maechler, M. and Bolker, B., "lme4: Linear mixed–effects models using S4 classes", R package version 0.999999–0, 2012.

[9]   Goedemans, R.W.N. and van Zanten, E.A., "Stress and accent in Indonesian", in V.J. van Heuven and E.A. van Zanten [Eds.] Prosody in Indonesian Languages, LOT: Occasional Series 9: 35-62, 2007.

[10] Arvaniti, A., D. R. Ladd and Mennen, I., "Tonal association and tonal alignment: evidence from Greek polar questions and contrastive statements", Language and Speech, 49: 421–450, 2006

[11] Grice, M., Ladd, D. R. and Arvaniti, A., "On the Place of Phrase Accents in Intonational Phonology", Phonology, 17: 143–185, 2000.

[12] Ultan, R., "Some general characteristics of interrogative systems", in J. Greenberg, [Ed.], Universals of human language, Vol. 4: Syntax. Stanford: Stanford University Press, 211–248, 1978.

[13] Gosy, M. and Terken, J., "Question marking in Hungarian: Timing and height of pitch peaks", Journal of Phonetics, 22: 269-281, 1994.

[14] D'Imperio, M., and House, D., "Perception of questions and statements in Neapolitan Italian", In Proc. Eurospeech, 97: 251-254, 1997.

[15] House, D., "Perceiving question intonation: The role of pre-focal pause and delayed focal peak", Proc. of 15th ICPhS, 755–758, 2003.

[16] Makorova, V., "The Effect of Pitch Peak alignment on sentence type identification in Russian", Language and Speech, 50: 385-422, 2007.

[17] Gussenhoven, C., "The phonology of tone and intonation", CUP: Cambridge, 2002.