In other words, the surface differences in pitch derive solely from the mapping of intonational tones to different types of feet; no tonal information is stored in the lexicon. As I show, there is little evidence to assume that the language is tonal. For example, the distribution of stress and pitch is largely predictable on the basis of vowel quality (as [1] state themselves, see e.g. pp. 639). It is argued the foot-based approach not only captures the patterns more directly but also leads to a considerably simpler and more straightforward analysis.

From a broader perspective, the purpose of this paper is to demonstrate on the basis of the Uspanteko data that pitch-based surface contrasts do not necessarily imply that tonal information is stored in the lexicon. In doing so, the analysis contributes to an ongoing discussion on the nature of tone accent oppositions, particularly in Franconian and Scandinavian. While traditional autosegmental analyses attribute the respective accent contrasts to the presence of lexical tone in the lexicon (e.g. [7], [8], [9], [10], for Franconian, [11], [12], [13], [14], [16], for Scandinavian), it has recently been argued that the tonal surface contrasts should rather be analyzed as metrical, foot-based contrasts ([3], [6], [17], [18], [19], [20], [21] for Franconian, [21], [22], [23] for Scandinavian; see also [24] for a metrical analysis of pitch accents in Scottish Gaelic).

1. Introduction

It has long been known that the relationship between phonological units and their phonetic expression need not be one-to-one. This is particularly obvious with respect to the metrical organization of words: for instance, it is now widely acknowledged that linguistic stress is not a segmental feature with invariant phonetic correlates but rather the manifestation of the organization of syllables into feet; i.e., the strong branch of a foot (the foot head) is in some way more prominent than the weak part of a foot (the dependent). The correlates of stress differ across languages, but the most common ones are pitch, duration, and articulatory effort (e.g. vowel reduction in unstressed syllables).

Uspanteko, a Mayan language spoken in Guatemala, shows a remarkably rich interaction between the abovementioned factors, i.e., location of stress, vowel quality and pitch accents / tone. [1] give a detailed analysis of the patterns. Combining insights from previous research ([2], [3], [4]) with data from their own fieldwork, the authors provide an excellent description of the patterns and also show how, from their point of view, the surface contrasts can be derived from the interaction of lexical tone, foot structure (iamb and trochees), and vowel quality. I would like to argue, however, that there is one incorrect basic assumption underlying the approach, viz. that the language has lexical tone. Instead, I propose that both the tonal contrasts as well as the interactions with vowel quality derive from contrastive foot structure in the underlying representation, modeled along the lines of [5], [6].

In other words, the surface differences in pitch derive solely from the mapping of intonational tones to different types of feet; no tonal information is stored in the lexicon. As I show, there is little evidence to assume that the language is tonal. For example, the distribution of stress and pitch is largely predictable on the basis of vowel quality (as [1] state themselves, see e.g. pp. 639). It is argued the foot-based approach not only captures the patterns more directly but also leads to a considerably simpler and more straightforward analysis.

From a broader perspective, the purpose of this paper is to demonstrate on the basis of the Uspanteko data that pitch-based surface contrasts do not necessarily imply that tonal information is stored in the lexicon. In doing so, the analysis contributes to an ongoing discussion on the nature of tone accent oppositions, particularly in Franconian and Scandinavian. While traditional autosegmental analyses attribute the respective accent contrasts to the presence of lexical tone in the lexicon (e.g. [7], [8], [9], [10], for Franconian, [11], [12], [13], [14], [16], for Scandinavian), it has recently been argued that the tonal surface contrasts should rather be analyzed as metrical, foot-based contrasts ([3], [6], [17], [18], [19], [20], [21] for Franconian, [21], [22], [23] for Scandinavian; see also [24] for a metrical analysis of pitch accents in Scottish Gaelic).

The distinct peak in the forms (1e-h) is traditionally assumed to derive from a lexical tone; yet this terminology is misleading, in the sense that the occurrence of the peak is highly predictable: first, it is restricted to the pre-final vocalic
mora, and only permitted if this mora is located in a stressed syllable (generally, only vowels are moraic; yet according to [1], glottal stops in final syllables are moraic as well iff they are followed by another coda consonant). In other words, no distinct peak can occur in final syllables, unless these syllables contain two vocalic moras, i.e., a long vowel.

Second, all words with penultimate stress obligatorily have a distinct peak. Notably, in dissyllabic words with two short vowels, the position of stress (and thus also the occurrence of a distinct peak) is largely dependent on the sonority of the vowels (relatively lower peripheral vowels are more sonorous than relatively higher vowels and schwa, the scale being a >> e, o >> i, u, o); words with falling sonority from the penultimate to the final syllable have penultimate stress (1g), words with rising sonority have final stress (1c). Words with equal sonority can have both patterns (1d, h), with a statistical preference for penultimate stress. Finally, long vowels in final syllables are always stressed (1a, b, e, f), and some of them have a distinct peak (1e, f); long vowels can only occur in word-final syllables. In sum, the distribution of ‘tone’ is largely determined by the position of word stress, which in turn often depends on the sonority of vowels. Consequently, there are only few cases where the occurrence of tone does not follow from external factors.

This largely predictable distribution seems to suggest that the alleged lexical tone may in fact be an intonational tone, predictably assigned to (some) stressed syllables. While the intuition itself is straightforward and simple (after all, it is predictably assigned to (some) stressed syllables. While the alleged lexical tone may in fact be an intonational tone, the position of stress (and thus also the occurrence of a distinct peak) is largely dependent on the sonority of the vowels (relatively lower peripheral vowels are more sonorous than relatively higher vowels and schwa, the scale being a >> e, o >> i, u, o); words with falling sonority from the penultimate to the final syllable have penultimate stress (1g), words with rising sonority have final stress (1c). Words with equal sonority can have both patterns (1d, h), with a statistical preference for penultimate stress. Finally, long vowels in final syllables are always stressed (1a, b, e, f), and some of them have a distinct peak (1e, f); long vowels can only occur in word-final syllables. In sum, the distribution of ‘tone’ is largely determined by the position of word stress, which in turn often depends on the sonority of vowels. Consequently, there are only few cases where the occurrence of tone does not follow from external factors.

This largely predictable distribution seems to suggest that the alleged lexical tone may in fact be an intonational tone, predictably assigned to (some) stressed syllables. While the intuition itself is straightforward and simple (after all, it is common across languages that stressed syllables receive high tone), one obvious question emerges: why is the distinct peak banned from all word-final stressed syllables with short vowels and from some word-final stressed syllables with long vowels (1a-d)? As I shall demonstrate in the following section, the non-occurrence of the distinct peak in some words follows from the interaction of different foot types with intonational tones.

3. Analysis: a metrical approach

3.1. Foot inventory

To analyze the interaction of intonational tone, stress, and vowel quality, I follow the approach put forward in [5], [6] for Franconian: the basic analytical assumption is that foot structure can be contrastive in languages, which in turn can affect the mapping of intonational tones and the interaction with segmental structure. As discussed in detail in [1], Uspanteko provides evidence for both iambic and trochaic footing, and the language has exactly one foot per word. I argue that there are two types of trochees in the language: first of all, there are ‘syllabic’ trochees ($\sigma$-trochees), which span across two syllables; the foot head is the initial syllable. Furthermore, there are ‘moraic’ trochees ($\mu$-trochees), whose foot head is the leftmost mora of a bimoraic syllable. In iambic, the head is always the rightmost syllable, which is either monomoraic or bimoraic; these general foot types, which are in line with standard assumptions on foot structure, are shown in Figure 1.

Intonational tones associate with these feet in different ways, depending on whether a TBU (here: vocalic mora) is part of a foot head domain, or not. The foot head domain comprises the foot head (indicated in Figure 1 with a superscript plus) and all prosodic structure it dominates. In Figure 1, this is illustrated with two types of association lines: moras with straight association lines are part of a foot head domain (either dominated by a foot head, as in a. and c., or a foot head themselves, as in b.). Intuitively, we can refer to these moras as strong moras at the foot level.

Dashed lines indicate that the moras in question are either dependents of a foot (b.) or dominated by a dependent (a., c.): at the foot level, these moras are therefore weak. In Uspanteko, foot type a. is found in words with penultimate stress and obligatory tone (1g, h). Foot type b. occurs in words with final stress and tone (1e, f); the initial syllable in (1e) remains unparsed. Lastly, foot type c. is found in ‘toneless’ words with final stress, which have either one or two moras (1a-d). Following [1], assume that the default foot is the iamb.

The two trochaic feet are either lexically marked or enforced by the grammar, as e.g. in interactions with vowel sonority. Unpredictable stress patterns are stored as metrical templates (lexical stress, see [5] for a detailed discussion).

3.2. Tones and their interaction with foot structure

I argue that Uspanteko has a prominence tone $H^*$, aligned with the foot head, and a word-final boundary tone $L$, which serves to mark the right edge of the (prosodic) word. The mapping of these two tones differs depending on the foot type of a specific item. The different types of tonal associations are shown in Figure 1. Type a. and b., the $\sigma$-trochee and the $\mu$-trochee, display a one-to-one association of tones to TBUs: $H^*$ is associated with the foot head, and $L$ links to the last mora of the word. This leads to a peak on the pre-final mora and a subsequent pitch fall to the final mora. This straightforwardly derives the forms in (1g, h), which are footed as $\sigma$-trochees, and (1e, f), which are $\mu$-trochees. This association is in line with the cross-linguistic observation that a) high tones and metrically strong positions attract each other, and b) low tones and metrically weak positions attract each other (e.g. [25]).

Crucially, I argue that this principle is responsible for the non-occurrence of distinct peaks in iambic feet, that is, in stressed word-final syllables with short vowels (1c, d) and some words with final long vowels (1a, b). In right-dominant iambic feet, the foot head is always a syllable. Therefore, the moras in the head syllable (be it one or two) are dominated by a foot head and therefore strong. Due to their metrical strength, these moras can repel low tone (though, depending on the language, this need not be the case); formally, this can be captured by a constraint stating that low tone is avoided in metrically strong positions (see [23]):

\[(5) \quad ^*H/L: \text{A metrical head is not associated with a low tone}\]

\[
\text{a. } \sigma \text{-trochee} \quad \text{b. } \mu \text{-trochee} \quad \text{c. iamb}
\]

\[
\begin{array}{c}
\text{F} \\
\sigma \sigma \\
\mu \mu \\
H^*L
\end{array}
\begin{array}{c}
\text{F} \\
\sigma \\
\mu \mu \\
H^*L
\end{array}
\begin{array}{c}
\text{F} \\
(\sigma) \sigma^+ \\
(\mu) \mu \\
\text{L}
\end{array}
\]

Figure 1: Association of the prominence tone $H^*$ and the word-final edge tone $L$ to different foot types in Uspanteko.

Expressed in terms of Optimality Theory ([26]), (5) is a high-ranked constraint in Uspanteko. Consequently, the low boundary tone cannot associate with any of the available moras in the stressed syllable of an iambic foot, since these moras are metrically strong. Instead, $L$ remains floating, which is phonetically implemented as pitch lowering in the stressed syllable (in Figure 1, this is represented with a circled floating
L below H*). This is lowering is responsible for the absence of a distinct peak in words of the type (1 e-h) and the absence of a pitch fall; no lexical tone is necessary.

In a next step, I shall show that the analysis can easily account for the predictable interaction of stress placement and vowel sonority; notably, similar interactions between stress and tone have repeatedly been observed across languages (see [27], [28]), while interactions of tone and vowel quality are largely unattested. With respect to the sonority effects in question, it should be mentioned that penultimate stress and distinct peaks are absent in trisyllabic (or longer) underived words; such items always surface with final stress and without tone, i.e., with an iambic foot. This is the case even when the sonority of the last two vowels is falling:

(6) la.xo.ři ‘today’ *la.řo.ři

If the word in (6) were disyllabic, we would expect stress to be located on the penult, since [o] is more sonorous than [i]. How can it be that sonority effects are important for determining the foot structure of disyllabic words but entirely ignored in longer words?

The answer is as follows: in Uspanteko, a foot head needs to coincide with the first or last syllable of a prosodic word (in [26], this is expressed with the constraint EDGEMOST). Furthermore, recall that Uspanteko only allows for final and pre-final stress, which can be attributed to a constraint stating that all feet have to be aligned with the right edge of the word: ALL–FT–RIGHT (this is in line with the assumptions in [1]). It follows from the interaction of high-ranked EDGEMOST and ALL–FT–RIGHT that, in a trisyllabic word, only the last syllable can be stressed. This is demonstrated in the OT tableau in (7): stressing the first syllable would imply that the foot is not aligned with the right edge of the word (cf., the losing candidate c.); stressing the second syllable would mean that the foot head is neither in the initial or final syllable of the prosodic word (losing candidate b.). As a consequence, a constraint requiring the most sonorous vowel to be stressed has to be violated (here, I will simply refer to this constraint as Sonority; for more elaborate discussions on the interaction of stress and vowel quality, see [27] and [28]):

(7) No sonority effect in trisyllabic monomorphemic words

<table>
<thead>
<tr>
<th>/laxori/</th>
<th>ALL–FT–RT</th>
<th>EDGEMOST</th>
<th>SONORITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. → la(xo.ři)</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. la(xo.ři)</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| c. (la.řo.ři) | | | *

In disyllabic words, however, Sonority plays a crucial role: note that feet in such words are always aligned with the right edge of the prosodic word and also have a head that coincides with the first or last syllable (no matter whether the feet are iambic or trochaic). In (8), this is demonstrated for the trochaic word [wa.lib], where the vowel with higher sonority receives stress (the occurrence of a distinct peak follows from the tonal mapping, depicted in Figure 1a).

(8) Trochaic footing enforced by vowel sonority

<table>
<thead>
<tr>
<th>/walib/</th>
<th>ALL–FT–RT</th>
<th>EDGEMOST</th>
<th>SONORITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. → (wa.lib)</td>
<td></td>
<td></td>
<td>*!</td>
</tr>
</tbody>
</table>
| b. (wa.lib) | | | *

Likewise, a word of the type [ʧu.ke] ‘cramp’ will receive final stress / iambic footing because the second vowel has a higher sonority than the first one. By virtue of (5), this word will not have a distinct peak: the boundary L is blocked from the stressed syllable and has to remain floating – it lowers the peak of the associated H (as depicted in Figure 1c). In sum, we can conclude that the foot-based approach provides a simpler explanation for these seemingly complex interactions.

(9) Iambic footing enforced by vowel sonority

<table>
<thead>
<tr>
<th>/ʧu.ke/</th>
<th>ALL–FT–RT</th>
<th>EDGEMOST</th>
<th>SONORITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. → (ʧu.ke)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| b. (ʧu.ke) | | | *!

4. Comparison to the tonal approach

In this section, I would like to provide some arguments why my foot-based is to be preferred over the ‘standard’ tonal approach to Uspanteko, for which I take the analysis in [1] as a point of reference. Most importantly, I would like to argue that the tonal approach needs to rely on some (ad-hoc) assumptions that are unnecessary under the metrical approach.

While some aspects of the analysis might be treated equally well with reference to foot structure (as proposed here) or tone (as in [1]), there are various properties of the system that seem to favor a metrical solution: this concerns for instance (i) the absence of distinct peaks from trisyllabic underived words, (ii) the interaction of stress and vowel sonority, (iii) the fact that all words with penultimate stress have a distinct peak, and (iv) the restriction of distinct peaks to the penultimate TBU. Let us begin with the points i) to iii); [1] derive all these effects from one complex OT constraint, which is given in (10):

(10) Perfect prosodic word (PPW)

Assign one violation mark for every prosodic word ω that does not meet all of the following criteria:

– ω is coextensive with a single foot F.
– The head syllable of F (σ_F) bears tone.
– F is bisyllabic.
– The nucleus of σ_F is at least as sonorous as the nucleus of σ₀, the syllable occupying the weak branch of foot F.

[1] themselves acknowledge that constraint (10) goes against basic principles of OT (or, for that matter, virtually any theory of grammar); they state that “we should prefer an account with many separate constraints over an equivalent account that has one constraint with many clauses” (pp. 622).

As I have shown in section 3, my analysis makes it possible to express the patterns with a small set of separate, general constraints – therefore, it is preferable over the competing tonal analysis. Furthermore, such powerful constraints have the tendency to overgeneralize, and this also seems to be the case with respect to PPW: consider again the forms in (1a, b), which surface with stress on a final syllable that contains a bimoraic vowel. The first mora of the stressed syllable is coextensive with the weak branch of foot F.

The emerging question is: what stops (1a, b) from carrying a bimoraic vowel. The first mora of the stressed syllable is a trochaic [wa.lib], where the syllable occupying the weak branch of foot F.

Comparison to the tonal approach

In this section, I would like to provide some arguments why my foot-based is to be preferred over the ‘standard’ tonal approach to Uspanteko, for which I take the analysis in [1] as a point of reference. Most importantly, I would like to argue that the tonal approach needs to rely on some (ad-hoc) assumptions that are unnecessary under the metrical approach.

While some aspects of the analysis might be treated equally well with reference to foot structure (as proposed here) or tone (as in [1]), there are various properties of the system that seem to favor a metrical solution: this concerns for instance (i) the absence of distinct peaks from trisyllabic underived words, (ii) the interaction of stress and vowel sonority, (iii) the fact that all words with penultimate stress have a distinct peak, and (iv) the restriction of distinct peaks to the penultimate TBU. Let us begin with the points i) to iii); [1] derive all these effects from one complex OT constraint, which is given in (10):

(10) Perfect prosodic word (PPW)

Assign one violation mark for every prosodic word ω that does not meet all of the following criteria:

– ω is coextensive with a single foot F.
– The head syllable of F (σ_F) bears tone.
– F is bisyllabic.
– The nucleus of σ_F is at least as sonorous as the nucleus of σ₀, the syllable occupying the weak branch of foot F.

[1] themselves acknowledge that constraint (10) goes against basic principles of OT (or, for that matter, virtually any theory of grammar); they state that “we should prefer an account with many separate constraints over an equivalent account that has one constraint with many clauses” (pp. 622).

As I have shown in section 3, my analysis makes it possible to express the patterns with a small set of separate, general constraints – therefore, it is preferable over the competing tonal analysis. Furthermore, such powerful constraints have the tendency to overgeneralize, and this also seems to be the case with respect to PPW: consider again the forms in (1a, b), which surface with stress on a final syllable that contains a bimoraic vowel. The first mora of the stressed syllable is coextensive with the weak branch of foot F.

The emerging question is: what stops (1a, b) from carrying a bimoraic vowel, i.e., why do the forms disobey PPW: the words are coextensive with a single foot (an iamb), the foot is disyllabic, and the nuclei of the two syllables are of equal sonority. Following [1], PPW outranks D+P–T, a constraint prohibiting insertion of tone; accordingly, there is no force in the proposed grammar that would stop the insertion of tone in forms of the type (1a, b). In other words, the analysis wrongly predicts that disyllabic words with long vowels in stressed final syllables should always have a high tone, while tone on monosyllabic words with long vowels would be facultative (because these words are not bisyllabic and thus violate (11) anyway). Of course, one might possibly add other clauses to PPW to ensure...
the ‘correct’ outcome; yet this would merely be a technical solution. Unlike the tonal approach, my metrical analysis does not make wrong predictions with respect to these patterns.

Let us now turn to the solution that [1] provide for (iv), the restriction of tone to the penultimate vocalic mora in the prosodic word. From a typological perspective, it seems unusual that a language would limit the occurrence of a lexical tone to the penultimate TBU. [1] derive the effect from three independent constraints: on the one hand, they argue that lexical tones are prohibited in word-final position (NONPIN TBU). Furthermore, they restrict the occurrence of (high) tone to stressed syllables (*UNSTRESSED H). Since stress can never occur more than two moras away from the right edge of a word (ALL FT RIGHT), the combination of these three high-ranked constraints derives the desired pattern that tone will always occur on the pre-final vocalic mora. Yet one might argue that the effect follows from a combination of factors that have no inherent relation with each other – that is, it is more or less coincidental. Arguably, this can be captured more elegantly in a metrical approach, where all tones are by definition tonational, as shown in section 3 – the problem of restricting tone to the penultimate TBU does not arise, and the analysis remains considerably simpler.

5. Conclusion

The purpose of this paper was to show on the basis of data from Uspanteko that tone-based surface contrasts do not necessarily have to derive from underlying tone but can successfully be analyzed as emerging from metrical oppositions, i.e., contrastive foot structure. In the case at hand, I have further argued that the metrical analysis is more elegant than the competing tonal one. Due to its richness of interactions, the Uspanteko case is a particularly clear example of foot-based segmental and tonal alternations: foot structure simultaneously interacts with tone and vowel quality. Similar observations have been made for Franconian, which also shows accent-based vowel alternations, as discussed in [6], [18], and [19].

Due to lack of space, a few further aspects of Uspanteko have not been addressed in this paper: most importantly, some affixes enforce trochaic rhythm, which can override other restrictions; such patterns can be dealt with by assuming metrical templates in the lexicon (similar to how [1] do it, the difference being that they rely on tonal markings (H), while in my approach, this would be expressed by stating that these templates enforce trochaic rhythm). The exact way these trochee-enforcing affixes affect the output differs across word groups. [1] analyze the patterns by assuming different co-phonologies, which could easily be incorporated in my analysis. Alternatively, one could also consider a representational approach, where the surface contrasts would be derived from different underlying structures; here, I will remain agnostic with respect to this issue.

Lastly, it should be noted that an alternative metrical treatment of the patterns has been briefly suggested in [29]; yet the approach does not discuss the sonority effects, which seem crucial for the analysis – therefore, it is, at least for now, difficult to compare these two (metrical) approaches.

6. References