



ICA 2013 Montreal

Montreal, Canada

2 - 7 June 2013

Speech Communication

Session 4pSCb: Production and Perception I: Beyond the Speech Segment (Poster Session)

4pSCb41. Phonetic alignment and phonological association in Tashlhiyt Berber

Timo B. Röttger*, Rachid Ridouane and Martine Grice

***Corresponding author's address: Ifl Phonetik, University of Cologne, Cologne, D-50931, Germany, Germany, timo.roettger@uni-koeln.de**

Although, Tashlhiyt Berber uses intonation to mark sentence modality, the location of f_0 events is severely constrained by its notorious predominance of consonantal nuclei (e.g. /tsskʃtst/ 'you dried it (fem)'). Here we report on the alignment of f_0 peaks in disyllabic target words in polar questions and contrastive statements in the language. Data from four native speakers revealed that questions tend to have later f_0 peaks than statements. This was reflected in discrete association patterns when more than one tone bearing unit was available: in questions the f_0 peak occurred significantly more often on the final syllable than in statements. Interestingly, if no association distinction was made, there was a difference in alignment of this peak within a tone bearing unit: the peak was aligned significantly later in questions. Thus, discrete phonological association patterns were mirrored by phonetic alignment detail. These data question the traditional dichotomy between phonological association and phonetic alignment.

Published by the Acoustical Society of America through the American Institute of Physics

PHONETIC ALIGNMENT AND PHONOLOGICAL ASSOCIATION IN TASHLHIYT BERBER

Introduction

Traditionally, tonal placement can be specified at different levels of structure and granularity and in both qualitative (phonological) and quantitative (phonetic) terms. The standard dichotomy is as follows: PHONOLOGICAL ASSOCIATION – the phonological linking of a tone with a Tone Bearing Unit (TBU), such as a syllable. PHONETIC ALIGNMENT – the fine phonetic detail in terms of the exact position of a tonal event (such as an f_0 peak or valley) relative to an acoustic or articulatory landmark *within* a tone bearing unit. In this study we investigate tonal placement in Tashlhiyt Berber (henceforth Tashlhiyt), a language that has (i) an abundance of syllables without vowels or even sonorants, (ii) alternations in syllabicity and (iii) no clear word stress (e.g. Dell & Elmedlaoui 2002), presenting a number of challenges for phonological association of tones and phonetic realisation of tonal events.

In Tashlhiyt any consonant, even a stop, can be the syllable nucleus; see (1), where /b/ and /g/ serve as syllable nuclei:

(1) tbdgt = **tb**.dgt ‘you are wet’

Words like this may surface phonetically with one or more schwa-like elements (henceforth @) between the stops. A number of phonetic and phonological studies have shown convincing evidence that these elements are transitional vocoids, and thus a matter of phonetic detail rather than phonological segments in their own right (Dell & Elmedlaoui 2002, Ridouane 2008, Ridouane & Fougeron 2011). Thus they do not constitute syllable nuclei and therefore cannot be TBUs (but see Coleman 2001 for an alternative view).

According to Dell and Elmedlaoui “the main pitch event” in an intonational phrase is generally on the final or penultimate syllable, as long as it has a sonorant nucleus (2002: 14). The location of this pitch event involves a certain amount of free variation Intonation Phrase finally (Dell & Elmedlaoui 1985). For example, when the word *igidr* ‘eagle’ is at the end of a phrase with question intonation, a rising tone can occur on the second vowel /i/ or on the final /r/. In their analysis, the location of this tone can affect the syllabicity of the final sonorant.

Gordon and Nafi (2012) have recently found that Intonation Phrase final syllables in statements have a higher f_0 than other syllables. They interpret these findings as a phrase final pitch accent, implying a primary association of a tone to a TBU. Grice et al. (2011) found a different contour in statements involving implicit contrast on an utterance medial word. Here there was a f_0 peak on the last available sonorant nucleus in the target word. They analysed the peak as an edge tone with an association to the rightmost available TBU (this is referred to as secondary association, cf. Pierrehumbert & Beckman 1988). If the target word contained no sonorants, both Grice et al. (2011) and Gordon and Nafi (2012) observed highly variable peak alignment patterns: the peak was either on a transitional vocoid within the target word (in both studies) or on a sonorant nucleus preceding it (Grice et al. 2011). The former case presents a problem, as the transitional vocoid is not able to function as a TBU if it is not a syllable nucleus

In this paper we present new data on different sentence modalities, investigating not only the placement of f_0 peaks on specific syllables (interpreted as phonological association) but also the phonetic alignment detail within the syllables on which they occur. To this aim we compared polar questions and contrastive statements where the target word is in Intonation Phrase final position.

Methodology

Speakers and Stimuli

Four native speakers of Tashlhiyt Berber (2 male, 2 female) were recorded. Three of the speakers live in Paris, but have spent at least their first 24 years in Morocco, and report using the language frequently. The other speaker (M2) lives in Morocco (in the area of Agadir). The age of the subjects ranged from 31 to 46 (mean = 40.5).

Stimuli contained 56 disyllabic target words and were selected to cover a wide range of segment types. The target words were incorporated into a mini-monologue of the following type:

- (2) a. Is inna **tiri**? ‘Did he say *tiri*?’ (target word in polar question)
 b. Ur inna tiri. ‘He did not say *tiri*.’
 c. Inna **tnza**. ‘He said *tnza*.’ (target word in contrastive statement)

In each mini-monologue the target words in the polar question and answer (2a and b) was different from the one in the contrastive statement (c). In (a) the target word is in a question. In (c) a different target word is in a statement. Because of the negation in (b), the target word in (c) is explicitly contrasted (corrective focus). In both cases the target word is located at the right edge of the Intonational Phrase. Each target word appeared in both contexts (questions and statements) three times resulting in 336 tokens per speaker (56 target words * 2 contexts * 3 repetitions, total data = 1344 tokens).

Recordings and Analyses

Participants were seated in front of a computer screen and read aloud sentences presented orthographically one mini-monologue at a time. The words were syllabified according to the analysis of Dell & Elmedlaoui (2002) and the syllable nuclei referred to in what follows is based on this analysis. All acoustic material was manually annotated by the first author.

The following labelling criteria were used, involving the annotation of the f0 pitch contour and the identification of the segmental boundaries:

F0 labels: The location of the peaks under investigation was judged by ear and subsequently confirmed by visual inspection of the f0 contour and spectrogram provided by Praat 5.3 (Boersma & Weenink 2012). If the tonal prominence was ambiguous (4.8%) or not detectable at all (6.6%), the datum was excluded. Around the rise contour area we identified local turning points in the f0 contour by hand. We labelled the end of the rise as the high peak (H).

Segment labels: To identify segmental boundaries of the target word in the acoustic waveform, we displayed a speech waveform and a wide-band spectrogram simultaneously. All segmental boundaries of vowels and consonants were labelled at abrupt changes in the spectra at the time 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). Based on these boundaries, we measured the temporal lags between the f0 peak (H) and the beginning of the syllable with which the peak co-occurred.

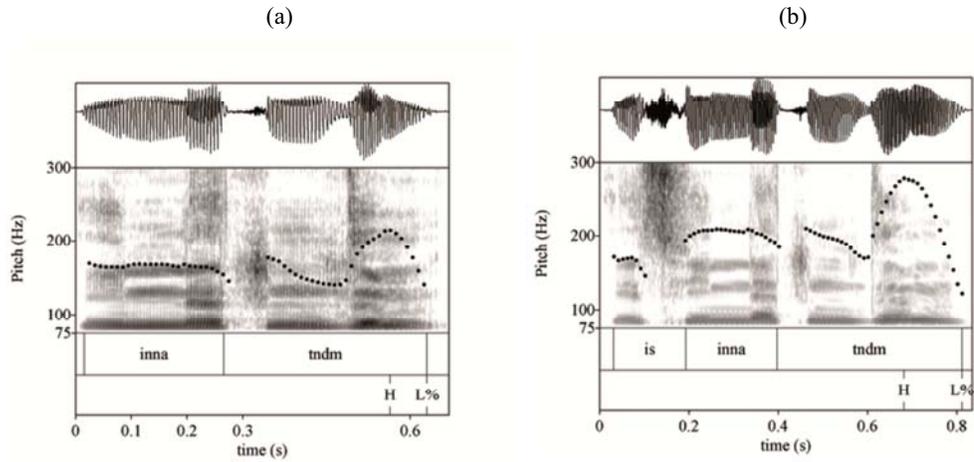
Results

In both sentence types there is a rise to a f0 peak (represented as a H tone) followed by a fall in pitch to a (relatively) low point at the end of the phrase (represented as L%). In general, questions reveal an overall greater pitch range and register and a concomitant steeper rise to the H peak than statements (see Figure 1)

In terms of the location of the f0 peak, the situation is as follows. In productions of target words with only one sonorant nucleus (vowel or sonorant), the H peak is almost always located on this syllable (92%) (this is also consistent with Grice et al. 2011).

Looking at words with two sonorant nuclei available, there is a great degree of variation in peak placement; even across the three repetitions by the same speaker. Part of the alternation is modulated by the weight of the final syllable, degree of sonority of the sonorant nuclei, as well as sentence modality. Since the former aspects of phonological association (sonority and syllable weight) have been discussed in Röttger et al. (2012), we shall focus on the latter aspect here.

FIGURE 1: Waveform, spectrogram, and f0 contours of (a) statement *inna tndm* ‘he said ‘she regretted’ and (b) question *is inna tndm* ‘did he say ‘she regretted?’ produced by speaker M1.



Of the 576 productions of target words with two available sonorants (e.g. *tndm*, cf. Figure 1), there was a general tendency to place the f0 peak on the final syllable, regardless of sentence modality (75% on the final syllable as opposed to 25% on the penult). Along the lines of previous work, we interpret this placement on specific syllables as phonological association. The tendency to locate the peak on final syllables was even stronger for questions (average across all four speakers 87%).

FIGURE 2 (a) Association: mean frequency of occurrence of the f0 peak on the final syllable (U); **(b)** Alignment: mean relative position of the f0 peak within the tone bearing syllable (penultimate and final syllable data pooled). 0% indicates the beginning of the syllable, 100% the end of the nucleus)

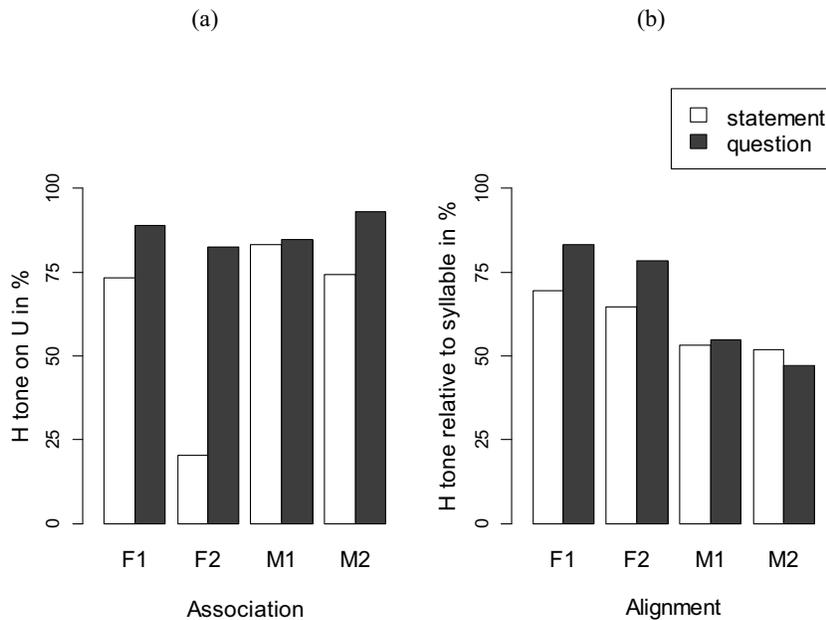
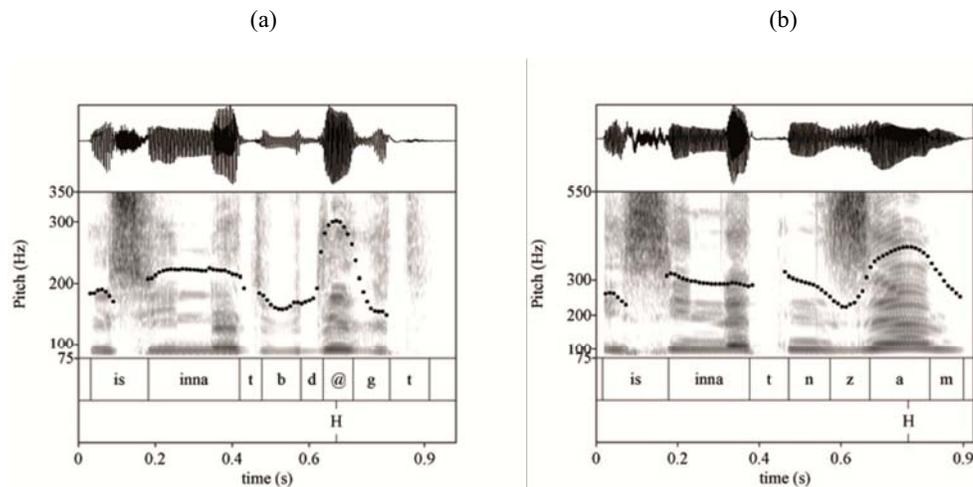


Figure 2 visualises the results for each speaker separately. The graph in 2(a) plots how frequently each speaker places the f₀ peak on the final syllable (U) in statements and questions with target words containing two sonorant nuclei. Although there is a great deal of variation, there is a clear difference in frequency of occurrence at least for three of the four speakers: in questions the peak is more often on the final syllable than in statements (Figure 2a). The strength of this preference is speaker dependent, with F2 showing the strongest asymmetry, M1 showing almost none, and F1 and M2 in between.

The analysis of where exactly in the syllable the f₀ peak is located is interpreted as phonetic alignment. Figure 2(b) plots the relative position of the peak within the tone bearing syllable for each speaker separately. In other words 0% indicates the peak being realized at the beginning of the syllable, 100% indicates the peak being realized at the end of the nucleus. On average, f₀ peaks occurred in the second half of the syllable. In regard to sentence modality, the two female speakers (F1, F2) place the peak later in questions than statements, whereas the male speakers (M1, M2) do not.

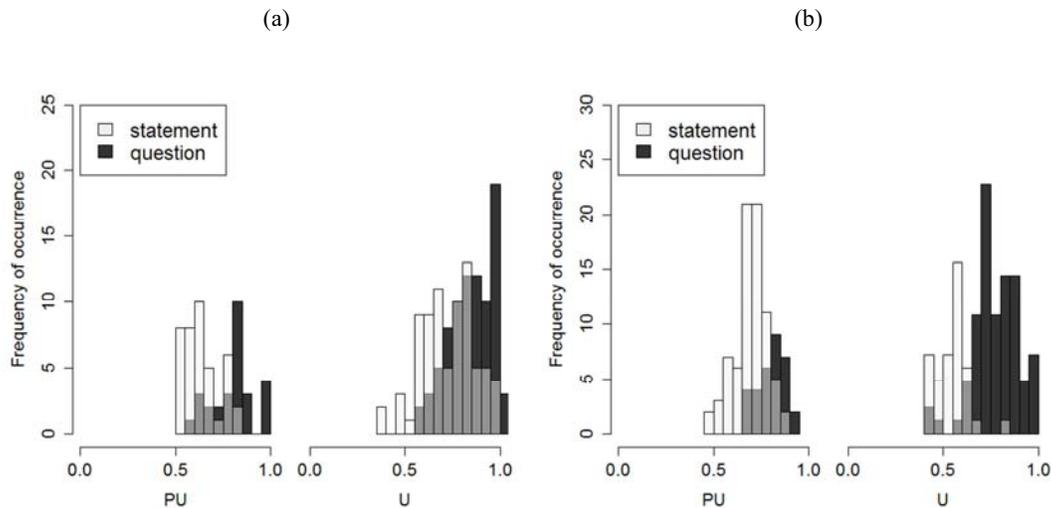
If the target word contained neither a vowel nor a sonorant consonant, it sometimes had no f₀ peak at all. Alternatively, the peak was either on a vocalic nucleus of the previous word (i.e. on the final vowel of *inna*) or on a transitional vocoid (@) within the consonant string of the target word (see Figure 3a). This pattern is consistent with the above mentioned study in which the target word was phrase medial (Grice et al. 2011). From these results, it appears that speakers are able to exploit non-phonological elements to place the peak on the target word. This might be a reflex of the general preference to place the peak towards the right edge of the phrase in our study, as the word is final in the phrase. However, this is not the whole story, as, looking at target words with at least one sonorant nucleus available, the peak was almost always aligned to one of the sonorant nuclei rather than to the last element in the syllable that is periodic enough and possibly loud enough. Even in the presence of a sonorant coda, the f₀ maximum was aligned to the nucleus (e.g. in *tnzam*, see Figure 3b).

FIGURE 3: Waveform, spectrogram, and f₀ contour of (a) question *is inna tbdgt* ‘did he say ‘you are wet?’’ produced by speaker M1 and (b) *is inna tnzam* ‘did he say ‘you were sold?’’ produced by speaker F2.



Looking in more detail at the two female speakers, Figure 4 shows that the distributions of peaks in statements and questions are overlapping. Nonetheless, it is possible to see that these two speakers produce different phonological association patterns, as well as different phonetic alignment patterns for statements and questions, although the strength of the distinction as well as the cue weighting seems to be speaker dependent.

FIGURE 4: Relative position of the f_0 peak within the tone bearing syllable (penultimate and final syllable data pooled). 0% indicates the beginning of the syllable, 100% the end of the nucleus, for F1 (a) and F2 (b). Grey bins indicate overlap.



In sum, the categorical tendencies of the peak to be placed on specific syllables, interpreted as phonological association is reflected in the tendency of speakers to align the peak earlier in statements than in questions (except for speaker M2).

Discussion

Our results point to different competing factors affecting the location of the H peak: Generally, there is a preference to place the peak on the rightmost syllable. This might be accounted for by phonetic grounding: a pitch event near the end of the phrase being a salient position due to recency. This right edge preference competes with communicative function: If the peak is on the final syllable, there is a RISE in pitch over the target word (with a following fall which varies in extent, and can sometimes be absent altogether). Such a rise is cross-linguistically common in polar questions (Grice, Ladd & Arvaniti 2000). If the peak is on the penult, there is a FALL in pitch over the target word, a pattern common in statements (Ohala 1983, Gussenhoven 2004). While this is a strong generalization, the actual data indicate a rather probabilistic picture. Moreover, this discrete asymmetry is reflected in more gradual phonetic detail reflected by phonetic alignment within the TBU. Peaks in questions are aligned later in the syllable than statements. This tendency is highly speaker dependent. The different preferences in questions and statements have to be considered as functionally motivated to distinguish sentence modality. The remaining variability, however, reflects the redundant nature of peak location. Both morphosyntactic devices (question particle) and global pitch cues could already be sufficient for marking sentence modality. The actual peak location phrase finally is thus only one of a number of cues.

The observed tonal events in Tashlhiyt have to be considered as tones associated to prosodic constituents above the word, which may correspond to what is referred to as “a property of units larger than words” (Dell & Elmedlaoui 2002: 14). The fact that there are different preferences for peak placement dependent on sentence modality speaks in favour of such an analysis. In line with the proposed analysis of Intonation Phrase medial tones in (Grice et al. 2011), one could account for this association as a secondary association of IP right edge tones to a TBU. In this analysis, both the H peaks in statements and in questions are edge tones seeking secondary association to a preceding syllable. If the H tone is primarily an edge tone, the distance from the right edge should be kept to a minimum, as this peak also marks the end of a prosodic constituent, as suggested in Grice et al. (2011).

However, this phonological analysis is unable to account for H obtained on transitional vocoids. Since these elements have not been found to be relevant at any other level of prosodic phonology, they should not be able to function as a syllable nucleus, let alone a TBU.

Another way of dealing with the location of H in the absence of sonorous syllables involves the division of labour of phonology and phonetics. If no TBU is available, then the phonological planning does not involve a secondary association of the H tone. The H tone would only have its primary association to the phrase as a right edge tone. This tone would still need to be realised, however. This could be achieved by placing the H peak on a part of the signal which is periodic enough and possibly loud enough to make the pitch peak clearly perceptible. As will have become clear, both phonological and phonetic aspects of tonal placement are at work, and both aspects have to be taken into account for a full account of tonal placement in this language.

REFERENCES

- Boersma, P., and D. Weenink (2012). *Praat: doing phonetics by computer (Computer program)*. Version 5.3, retrieved 10 November, 2012, from <http://www.praat.org/>
- Coleman, J. (2001). "The phonetics of vocoids associated with syllabic consonants in Tashlhiyt Berber", *Transactions of the Philological Society* **99**, 29–64.
- Dell, F., and Elmedlaoui, M. (1985). "Syllabic consonants and syllabification in Imdlawn Tashlhiyt Berber," *Journal of African Languages and Linguistics* **7**, 105–130.
- Dell, F., and Elmedlaoui, M. (2002). *Syllables in Tashlhiyt Berber and in Moroccan Arabic* (Dordrecht: Kluwer).
- Gordon, M., and Nafi, L. (2012). "Acoustic correlates of stress and pitch accent in Tashlhiyt Berber", *Journal of Phonetics* **40**, 706–724.
- Grice, M., Ladd, D.R. and Arvaniti, A. (2000). "On the Place of Phrase Accents in Intonational Phonology", *Phonology* **17**, 143–185.
- Grice, M., Röttger, T. B., Ridouane, R., and Fougeron, C. (2011). "Tonal association in Tashlhiyt Berber", *Proc. ICPhS*, pp. 775–778.
- Gussenhoven, C. (2004). *The phonology of tone and intonation*. (Cambridge University Press).
- Ohala, J. (1983). "Cross-language use of pitch: An ethological view". *Phonetica* **40**, 1–18.
- Pierrehumbert, J., and Beckman, M. (1988). *Japanese Tone Structure* (MIT Press).
- Ridouane, R. (2008). "Syllables without vowels: phonetic and phonological evidence from Tashlhiyt Berber", *Phonology*, **25**, 321–359.
- Ridouane, R., and Fougeron C. (2011). "Schwa elements in Tashlhiyt word-initial clusters", *Journal of Laboratory Phonology* **2**, 1–26.
- Röttger, T. B., Ridouane, R., and Grice, M. (2012). "Sonority and syllable weight determine tonal association in Tashlhiyt Berber", *Proc. Speech Prosody*.
- Stroemer, H. (2008). "Three Tashlhiyt Berber texts from the Arsène Roux archives", in: *Evidence and Counter-Evidence. Essays in Honour of Frederik Kortlandt, Vol. 2: General Linguistics*, edited by A. Lubotsky, J. Schaeken, and J. Wiedenhof (Amsterdam: Rodopi), pp. 389–397.