Practical introduction to stress, accent and tone

23rd ABRALIN Summer Institute, UFF Niterói, March 2017

Course materials – Part Two

http://www.lel.ed.ac.uk/~bob/ABRALIN/part-two

Prof. D. Robert Ladd, University of Edinburgh

bob.ladd@ed.ac.uk

Examples indicated by the symbol ⊛ have accompanying sound files. These can be found in the public directory http://www.lel.ed.ac.uk/~bob/ABRALIN/. The filename is given next to the ⊛ symbol.

These materials have been prepared for a three-day mini-course at the ABRALIN summer institute. As such they combine textbook presentation with a variety of original material, especially in Part Two. They lack the scholarly completeness that would be expected of a more formal publication, and some of the ideas still await further development. In the coming months I expect to prepare a journal article based on Part Two or a short monograph based on Parts One and Two. Until then, these materials can be used and cited like any conference handout, but please check with me if you want to quote them verbatim.

Part Two: Reconsidering 'stress'

I. European perceptual biases and illusions

1. The problem

Stress is a real phenomenon in many European languages. Examples from English, Portuguese, Russian and Greek in Part One make it clear that we need to have some way to describe the phonetic/phonological difference between members of minimal pairs like *esta/está*. There is a reasonably clear phonetic basis in many European languages for saying that one syllable in a word or short phrase stands out among the others. Moreover, identifying the primary stress – the single most prominent syllable – in a word or short phrase is a metalinguistic task that many speakers of those languages find intuitive and uncomplicated.

However, there is increasing recognition that stress may not be a useful concept in describing many non-European languages. There is also increasing recognition that speakers of European languages often perceive phonetic differences between syllables in another language and interpret them as differences of relative prominence – 'stress' – even when they are not. In effect, speakers of many European languages **can't help** perceiving the phonetics of a string of syllables without imposing a framework of relative prominence and/or rhythmic organisation. It is now becoming clear that this has affected descriptions of non-European languages written by Europeans; Tabain, Fletcher and Butcher 2014 suggest the term 'stress ghosting' for this illusory perception of stress.

The Australian aboriginal languages, which are Tabain et al.'s main concern, are likely to be a prime example of stress ghosting. Many descriptions of these languages (e.g. Dixon 1980, section 6.2) include statements about rhythmic stresses – alternating sequences of stressed and unstressed syllables in long words. These descriptions are now beginning to be challenged; Tabain et al. show that in Pitjantjatjara, contrary to earlier claims of rhythmic stress, the first syllable of a word has higher pitch and longer duration, but that there are no stress-related acoustic properties of other syllables, nor any evidence of differences of vowel quality or spectral tilt on the word-initial syllable. They explicitly attribute earlier descriptions to the perceptual biases of speakers of European languages:

> It is possible that the expectation that stress will be regular and foot-based (i.e. rhythmic) has a cognitive basis in the case of field linguists who are native speakers of prototypical stress languages, such as English or German. Because of the largely regular and rhythmic structure of lexical stress in their languages, field linguists may expect to hear a similarly regular and rhythmic structure in the languages they are studying – and they hear this stress whether the acoustic cues are present or not. This would then be a case of stress ghosting comparable to stress deafness for speakers whose native language does not have lexical stress. (Tabain, Fletcher & Butcher 2014: 64)

A rather different type of perceptual influence seems to have affected descriptions of the vowel quantity system of Dinka. Alternations of vowel quantity play an important role in Dinka inflectional morphology, and it has now become clear (Remijsen & Gilley 2008) that Dinka has a typologically unusual three-way surface phonetic contrast of vowel length (short, medium, and long); the short vowels are found in the morphologically short 'grade' of a lexically short stem. At least since the mid-20th century, however, there have been suggestions that the short grade involves 'stress' in some way, and Remijsen & Gilley's study aimed to establish acoustic correlates for this elusive property

of the short vowels. Yet they found no phonetic evidence at all for stress, only extremely short vowels (typically 60-80 ms.). What then, they wondered, gives rise to the impression of stress? They suggest that 'it may be due to the native language (L1) of most of the researchers involved. In many languages, including English, vowels with a duration of 30-60 ms. are most likely to be found in weak positions – in syllables that are reduced and unstressed. This is not the case in Dinka…' They go on to discuss an example in their data of a vowel with a duration of 57 ms that nevertheless represents the amplitude maximum of the utterance in which it occurs, and say that because of this combination of extremely short duration and high amplitude it 'may appear to be unusually salient in the context of the framework of reference that an L1 like English imposes on the auditory system.' (p. 338).

Finally, perhaps the clearest case where 'stress' has confounded Eurocentric researchers for decades is that of several Malay languages, including standard Indonesian. During the colonial period of the 19th and 20th centuries Dutch linguists gave contradictory descriptions of the location of stress in Indonesian. They generally noted that stress was 'weak' and not lexically distinctive, but none of them doubted that it was present. More recently, other linguists armed with metrical theories have attempted to provide theoretical accounts of supposedly unusual aspects of Indonesian stress and rhythm (e.g. Cohn 1989, 1993). But experimental work seems to make clear that Indonesian simply **doesn't have stress**.

For example, Odé 1994 asked ten Indonesian speakers (who had all studied linguistics!) to mark boundaries and prominent words in a short recorded story in Indonesian. They found the task very difficult and showed little agreement among themselves on which words were marked as prominent. At that time, Odé was not yet ready to give up on what she calls 'the problem of prominence in Indonesian' – she clearly acknowledged that Indonesian is different from Dutch or English, but continues to assume that given the right methods, we will be able to understand how 'prosodic events' signal prominence in Indonesian. She didn't seem to consider the possibility that Indonesian might not use prosodic events to signal prominence at all. But more recent work carried out by Dutch phoneticians with the same general theoretical background as Odé – in particular, a series of papers by van Zanten and her colleagues (e.g. 2003) – has provided clear evidence that that is the right way to look at things. The same conclusion, for a different Malay language, is announced in the title of a paper by Maskikit-Essed & Gussenhoven (2016): 'No stress, no pitch accent, no prosodic focus: the case of Ambonese Malay'.

In what follows I explore the implications of acknowledging that there may be many languages without anything usefully called stress, and, more specifically, that impressionistic observations by speakers of English or Dutch are not necessarily a reliable basis for describing the prosodic properties of other languages. Accepting this idea leads to simpler and more empirically defensible analyses of individual languages and also to a richer typological framework for understanding the way languages may use pitch and prosodic features.

2. Two illustrations (with sound files)

Before continuing it will be useful to illustrate the effect of native language experience on the perception of stress in other languages. The following two cases involve languages where there is no doubt about the existence of stress but where speakers of many Western European languages (including English and Portuguese) are likely to have trouble hearing stress 'correctly'. If possible, in following the discussion in the next paragraphs you should listen to the linked sound files before reading the explanations.

The first example comes from Modern Greek. Here we have two different renditions of the phoneme string /ɛfijɛ/:

(1a) version 1: ⊛ efije1.wav

(1b) version 2: ⊛ efije2.wav

Most people whose native language is a Western European language hear stress on the first syllable in version 1 and on the final syllable in version 2. However, this is not what Greek speakers think, as we can see from the written form, which is identical in both versions:

(2) έφιγε

Greek orthography consistently indicates stress with an acute accent. The two utterances are simply a statement utterance and a question utterance of the same word, meaning 'S/he left' (version 1) and 'Did s/he leave?' (version 2).

The problem here, for the Western European listener, is that Greek question intonation involves a **low** pitch accent on the stressed syllable together with an edge tone sequence HL% at the end. Such an intonational sequence does not occur in most Western European languages. Consequently the Western European listener tends to interpret the HL% tonal sequence as an accentual H followed by a L boundary tone, which automatically implies that the syllable with the accentual H is the stressed syllable.

(3a) version 1:
```
H* L  L%
|     |
ɛ f i j ɛ
```

(3b) version 2:
```
L*   HL%
|     |/
ɛ f i j ɛ
```

The Greek listener, of course, hears the final peak-and-fall as the mark of a question, and the location of stress is unaffected.

The second case, from Romanian, is more subtle. Here the Western European listener hears the 'correct' syllables as stressed but is usually unable to tell which one is subordinated to the other in the phrase. The text is identical in both versions:

(4) *o să venim la voi?* 'shall we come to your place?'
(FUT COMP come.1PL to you.PL)

There are two accented words, *venim* 'we come' and *voi* 'you (pl.)', and most Western European listeners have no trouble hearing those as prominent (and indeed, no trouble determining that the stress on *venim* is on the second syllable *-nim*). However, there are two versions, a broad focus version questioning the whole sentence (i.e. 'shall we come to your place (or not)?') and a narrow focus version emphasising 'your place' (i.e. 'shall we come to your place (or somewhere else)?'). As in many Eastern European languages, the primary sentence stress in a broad focus question is on the verb, so in the broad focus version the primary accent is on *venim*; also, as in Greek, the primary accent in a question is low, and is followed by a boundary high (on *voi*) which Western European listeners interpret as an accent. In the narrow focus version, not surprisingly, the (low) primary accent is on *voi*, immediately followed by the boundary high, and with a second not-very-high accent on *venim.* Western European listeners are generally unable to hear

which of the two accents is primary, or even to hear any difference between the two versions at all; both appear to have the primary sentence accent on *voi.*

(5a) version 1: L* H*(L%) ⊛ venim-voi1.wav
| |
o să venim la voi? 'shall we come to your place (or not)?'

(5b) version 2: H* LH% ⊛ venim-voi2.wav
| |/
o să venim la voi? 'shall we come to **your** place?'

(Close phonetic analysis shows clearly that the pitch on *voi* peaks earlier in version 1 than in version 2, because the pitch peak in version 2 is preceded by a low accentual target. For a fuller discussion of the phonetic details of this typically Eastern European question intonation, see Grice et al. 2000, and Ladd 2008, section 2.5.)

Together, these two examples illustrate the powerful effect of native speaker biases in perceiving relative prominence, even in listening to related languages with broadly similar prosodic systems. This should alert us to the possibility that the way we perceive prosodic properties of genealogically and typologically more distant languages may be seriously distorted compared to what the native speaker hears.

II. Three case studies

Although I have little direct experience with the indigenous languages of the Americas, my impression is that many descriptions of these languages treat stress and other word-prosodic features in ways that depend on Eurocentric ideas and Eurocentric perception. There appear to be many cases in which field linguists have attempted to incorporate their phonetic impressions of 'stress' into the phonological analysis of American languages where stress actually plays no role in the phonology. Because these impressions often involve the location of local pitch peaks, the confusion has certainly been made worse by typological uncertainty about what 'pitch accent' means.

In this section, I sketch three case studies of American languages in which a less Eurocentric understanding of stress seems to lead to a simpler analysis and a clearer typological picture. Two of these (Chickasaw and Mapudungun) are languages in which linguists have talked about stress but which can easily be reanalysed without any reference to stress whatsoever, only to pitch accent. The other (Mixtec) involves a group of languages with complex lexical tone systems and dense tonal specifications that nevertheless appear to have prosodic structures with phonetic correlates that resemble many aspects of European stress systems.

1. Chickasaw

Chickasaw is a nearly–extinct Muskogean language, originally spoken in the Southeastern United States and then, for two centuries following the forced relocation of its speakers, in Oklahoma. In the last decades of the 20th century the language was extensively documented by Pamela Munro of UCLA and Catherine Willmond (a Chickasaw native speaker), and its prosodic features have been investigated in detail by Matthew Gordon of the University of California at Santa Barbara. The following summary (and some cases reanalysis) is based on my reading of Gordon's papers and correspondence with him.

Syllable weight and rhythmic lengthening: Chickasaw has a phonemic distinction between long and short vowels, and its prosodic system makes reference to the distinction between light syllables (CV) and heavy syllables (CVV, CVC, CVVC). Specifically, in most sequences of two light syllables (CVCV), the second vowel is lengthened, so that in effect the second syllable becomes heavy. However, this 'rhythmic lengthening' is affected by morphological structure; it does not apply to word-final vowels, and within the word it does not apply to prefixes. Gordon's phonetic research (especially Gordon & Munro 2007) shows that rhythmic lengthening generally results in a vowel that is not quite as long as a phonemic long vowel but clearly different from an unlengthened short vowel.

(6) *Examples of vowel length distinction:*

/waka:/ 'fly [verb]'	/wa:ka/ 'be spotted'
/fala/ 'crow'	/fala:/ 'it is long'
/kola/ 'it is dug'	/ko:li/ 'he breaks it'

(7) *Examples of rhythmic lengthening* (indicated by IPA half-long [ˑ]):

/asabikatok/ 'I was sick' ➔ [asaˑbikaˑtok]
/asabika/ 'I am sick' ➔ [asaˑbika] (no lengthening in final syllable)
/imapilalitok/ 'I helped him/her for him/her' ➔ [imapiˑlaliˑtok]
(not [imaˑpilaˑlitok], because /im-/ is a prefix)

Intonation and pitch accent: Unlike the typical utterance-final intonational patterns seen in European languages, in Chickasaw ***non*-interrogative** utterances (including words in citation form) generally end on a **high pitch**; the pitch on **questions** (including WH-questions) generally **falls sharply** from a peak near the end. In **non-final prosodic phrases** there is usually a rise at the beginning followed by a high plateau and a fall at the end. Gordon (2005: 308) notes that 'the nuclear pitch accent [is] consistently (even in different semantic contexts) realised as a high tone', unlike some European intonation systems, and like some 'pitch accent languages' (e.g. Japanese). In addition, some verb forms involve a local high tone on a specific syllable, which Gordon calls a 'morpholexical pitch accent'. Morpholexical accents can only occur on heavy syllables.

Figure 1 on the next page shows both the non-final and the non-interrogative final intonation patterns: the plateau-plus-fall pattern of non-final prosodic phrases can be seen on /aba:nompiʃtanompoliat/ 'the preacher'; while the non-interrogative final rise can be seen on /maliˑli/ 'he runs'. Gordon describes the latter pattern as resulting from a H pitch accent and a H boundary tone on the final syllable; I return to this point in the alternative analysis given below.

Figure 2 shows two different examples of the interrogative final fall, one with the peak on the final syllable and one with the peak on the antepenult. Together they show that the final fall begins with high pitch (presumably an intonational pitch accent) on the last heavy syllable in the word. This is essentially the description of interrogative intonation given by Gordon (2005: 310ff).

Figure 3 shows two different questions involving verbs with morpholexical pitch accents. In Figure 3a, the expected intonational high tone on the last heavy syllable /-jaʔ-/ appears to be suppressed by the immediately preceding morpholexical high, whereas in Figure 3b the high tone on the last heavy (rhythmically lengthened) syllable is present, but lowered (downstepped) relative to the earlier morpholexical accent on /hoj-/.

Stress: The analysis presupposed in Gordon's work, which appears to be based largely on the work of Munro and her colleagues, refers to 'stress' on certain syllables. In general, **all heavy syllables** (including **rhythmically lengthened** syllables) and **all word final syllables** are said to be stressed. (This description can be expressed in a metrical analysis in which the syllables are parsed into iambic (weak-strong) feet and heavy syllables can (but need not) constitute a foot on their own; rhythmic lengthening can be seen as a natural concomitant of such a metrical structure.) One syllable in the word is further designated as having primary stress, and the other stressed syllables are considered to be secondary. Gordon (2004:7-8) states that the last CVV syllable in the

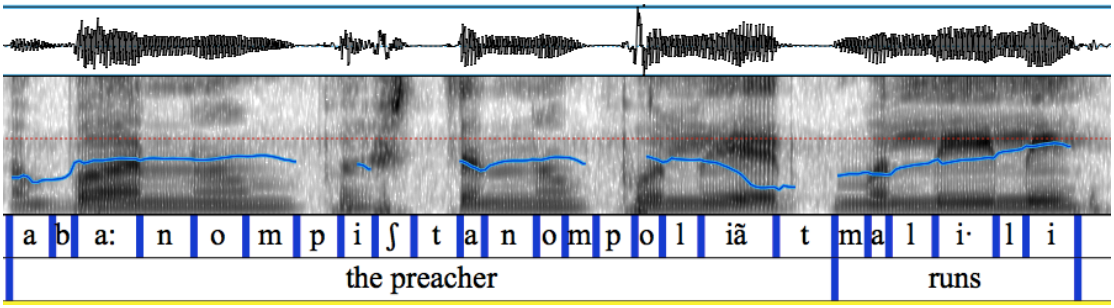

Figure 1

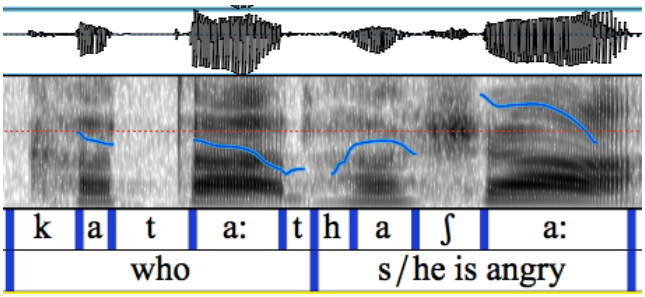

Figure 2a

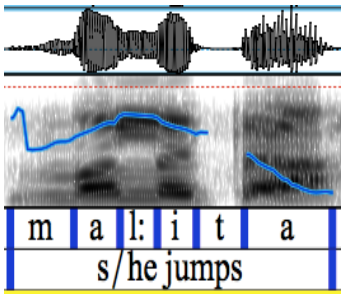

Figure 2b

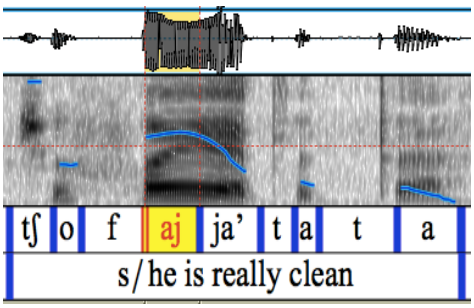

Figure 3a

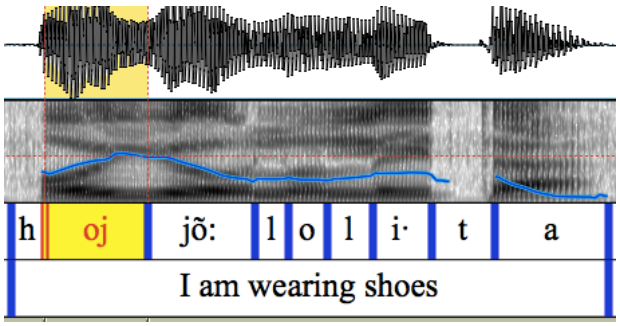

Figure 3b

Segmented pitch tracks of selected utterances from Gordon 2005. Displays are adjusted to approximately the same time scale on the x-axis; the utterance in Figure 2a is just under 1 second long. In Figure 3a and 3b the vowel bearing the morpholexical pitch accent is highlighted in yellow. See text on previous page for details.

word has primary stress, and that if there is no CVV syllable then the final syllable is primary. However, under certain intonational conditions a light final syllable can lose its primary stress and become secondary, with an earlier heavy syllable taking primary stress. For the most part these generalisations appear to be based on impressionistic transcription by phonetic fieldworkers who are native speakers of English.

Gordon 2004 is a serious attempt to get beyond such impressions. He reports a very thorough study of the contribution of duration, intensity, and fundamental frequency to signalling stress; unusually for such studies, he also considers vowel centralisation and various consonant lenition effects. Nevertheless, his conclusions are disappointingly typical: he reports that there are small but statistically significant differences in mean duration, intensity, fundamental frequency, and vowel quality between stressed and unstressed syllables, but he also notes that different speakers appear to use the various acoustic cues inconsistently or in different ways. He concludes that his results 'point to three levels of stress: primary stress, secondary stress, and lack of stress' (23), though he actually fails to find consistent correlates of the distinction between primary and secondary stress and observes that speakers' productions vary with respect to which syllable has primary stress. He also notes that his speakers 'do not have strong or consistent intuitions about which syllables are stressed' (2), and elsewhere (Gordon 2005: 304) he suggests that 'the most reliable diagnostic for distinguishing between secondary stressed and unstressed syllables is a series of syncope processes affecting light, unstressed syllables, i.e. non-final CV.'

An alternative analysis: This rest of this section briefly sketches an analysis of Chickasaw that makes no reference to stress at all, consistent with the fact that native speakers lack intuitions about stress. I have no first-hand experience with the language; this proposal is intended only as a demonstration that the idea is not *prima facie* ridiculous.

1. Weight, not stress: The prosodic system of the language depends in a number of ways on syllable weight. The only prosodic distinction that it is absolutely necessary to make use of is 'heavy' vs. 'light', though it may be, as Gordon suggests, that a further division of heavy syllables into CVV and CVC is phonologically important. Rhythmic lengthening is a real phenomenon, but it is a matter of syllable weight. Nothing is added by describing all heavy syllables as stressed. I do not assume that descriptions in terms of stress have no basis in phonetic reality, but only that researchers who talk about stress in Chickasaw are mostly native speakers of English for whom attending to fine distinctions of stress is an essential part of their linguistic competence. Specifically, four claims about stress in Chickasaw plausibly have their basis in Anglophone percepts.

1. The claim that word-final syllables are stressed is probably based on the fact that in citation forms these syllables are normally the highest in pitch. An alternative account of this pitch pattern is given below.
2. The claim that primary stress preferentially seeks the last CVV syllable of a word, but not a CVC syllable, is probably based on the greater sonority of CVV syllables and the effect it has on Anglophone percepts.
3. The claim that *in questions* primary stress may 'shift' from the final syllable to an earlier heavy syllable is based on the fact that the last heavy syllable in questions bears a pitch pattern very similar to that of a prominent sentence-final pitch accent in English. (It is also due to the theoretical expectation, again based on European languages, that nuclear pitch accents associate with primary stressed syllables.)
4. More generally, the claim that there is a distinction between primary and secondary stress is based on Anglophone percepts, reinforced by the strongly hierarchical assumptions of metrical stress theory – which itself may simply embody Anglophone expectations. In correspondence with me (February 2017) Gordon acknowledges that the evidence for culminative primary stress in many American languages is weak.

2. Location of intonational pitch accents: Certain features of Gordon's account of the location of intonational pitch accents are affected by the attempt to make typological

expectations fit the transcription-based stress data. He naturally expects (see point 3 immediately above) that the nuclear pitch accent will associate with a primary stressed syllable. Since the question pattern clearly requires two syllables or a long vowel for its realisation (Gordon 2005: 311f), there is a problem in sentences that end in a 'stressed' word-final light syllable. Gordon's solution, as we just saw, is to assume that the primary stress 'shifts' to an earlier heavy syllable, which is, at the very least, typologically unexpected. But if we do not assume that the final syllable is 'stressed' and instead simply assume that the nuclear pitch accent seeks the last heavy syllable, this problem disappears.

However, this solution creates a different problem, because in non-interrogative intonation, as noted earlier, the highest pitch is normally on the final syllable (e.g. /mali·li/ in Figure 1). As we saw, Gordon treats this final high pitch as the reflex of a H pitch accent and a H boundary tone both occurring on the final syllable. If his analysis is correct, then the suggestion that the nuclear pitch accent seeks the last heavy syllable fails. However, if we accept that the supposed primary stress on a light final syllable is merely an Anglophone illusion, then there is nothing to prevent us from saying that the nuclear pitch accent **invariably** associates with the **last heavy syllable**. Statement intonation on /mali·li/ would be analysed as involving a H pitch accent on the penultimate (rhythmically lengthened) syllable and a H boundary tone on the final syllable. Question intonation is the same, but with a L boundary tone instead of H. This is obviously much more in line with normal typological expectations about intonational pitch accents.

The fact that morpholexical pitch accents can override or suppress intonational ones and that rhythmic lengthening is subject to morphological constraints is perfectly consistent with typological expectations about 'pitch accent languages'.

2. Mapudungun

Background: Mapudungun is an endangered but still widely spoken language of Chile and Argentina. It is considered an isolate (i.e. it has no known genealogical relation to any other language). Descriptions of the language, many of which mention stress or accent, date back to the early colonial period (early 17th century). Most of them (all by speakers of European languages) say that stress is 'weak' or 'inconsistent'. There does not seem to have been a great deal of change in the system since the earliest descriptions. Most of what I say here is based on discussions with Ben Molineaux, and on his papers both published and still unpublished. Molineaux's phonetic data establish clearly that the only consistent acoustic correlate of stress is a pitch excursion, which means that in Beckman's terms we are dealing with a 'non-stress' accent. He also notes that, in words with more than one stressed syllable, native speakers have no intuitions that one of the stresses is 'primary'.

Two incompatible analyses: Under the name Araucanian, the language has been repeatedly mentioned (and occasionally discussed in more detail) in the theoretical literature on metrical stress (e.g. Hyde 2002, Gordon 2002, McGarrity 2003, Hermans 2011) because of the supposed typological interest of its stress system. However, virtually all of this theoretical discussion is based on a single short paper (Echeverria & Contreras 1965) that seems to contradict most of what other writers (including those of the colonial period) have said about stress in Mapudungun. Only recently (de Lacy 2014, Molineaux 2016) has attention been drawn to the doubtful empirical adequacy of the Echeverria and Contreras paper (henceforth E&C). Meanwhile, the theories march on.

The basic empirical disagreement between E&C and everyone else is as follows. E&C say that primary stress goes on the **second** syllable of the word and secondary stresses then

occur on every other syllable thereafter (e.g. the 4th and 6th syllables). Most other writers say that stress goes on the **last vowel that is followed by a consonant**, i.e. on the final syllable if it is closed, and on the penultimate syllable if the final syllable is open. From the point of view of metrical stress theory, the E&C description makes Mapudungun look like a system based straightforwardly on **iambic** feet built from the **beginning** of the word. Everyone else's version might be treated as **trochaic** (which is what Molineaux prefers), and clearly prioritises the **end** of the word. Although E&C's description would appear to be utterly incompatible with others, Molineaux points out that starting from the beginning and starting from the end need not always yield different results, especially in short words:

(8) na'mɨn 'foot' ḻaf'keṉ 'sea'
ma'wiθa 'woodland' pi'fɨʎka 'two-tone flute'

However, in an empirical study of stress judgements (2014) he also notes that two-syllable words with final open (CV) syllables seem to be variable, with stress sometimes on the first syllable and sometimes on the second. This variability is not expected under either description.

Additional stresses: A further empirical issue is that no one except E&C has suggested the presence of rhythmic secondary stresses. Many of the descriptions that place stress on the final or penultimate syllable note that there may be other stresses earlier in long words, but they point out that these seem to be sensitive to morphology, not to rhythm. Molineaux (ms.) describes the morphological conditions in some detail; I mention only the key points of stress in verb forms here.

- First, in long verb forms – which begin with the verb root followed by various suffixes – there is an additional stress on the second syllable of the root. The root is sett between curly brackets in the transcriptions. For example:

  {θew'ma}ka'kij 's/he is usually making'
  {ɨ'ʈʂif}tu puke'lajmi 'you don't usually throw ___ back here'

  Note that the second-syllable stress in such forms is consistent with E&C's description.

- Second, if the root has only one syllable, the additional stress is on that syllable:

  {'lef}pu'lej 's/he is running here'

  This means that we might equally say that the additional stress goes on the **last** syllable of the root rather than the **second** syllable.

- Third, there are certain verb suffixes that attract stress. Molineaux suggests that these suffixes, which immediately follow the verb root, combine with the root to form an extended stem, and that the stress goes on the last syllable of the stem (the extended stem is enclosed in curly brackets and the stressed suffix is underlined):

  {tuku'ŋe}la'fuj 'he/she/it was not worn'

  However, I will suggest an alternative analysis for these forms below.

- Fourth, if stress on the root (or the 'extended stem') would be on a syllable immediately adjacent to the sressed final or penultimate syllable of the word,

one of them (usually the one on the root) is deleted, and the form has only a single stress:

{leli}'fijmi 'you watch him/her/it' < {le'li}'fijmi
{eluŋ'ma}fijmi 'you give him/her/it ___' < {eluŋ'ma}'fijmi

A pitch accent analysis: Whatever is going on with these additional stresses, it is clearly not based on parsing the string of syllables into trochaic feet. The additional stresses are governed by morphology, not prosody. Also, as mentioned above, the acoustic cues to stress consist exclusively of local pitch peaks on the affected syllable, and Molineaux finds that native speakers have no intuitions about which of the stresses in a long word is primary or more prominent. Consequently, I propose that we are dealing with a morphologically conditioned pitch accent, and that there is no need to invoke stress – or metrical feet – at all. The principles are simple: A word-level 'right-edge' pitch accent goes on the last syllable of the word unless it is an open syllable, in which case it goes on the penultimate. A word-level 'left-edge' pitch accent goes on the last syllable of the root. Pitch accents on adjacent syllables are not allowed, and one of two adjacent accents is deleted.

Interestingly, this analysis suggests a reason for the prosodic variability of two-syllable words, which in unexplained in stress-based analyses. In a pitch accent analysis, the variability results from a conflict between the 'right-edge' and 'left-edge' pitch accent principles. In a two-syllable word with a final open syllable, the 'right-edge' accent (on the last vowel followed by consonant) will appear on the initial syllable; the 'left-edge' accent (on the last syllable of the root) will appear on the final syllable. Only one of these accents can be realised because of the prohibition on adjacent pitch accents; the variability arises from opposite choices of which accent to delete.

The pitch accent analysis, combined with the prohibition on adjacent accents, also suggests an alternative approach to the accented suffixes that Molineaux sees as part of an 'extended stem'. Perhaps these suffixes are inherently accented, i.e. perhaps they are specified in the lexicon as having an associated H tone, similar to the morpholexical accents in Chickasaw. Since they immediately follow the verb root, they will result in two adjacent pitch accents, namely the left-edge accent on the last syllable of the root and the lexically specified accent on the suffix; this clash is resolved by deleting the left-edge accent.

The analyses just suggested are obviously based on extremely limited data and would need to be tested against a larger set of morphological structures. For example, if there are any prefixes or other proclitic elements on verb forms, we could determine whether the left-edge pitch accent seeks the second syllable of the word (as claimed by E&C) or more specifically the second syllable of the verb root. As with the alternative analysis of Chickasaw, my proposal is only intended to show that the Mapudungun data might be accounted for with a relative simple set of principles for locating phonetically 'non-stress' pitch accents, and that no independent notion of stress or any metrical structure is needed.

3. Mixtec

Background: What is frequently referred to as 'Mixtec' is not a single language but a group of related Otomanguean languages spoken in Mexico. As is frequently true in such cases, there is some disagreement over how many different Mixtec languages need to be distinguished, but it is roughly accurate to say that variation within the Mixtec group is

comparable to that within the Romance group (di Canio et al., ms). This means that, while we may expect broad typological similarity in many respects, specific claims about one Mixtec variety may not carry over to others. Mid-20th century impressionistic work on Mixtec was based primarily on the varieties of San Miguel el Grande (ISO code MIG), San Esteban Atatláhuca (MIB) and Santo Tomás Ocotepec (MIE); the three varieties under consideration here are Yoloxóchitl (XTY), Ayutla (MIY), and in particular Southeastern Nochixtlán (MXY).

Stress: All Mixtec languages have rich lexical tone systems, with complex tone sandhi in many varieties, and with tone specifications on virtually every syllable in all varieties. Some typological preconceptions might therefore lead us to expect that stress would play no role in Mixtec, but in fact there is good evidence for stress in at least some varieties (di Canio et al. ms. on Yoloxóchitl, McKendry 2013 on SE Nochixtlán.) Because pitch is 'busy' conveying tonal contrasts, the acoustic basis of Mixtec stress seems to involve primarily duration and some segmental effects (e.g. fortition of /j/ to [ʒ] in stressed syllables in SE Nochixtlán). Di Canio et al.'s study shows that Yoloxóchitl has fixed stem-final stress and McKendry's study shows that in SE Nochixtlán stress is stem-initial.

However, an influential impressionistic study on Ayutla by Pankratz & Pike (1967) claimed that there is a major interaction between stress and tone. Simplifying slightly, their word stress rules place stress on a H-tone syllable followed by a L-tone syllable; failing that, on a M-tone syllable followed by a L-tone syllable; failing that, on the first H-tone syllable of the stem; failing that, on the first syllable of the stem. In hindsight these rules look remarkably like predictions about where native speakers of English would be likely to hear stress given specific pitch contours. Nevertheless, the Pankratz-Pike version of Ayutla stress is central to an influential theoretical proposal by Paul de Lacy (2002) about supposedly universal principles of tone-stress interaction. As with the theoretical over-reliance on Echeverria & Contreras's questionable data from Mapudungun, it appears likely that here, too, theoretical speculation has run far ahead of any reasonable empirical base.

Stress in SE Nochixtlán: The following sections summarise the findings of McKendry's 2013 study, which was based on careful instrumental measurement of controlled speech material. McKendry's first instrumental study, based on laboratory productions of 66 carefully controlled sentences spoken by 4 different speakers, seems to confirm impressionistic observations about stress on the first syllable: in non-contrastive contexts, the vowel of the first syllable of the word is about 25% longer than the vowel of the second, and also has greater intensity. This association between duration and intensity is consistent with interpreting the duration difference as 'stress'. However, neither duration nor intensity is affected by tone. This suggests (contra Pankratz & Pike or de Lacy) that **stress is completely independent of tone**.

McKendry's second instrumental study makes an even stronger case for interpreting these durational effects as 'stress', by showing that Mixtec stress, like stress in many European languages, exhibits **sentence-level effects of information structure**. She placed test words in different discourse contexts designed to make them 'focused' or 'contrastive' or to place the contrast on a word other than the test word (i.e. contexts like *I bought X, not Y* or *I didn't sell X, I bought it.*) These are exactly the kinds of contexts in which many European languages manipulate 'sentence stress' to convey focus and contrast. She found that focus and contrast **increase the durational effects** of stress. As with sentence stress in some European languages, she also found that the effects differ between nouns and verbs.

- Results for **nouns:** When a test noun occurs in a position of focus or contrast, the difference in duration between the first and second syllable is substantially increased (first syllable vowel 60-70% longer than second syllable vowel). When the contrast is on the verb, the duration difference **on the test noun** is similar to the neutral contexts in the first instrumental study (approx. 25% longer).
- Results for **verbs**: When the noun is neutrally focused or when the verb is contrastive, the vowel of the first syllable of the test verb is about 25% longer than the vowel of the second syllable, exactly as with nouns in the first instrumental study. However, when the noun is explicitly contrastive (not simply neutrally focused), the test verb **loses its stress:** in these contexts there is no duration difference between the two syllables.

There are marginal effects of intensity in these contrastive conditions that are similar to what was observed in the first instrumental study. Once again, there is no effect of tone. Di Canio et al.'s study reports similar effects of focus and contrast in Yoloxóchitl, including effects on duration some interestingly complex effects on pitch range.

Finally, in a third instrumental study, McKendry looked at durational effects in compound verbs (two-verb sequences where the first verb becomes a kind of auxiliary). She found that the first syllable of the **main verb** (the second verb) is longer than its second syllable, but there is no difference in duration between the two syllables of the **auxiliary verb**; moreover, the syllables of the auxiliary verb are somewhat shorter than the **second** ('unstressed') syllable of the main verb. This is consistent with an analysis in which the main verb has the prominent stress in the compound as a whole, and retains the stress on its first syllable, while **the auxiliary verb loses its stress**. This is strikingly comparable to the stress subordination in an English compound like *babysitter*, except that the phonetic manifestation of the stress distinction involves duration in Mixtec and pitch accent in English: *baby* has the stress for the compound as a whole, and attracts the pitch accent to its stressed syllable (*ba*-), while there is no pitch accent on the stressed syllable of *sitter*.

McKendry also suggests that related phenomena are found in genitive noun phrases (*the man's dog*, etc.), but with a difference: first, the more prominent constituent **precedes** the less prominent one, and second, the less prominent one is not completely de-stressed. Specifically, the head (the possessed) precedes the genitive noun (the possessor); both nouns retain their internal durational difference between the first and second syllables, but the difference in the head noun is greater. However, she does not provide instrumental data in support of this observation.

Stress in tone languages generally: McKendry's data (and di Canio et al.'s data on Yoloxóchitl, which I have not presented in detail) seem to make clear that stress and lexical tone can coexist in the same language largely without interacting. Mixtec pitch is devoted almost exclusively to conveying lexical tonal distinctions; stress, cued principally by duration, serves to convey prosodic organisation and information structure. That is, the phrase-level function of stress in Mixtec is similar to that in many European languages, but unlike in the European languages, stress has no relevance for pitch phonology, and there is no analogue of European pitch accents. It may be significant that Mixtec stress appears to be fixed on a specific syllable of the stem, so that it is not being used at the word level to make distinctions like English *insight/incite* or Portuguese *esta/está*, only at the phrasal level.

III. Where do we go from here?

If stress is not universal, and if the wide application of metrical theories of prosodic structure is empirically suspect, there are important opportunities for rearranging our prosodic typology. I briefly explore these in this final section.

1. Accent

I assume a general definition of accent inspired by Beckman's work: accent is the **relative** phonological prominence or strength of a specific syllable in context. In the first instance the prominence is a phonological abstraction; its phonetic manifestations can be rather varied. Phonetically, a probably major typological divide is between 'stress' accents and 'non-stress' accents, the former being cued by a variety of phonetic properties related to greater articulatory effort and the latter being cued primarily by specific pitch configurations. Phonologically, the distribution of accents is based on principles that involve prosodic structure. It is probably possible for this structure to be extremely flat (e.g. a string of phonological words in which a H tone is attached to the penultimate syllable each phonological word) and it is certainly possible to have hierarchical prosodic structure of considerable depth. It seems plausible that there is an association between flat structure and non-stress accent (e.g. Japanese) and between stress accent and hierarchical structure (e.g. English), but that may an artefact of our typological sample.

A further important distinction between accentual systems is whether accent is obligatory or not. There are certainly systems in which domains can differ in whether or not they have an accent; Japanese unaccented words are a clear example of this possibility. Again, there seems to be a link between hierarchical depth of structure and obligatory accent: a prosodic domain that incorporates several levels of structure seems more likely to require an obligatory peak of prominence than a domain in a flatter structure. This connection between 'culminativity' and hierarchical structure seems obvious – the hierarchy both facilitates and forces the identification of a single peak – but again it may be an artefact of our typological sample.

2. Tune-text association

I start from the assumption that the relation between pitch and segments in any language is best formalised in roughly autosegmental terms, with a string of tones that is associated with the segmental string in specifiable ways based on a variety of principles. An important part of our typology then concerns the principles that govern tune-text association.

(1) TONAL PRINCIPLE

Prototypical tone language. Syllable tone is specified by lexicon or grammar; most syllables have tone specifications. There may be several phonemic contrasts of tone. (Chinese, Yorùbá).

(a) There may also be intonational boundary tones (Chinese, Thai)
(b) There may not be any boundary tones (Yorùbá)
(c) In context, tones may move from one syllable to another (some Bantu & Mixtec languages)

Limited tone language. Syllable tone is specified by the lexicon or the grammar, but many syllables may lack tone specifications. Tone contrast may be limited to H vs. L or even presence vs. absence of specific tone. Possible tone bearing syllables may be restricted, e.g. only certain affixes, only stems, only heavy syllables. This type grades

into 'lexical pitch accent language' (see 'Continuum (?) from tone to pitch accent' below). (Many Amerindian languages?)

(2) ACCENTUAL PRINCIPLE

Lexical pitch accent language. Accented syllable always bears a specific tone (e.g. H), regardless of intonational context; tone-bearing syllables may have no other features of phonetic prominence. Domain structure typically flat; unaccented domains may be permitted. (Japanese).

Culminative non-tonal ('stress') language. Accented syllables are ***potential*** bearers of tone. Presence or absence of tone and choice of tone (e.g. H vs L) is specified intonationally, not lexically. Syllables specified as potential tone bearers often have other marks of phonetic prominence ('stress') even when they do not have a tone associated with them. At some level (or levels) of structure, well-formed domains *must* have a potentially tone bearing syllable, i.e. unaccented domains are not permitted. (English, Portuguese).

Boundary-marking non-tonal language. 'Accent' gravitates to morphological or prosodic boundaries. Presence or absence of tone and choice of tone (e.g. H vs L) is specified intonationally. Syllables one which boundary tones are realised may sound 'stressed' to speakers of languages with stress. (French, Persian(?), Indonesian).

3. Mixed types and possibly frequent diachronic shifts:

Loss of stress to boundary tones: Stress/intonational accent can be reinterpreted phonologically as edge pitch if post-tonic syllables are lost through sound change (Latin > French). It's probably also possible for edge pitch to be reinterpreted as intonational pitch accent (Hungarian?).

Pitch accent becomes stress accent: Lexical pitch accent can be reinterpreted phonologically as stress/intonational accent if (a) accent is already culminative (i.e. if there are no unaccented words), and (b) intonationally specified tones come to associate with accented syllable (Ancient Greek > Modern Greek?).

Continuum (?) from tone to pitch accent: There may be a kind of continuum between a prototypical tone language and a prototypical lexical pitch accent language. The tone language end of the continuum has greater tonal density (tone/syllable ratio approaches 1:1) and more tonal phonemic distinctions; the pitch accent language end has minimal tonal density (less than one tone per word: maximum of one tone per word and possible toneless words) and minimal tonal contrast (e.g. any syllable either has or does not have a tone, usually H, but possibly L).

European 'pitch accent' languages: Languages like Swedish need not involve tonal specifications in the lexicon, though they are often analysed that way. I follow Morén-Duolljá 2013 is assuming that Swedish 'tonal' contrasts involve only abstract prosodic boundaries (of the kind that are needed to account for differences like *nitrate* vs. *night-rate* in English), which cause the intonational tonal specifications to be aligned phonetically in different ways with the segmental string. In either analysis the distinction involves the presence or absence of a phonological abstraction. In other respects these languages are culminative non-tonal stress languages.

Stress in tone languages: Culminative prominence structure is not incompatible with lexical tone; it's simply that in a language with tonal specifications on most syllables,

prominence cannot be realised by means of intonational pitch accent, but only by other phonetic cues to prominence and/or prosodic structure (e.g. various Mixtec languages, probably Mod. Std. Mandarin).

Non-culminative accent in complex morphological systems: There is no reason that languages with complex morphology could not have boundary-marking systems operating within the word, i.e. multiple edge tones associated with different structurally-defined edges within the multi-morphemic word (possible examples: Mapudungun, Navajo?). This is very likely to be perceived as stress by speakers of stress/intonation languages, but need not involve any culminativity, nor any non-pitch phonetic correlates of prominence.

References

Cohn, Abby (1989). Stress in Indonesian and bracketing paradoxes. *Natural Language & Linguistic Theory* 7: 167-216.

Cohn, Abby (1993). The initial dactyl effect in Indonesian. *Linguistic Inquiry* 24: 372-381.

De Lacy, Paul (2002). The interaction of tone and stress in Optimality Theory. *Phonology* 19: 1-32.

De Lacy, Paul (2014). Evaluating evidence for stress systems. In van der Hulst, H. (Ed.). *Word stress: theoretical and typological issues* (Cambridge University Press), pp. 149-193.

DiCanio, Christian; Benn, Joshua; Castillo García, Rey (ms.). Prosodic focus and information structure in Yoloxóchitl Mixtec. Under review.

Dixon, R. M. W. (1980). *The Languages of Australia*. Cambridge University Press.

Echeverría, Max; Contreras, Heles (1965). Araucanian phonemics. *International Journal of American Linguistics* 31: 132-135.

Gordon, Matthew (2002). A factorial typology of quantity insensitive stress. *Natural Language and Linguistic Theory* 20: 491-552.

Gordon, Matthew (2004). A phonological and phonetic study of word-level stress in Chickasaw. *International Journal of American Linguistics* 70: 1-32.

Gordon, Matthew (2005). An autosegmental/metrical model of Chickasaw intonation. In Sun-Ah Jun (ed.), *Prosodic Transcription and Typology: a Unified Approach* (Oxford University Press), pp. 301-330.

Gordon, Matthew; Munro, Pamela (2007). A phonetic study of final vowel lengthening in Chickasaw. *International Journal of American Linguistics* 73: 293-330

Grice, Martine; Ladd, D. Robert; Arvaniti, Amalia (2000). On the place of phrase accents in intonational phonology. *Phonology* 17: 143-186.

Hermans, Ben (2011). The representation of stress. In M. Van Oostendorp, C. J. Ewen, E. Hume, & K. Rice (Eds.), *The Blackwell Companion to Phonology, Volume II*, (Oxford: Blackwell), pp. 980-1002.

Hyde, Brett (2002). A restrictive theory of metrical stress. *Phonology* 19:313-359.

Ladd, D. Robert (2008). *Intonational Phonology* (2nd edition). Cambridge University Press.

Maskikit-Essed, Rachael & Gussenhoven, Carlos (2016). No stress, no pitch accent, no prosodic focus: the case of Ambonese Malay. *Phonology* 33: 353-389.

McGarrity, Laura (2003). *Constraints on patterns of primary and secondary stress.* PhD thesis, Indiana University. Available online from the Rutgers Optimality Archive at http://roa.rutgers.edu/files/651-0404/651-MCGARRITY-0-0.PDF

McKendry, Inga (2013). *Tonal association, prominence, and prosodic structure in south-eastern Nochixtlán Mixtec.* PhD thesis, University of Edinburgh.

Molineaux, Benjamin (2014). *Synchronic and diachronic morphoprosody: Evidence from Mapudungun and Early English.* DPhil thesis, University of Oxford.

Molineaux, Benjamin (2016). Native and non-native perception of stress in Mapudungun: Assessing structural maintenance in the phonology of an endangered language. *Language and Speech* (published online, DOI 10.1177/0023830916628899)

Molineaux, Benjamin (ms.) The diachrony of Mapudungun stress assignment. Under review.

Morén-Duolljá, Bruce (2013). The prosody of Swedish underived nouns: No lexical tones required. *Nordlyd* 40.1: 196-248 (*available online*: http://septentrio.uit.no/index.php/nordlyd/article/viewFile/2506/2320)

Odé, Cecilia (1994). On the perception of prominence in Indonesian. In C. Odé & V.J. van Heuven (eds.), *Experimental studies of Indonesian prosody.* Department of Languages and Cultures of Southeast Asia and Oceania, University of Leiden.

Pankratz, Leo & Pike, Eunice (1967).Phonology and morphotonemics of Ayutla Mixtec. *International Journal of American Linguistics* 33: 287-299.

Remijsen, Bert & Gilley, Leoma (2008). Why are three-level vowel length systems rare? Insights from Dinka (Luanyjang dialect). *Journal of Phonetics* 36: 318-344.

Tabain, Marija; Fletcher, Janet; Butcher, Andrew (2014). Lexical stress in Pitjatjantjara. *Journal of Phonetics* 42: 52-66.

Zanten, E. van; Goedemans, R.W.N.; Pacilly, J.J. (2003). The status of word stress in Indonesian. In: V.J. van Heuven & J. van de Weijer (eds.) *The Phonological Spectrum, volume 2*. Amsterdam/Philadelphia: John Benjamins, pp. 151-175.