Fiat Lingua

Title: Tone for Conlangers: A Basic Introduction

Author: Aidan Aannestad

MS Date: 02-24-2018

FL Date: 04-01-2018

FL Number: FL-00004F-00

Citation: Aannestad, Aidan. 2018. "Tone for Conlangers: A Basic Introduction." FL-00004F-00, *Fiat Lingua*, <http://fiatlingua.org>. Web. 01 April 2018.

Copyright: © 2018 Aidan Aannestad. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License.

http://creativecommons.org/licenses/by-nc-nd/3.0/

Fiat Lingua is produced and maintained by the Language Creation Society (LCS). For more information about the LCS, visit http://www.conlang.org/

Tone for Conlangers: A Basic Introduction

Aidan Aannestad

1 Introduction

So you want to make a conlang with phonemic tone.[1] That's good! Phonemic tone is quite common among the worlds languages, and is possibly the most common suprasegmental system there is.[2] Certainly, phonemic tone occurs all over the world—not just East and Southeast Asia, but also most of Africa, multiple places in both North and South America, and even in Scandinavia. A conworld with no languages anywhere that have phonemic tone would be very unrealistic.

A lot of the time, when we think of a 'tone language',[3] we think of what I'll call the Sino-Tibetan/Tai/Hmong tone typology (STTH from here on out), especially as exemplified by Mandarin Chinese. In these languages, there is a list of distinct tone classes, and (almost) every syllable has a particular tone class assigned to it. Mandarin, for instance, has four 'tones'—high, falling, low-falling-rising (mostly just low), and rising; along with the occasional syllable that has so-called 'neutral tone'. There's very little tonal allophony in the standard conception of Mandarin—low-falling tones become rising before other low-falling tones, and neutral tones have some predictable value assigned, and that's about it.

Most languages with phonemic tone barely even resemble this.

Part of this is due to other rather unique typological characteristics of the STTH (plus Austroasiatic) linguistic area—most languages have a largely one-to-one syllable-to-morpheme correspondence, with almost all polysyllabic words being clearly multimorphemic (or at least having been such recently). It's not hard to just assign a tone pattern to each syllable, and leave it at that—each syllable is almost entirely independent of its neighbours anyway. Most languages elsewhere, as we well know, don't work like this—morphemes are very often more than or less than one syllable, and words can at times become very long indeed. In these cases, tone systems tend to be a good deal more complex than what Mandarin presents. Tones might end up bleeding over a large number of syllables, or appearing in several different places in a word as their context changes, or changing based on whether the word is inside a phrase or at an edge of the phrase, or themselves conveying meaning independent of segments—all kinds of complex phenomena occur. This can easily turn off the casual student of tone; but I've found that tone, properly presented and explained, truly rewards the effort put in to understand it, and opens a vast new range of possibilities for creativity in conlanging.

This article is intended as an introduction to tone for conlangers. I'm going to start with a description of basic concepts and processes in tone, then I'll go into a more general overview of variation among tone systems. After that, I'll talk about tone in a diachronic context, and at the end, I'll give some thoughts more specifically about creating conlangs with phonemic tones. I'm mostly going to ignore STTH-style tone systems here; I'll dedicate a small section to them under the overview of tone typologies.

[1] If you don't, you should.

[2] We don't know for sure, but it's suspected that many underdocumented languages have or had phonemic tone systems that the English-speaking documenter just didn't hear.

[3] A term I don't like all that much, since the construction '(feature) language' feels a bit exoticising. It sounds a lot sillier when you try to describe English as a 'fricative language'.

2 The basics of tone

Tone is, fundamentally, pitch variation used phonemically. This does not mean that *individual pitches* are themselves phonemic—no language requires someone to say a vowel with, say, 178 Hz exactly. Instead, phonemic tone systems rely on *relative* pitch heights. For example, the two Japanese words *àmé* 'rain' and *ámè* 'candy' are differentiated by which syllable has the higher pitch of the two—beyond that, it really doesn't matter. Because tone is phonemic, it functions as a small set of perceptual categories rather than any sort of continuum—much like vowel length is always a discrete 'long' vs. 'not long' distinction,[4] tone is also a very small number of discrete distinctions. Actual, individual pitches are more directly influenced by the speaker's vocal tract and by phrase-level intonational contours, which still very much exist in languages with phonemic tone. As long as nearby tones are still clearly distinguishable from each other, larger-scale things are free to vary.[5] Tonal pitch variation is also constrained to within speakers' normal spoken pitch range—no language requires a tone to be spoken as high as a speaker could theoretically sing, for example.

It's worth taking a moment to discuss how tone is fundamentally different from stress. Stress as it's typically conceived tends to consist of some combination of the following features.

- There's a maximum of one and usually a minimum of one (primary) stress per word
- Prosodic feet—larger than a syllable—are the main unit of stress assignment
- Stress and syllable weight often go together (weight attracts stress, stress encourages weight)
- The stressed syllable has some sort of phonetic 'prominence', typically beyond pitch changes and including things like non-phonemic extra length in stressed syllables and neutralisation of contrasts in unstressed syllables

Tone is different in all four of these areas.

- Depending on the language, words can have any number of (marked) tones, including zero; and multiple tones are usually grouped into phonemic 'melodies'
- Tones are assigned to moras or syllables
- Tone has no correlation whatsoever with syllable weight, except that a heavier syllable can provide more moras to attach more tones to
- Tones have no additional 'prominence' associated with them, though they can occasionally correlate with phonation changes

If you're not familiar with phonemic tone systems, you may inadvertently bring along some expectations about how they work from your familiarity with stress systems, so it's worth being careful. There has also traditionally been a category called 'pitch-accent' which ostensibly shows some characteristics of both tone and stress; but as this tends to be no different from a tone system that has a maximum of one marked tone (or melody) per word, I don't believe the term 'pitch-accent' is at all useful.[6]

2.1 Autosegmental phonology

Explaining tone properly hinges on the concept of *autosegmental phonology*.[7] Autosegmental phonology is the idea that phonological interactions don't all occur on the same conceptual 'level'—sometimes certain

[4] The very rare instances of actual three-way length distinctions are still just three discrete boxes instead of two.

[5] What counts as 'nearby' and 'clearly distinguishable' may vary from language to language.

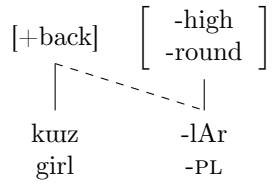
[6] If you'd like to read a fuller reasoning, Hyman 2009 is an excellent discussion on the topic. Scandinavian languages—often the poster children for 'pitch-accent' systems—are really best described as languages with both stress and tone simultaneously.

[7] The theory was originally published in Goldsmith (1976). Goldsmith (1990), an updated and much fuller description of the theory, is highly recommended as further reading for anyone who's interested in a deeper look at the theoretical mechanics of autosegmental phonology.

features may interact with similar features a good way away in the word, and ignore all of the intervening material. The name refers to the idea that some features behave as their own independent 'segments' on a tier separate from the base string of vowels and consonants. Vowel harmony is a good example of this—vowels harmonise only with other vowels, and none of the consonants in between affect the process at all. Effectively, there's another 'tier' above the basic string of segments where features like rounding and backness interact directly with each other.

Here's an example from Turkish of how autosegmental phonology describes vowel harmony.

(1)



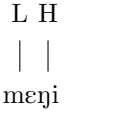
kɯzlar 'girls'

In this example, the solid line shows where in the string of segments the feature originally attaches, and the dashed line shows where it spreads. In this case, the [+back] feature on the vowel /ɯ/ in *kɯz* spreads to the suffix *-lAr*,[8] causing the allomorph *-lar* to appear. The intervening consonants are simply ignored, as they don't have any representation on the tier the vowel features exist on.

There are several phenomena that autosegmental phonology explains much more clearly than a simple single-line conception of phonology would, but tone is the place where it really shines. The core concept of autosegmental tone is that each morpheme can have a tone melody associated with it, and the individual tones within the melody then associate themselves with syllables. As an example from my own conlang Emihtazuu, the word *mɛŋí* 'this thing' has in effect two parts—the segmental sequence /mɛŋi/ and the tone melody [LH] (where L is 'low tone' and H is 'high tone'). Other words have other tone melodies—*pára* 'fast' is [HL], for example. Some words have melodies with fewer tones than the number of syllables—*kémí* 'watch, listen to' has a tone melody of simply [H], which is realised across the entire morpheme. Other words, such as *juna* 'cloud', don't have any tone melody associated with them, and their tone is entirely predictable from their environment.

This example shows a basic Emihtazuu word, with a tier for segments and a tier for tones.

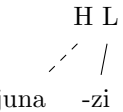
(2)



mɛŋí 'this object'

Sometimes a syllable can have more than one single tone associated with it, and there are two strategies for resolving that situation. The first is to simply move one of the tones to a neighbouring syllable that has no tone. When *juna* above is followed by the ablative marker *-zî*, which has an [HL] tone melody, the end result is *junázi* 'from the cloud'—the H tone gets moved onto the previous syllable, so that every tone is now associated with one syllable. There's the proper segmental sequence /junazi/, and you have an [HL] melody somewhere; so there's no problem—a morpheme's tone melody doesn't need to be coterminous with the segmental part of the morpheme. This is shown by the following diagram.[9]

(3)



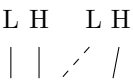
junázi 'from the cloud'

[8]The A here represents a low unrounded vowel that's unspecified for backness.

[9]The initial syllable is automatically L because unmarked syllables default to L in Emihtazuu.

Sometimes, though, you don't have a free syllable to offload an extra tone onto, and often these situations are where you get contour tones (like 'rising' and 'falling'). A contour tone in most autosegmental analyses is just two level tones on the same syllable. In Emihtazuu, if you want to combine *mɛŋí* and *-mǐ* 'inside', you get four total tones ([LHLH]) and only three syllables to assign them to. In this case, the result is that the middle syllable gets both the H tone from the first morpheme and the L tone from the second morpheme in order—so you get *mɛŋîmí* [L-HL-H], which is shown by the diagram below. Again, it preserves the segments and the tone sequence; you just have to squeeze the tone sequence a bit to get it to fit.

(4)

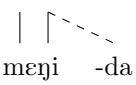


mɛŋîmí 'inside this object'

In some cases, there simply isn't enough space to fit all the tones a word is assigned—for example, if your language doesn't allow two tones on a syllable, what do you do when a monosyllabic word ends up with two tones? Sometimes, this results in tones getting pushed off the end of a word, with the surface only rendering part of the tone melody. Other times, the tone melody is simplified: an underlying [HLH] melody on a monosyllable might be realised as just a high tone, or a high-mid sequence, or something similar. Still other times, a language can realise a tone melody in a way that is entirely unexpected. Emihtazuu, for example, has a prohibition on contour tones on short word-final vowels, and thus contour tones on short monosyllables don't 'fit' properly. With a word like *nǐ* 'eat', the initial low tone is pushed off and you just get a high tone [ní]; with *nî* 'be not there', the whole word is given an unexpected realisation with *superlow* tone plus creaky voice and appears as [ǹ̰].

When you have morphemes that are unspecified for tone, there are a couple of strategies that can operate to assign them a tone. The first choice tends to be expanding contours and complex melodies onto otherwise unspecified syllables. If you still end up with unspecified syllables, one thing that can happen is so-called 'tone spreading'. Tone spreading is a phenomenon where a tone at the edge of a melody will spread in one direction or the other until it reaches some defined stopping point.[10] For example, in Emihtazuu, high tones spread rightwards across unmarked syllables—so *mɛŋí* plus *-da* 'genitive', which is unspecified for tone, becomes *mɛŋídá*, as shown by the diagram below. In standard Japanese, a high tone will spread leftward instead, and only up to the second syllable—so *hoɕi-kúzu* 'stardust' ([xx-HL] or [xx-Hx] underlyingly, depending on your interpretation) becomes *hòɕíkúzù* [LHHL]. This in some sense still preserves tone melodies, if you consider the melodies as defined by transitions between level tones rather than by the level tones themselves.

(5)



mɛŋídá 'for/of this object'

The other strategy is by having contextual tone assignment that doesn't necessarily reference other tones. In standard Japanese, an otherwise unmarked word is realised with the first syllable low and the rest high to the end.[11] In Emihtazuu, an unmarked first syllable becomes *mid* as long as it's more than one syllable away from a marked tone, and every other unmarked syllable becomes low (if a high tone hasn't spread to it). Cross-linguistically, unmarked tones tend to default to low, but this isn't always the case—some Athabaskan languages have high tone as the unmarked default.

[10] Which may or may not be another marked tone—in some languages, tone spreading pushes other marked tones along in front.

[11] This extends all the way to the end of an attached clitic, and some words are distinguishable only by whether or not the tone spreads to the clitic. The noun in *hàɕí=gá* 'edge-SUBJ' is unmarked, while the noun in *hàɕí=gà* 'bridge-SUBJ' has an [LH] melody.

3 Tone systems

There's quite a lot of variation in tone systems across the world, but just like in vowel systems, a good deal of that can be boiled down to some specific key factors. This section should give you a basic idea of what factors can vary and how.

3.1 Number of level tones

Languages can differ in exactly how many level tones can be used to form tone melodies. The minimum is generally two, though in some cases you might be able to argue for languages that have one marked tone while everything else is underlyingly unmarked; even in these cases, though, you still end up with a surface contrast between two tone levels. With two, you'll have high and low, with three you'll have high, mid and low. Two-tone systems are the most common cross-linguistically, with three-tone systems making up most of the rest. Some Oto-Manguean languages, though, are reported to have as many as five phonemic levels, from 'superlow' to 'superhigh'! Some, possibly most, languages allow some morphemes to have no marked tone; some languages assign a tone melody to everything.

Your language may also vary on exactly how far apart musically the tones typically are realised, though this is likely to be different in different social contexts rather than being a whole-language setting. Formal loudspeaker or radio announcements in Thai apparently exaggerate the tone contrasts dramatically;[12] in contrast, a friend of mine who does fieldwork on Mixtec regularly complains that her variety leaves less than a half-step's worth of distance between its three phonemic levels—making the differences very hard to hear for non-native speakers!

3.2 Set of allowed melodies

Languages can also vary by what melodies they allow morphemes to carry. Some languages only allow a maximum of two tones per melody (thus leaving [H], [L], [LH] and [HL] as the four possible melodies); other languages will allow three-tone melodies. Languages don't necessarily allow the theoretical full set of melodies these constraints would allow, and languages with three-tone melodies are especially likely to have some melodies disallowed—for example, a language might allow [HLH] but not [HMH], and Emihtazuu seems to have no marked [L] melody. Whether or not your language allows same-tone sequences in melodies (e.g. [HHL]) may depend on how it deals with the so-called 'Obligatory Contour Principle' described below; but keep in mind that a language can still have underlying same-tone sequences without necessarily having overt same-tone sequences anywhere.

3.3 Tone-bearing units and tone-stress combinations

We've been referencing tones on syllables so far, but it's much more accurate to talk about 'tone-bearing units', or TBUs. Some languages have syllables as the only TBU, but most languages have the mora as the basic TBU. Not all moras can necessarily bear tone; in Emihtazuu, for instance, only vocalic moras can bear tone—coda /h/ and nasals can't bear a tone. In most Bantu languages, nasals can bear tone just fine. Also in Emihtazuu, when two tones are squeezed onto a one-mora syllable, the short vowel will become long to make sure that each tone has a mora to attach to.

Some languages combine tone and stress, with varying results. You can theoretically have a language with both tone and stress, completely independent of each other; I might rather expect a system where either stress or tone restricts the distribution of the other. For example, in Norwegian and Swedish, only the stressed syllable contrasts for tone. This creates a contrast between /ˈ̀tːɑlə/ 'pronounce' and /ˈ̂tːɑlə/ 'pronunciation'—tone on all the other syllables is predictable, and the only contrast on the stressed syllable is low vs. falling.[13]

[12] According to Corley and Watkins (2015).

[13] This is the tone pattern in East Norwegian dialects. West Norwegian dialects invert the tones and instead have a contrast between high and rising!

3.4 Sino-Tibetan/Tai/Hmong tone typology

As I mentioned before, STTH languages tend to have a very unusual approach to phonemic tone. Some researchers have argued that these systems can still be described autosegmentally, but whether or not this is the case, it doesn't often seem to be the most helpful way to think about STTH-style tone systems. There are some prerequisites to having a STTH-style tone system; these mainly include having a strong tendency for every morpheme to be one syllable (or less), and having very minimal phonological interactions across morpheme boundaries. This tends to result in a preponderance of heavy syllables with tone contours that mostly don't interact with each other. In these cases, tone systems are better described by the set of allowed contours,[14] and which syllables the contours are valid on—Cantonese, for example, allows a much restricted set of tone contours on stop-final syllables. These systems often have certain tone contours coöccur with voicing qualities—Vietnamese has a low falling tone with breathy voice and several tones with creaky-voice-like phonation.[15] Some syllables can be unmarked for tone, such as Mandarin's suffixes *-mən* 'PL' and *-lə* 'PERF', but these tend to be found only in the language's very small set of bound morphemes.

These kinds of languages seem to be entirely restricted to East and Southeast Asia, as part of a single linguistic area that gained phonemic tone together over a period from about 500 AD to 1500 AD.[16] I've personally never heard of anything similar elsewhere in the world. Keep that in mind when determining what's 'normal' crosslinguistically—if you'd like your conworld to have Earth-like frequencies of various typologies, you probably shouldn't make all languages with phonemic tone work like Southeast Asian languages.

4 Types of tonal allophony

There is quite a range of allophonic changes that can happen to tones. In this section, I'll go over the main types of tonal allophony, and give you some descriptions of how they typically work.

4.1 Moving vs. Contouring

We saw above how tones in Emihtazuu can be moved or compressed to accommodate mismatches between the number of syllables and number of marked tones. Languages can vary on which of these strategies they prefer, and how exactly they work out. In the Emihtazuu example, the following hierarchy of repair strategies operates in these kinds of situations:

- If there's an unmarked syllable to the left, expand the melody onto it
- If there's none to the left but one to the right, expand the melody onto that one
- If both sides are marked or word boundaries, attach the leftmost tone in the melody to the syllable to the left, and make it a contour
- If the syllable to the left is already a contour (or a word boundary), make the current syllable a contour
- If the syllable to the left is a contour or word boundary, and the current syllable can't have a contour (or is already a contour due to tones coming from the right), shift the whole series of melodies leftwards until the above rules can make things work out

This sequence is just an example. Your language might disallow contouring entirely, or it might never have any unmarked syllables for melodies to expand to. You might allow three tones to attach to one syllable

[14] Which is probably the best area to try to apply autosegmental analyses to—asking the question of which underlying tone sequences result in the observed set of allowed contours.

[15] This can actually be realised as a glottal stop partway through the vowel in some speakers' speech. This is one piece of evidence in some scholars' argument that glottal stops can actually behave suprasegmentally at times, which is neat.

[16] Sino-Tibetan, Tai-Kadai, and Hmong-Mien are the main participants in this area, and have no languages left without tone. Austroasiatic languages besides Vietnamese and its neighbours have mostly remained without tone, though they share a good deal of the region's other common features.

under certain circumstances, which might then surface with a simplified or entirely unexpected realisation. You could end up pushing tones off the edge of words, where they might simply vanish, or they might appear instead on the next word! Overall, tone seems to strongly prefer shifting leftwards when it can, but your language might move it rightwards more often instead. There is quite a lot of variety possible in tone mismatch repair strategies.

4.2 Avoidance of same-tone sequences

There's a phenomenon in tone studies called the 'Obligatory Contour Principle', which states that two identical marked tones should not occur side-by-side.[17] Like most linguistic universals, this is quite violable; but it seems to be a very common phenomenon. Here's an example from kiRundi (taken from van Oostendorp 2005).

(6) a. nà-rá-zì-b**á**rììrà
1SG.SUBJ-PAST-CL10.PL.OBJ-sew
'I was sewing them'

b. nà-rá-b**à**rììrà
1SG.SUBJ-PAST-sew
'I was sewing'

In this case, the verb root for 'sew', *bárììrà*, has a high tone that attaches to the first syllable. When that follows a low tone, as with the class-10 plural object agreement marker *zì-*, this is fine; but when this high tone follows a high tone (as in(6b)), a rule takes effect that flips the second high tone to a low tone, to avoid a sequence of two marked high tones.[18]

Despite the name, this is almost certainly not an inviolable constraint, and is just another phenomenon that seems to happen frequently cross-linguistically. (Emihtazuu almost ignores it entirely.) Again, use it as you will, and be creative in deciding how it works out in practice.

4.3 Tone spreading

Once again, we've seen above how melody-final high tones in Emihtazuu spread rightwards until they reach another marked tone. This is another kind of common tonal allophony, where a marked tone will expand beyond a single syllable in one direction or another. This may not affect every tone level in your language, and may be limited to various domains (e.g. within a word; within a phrase; within just a morpheme)—again, see the above example from standard Japanese, where high tone spreads leftwards, but only to the second syllable of a word (never the first). It may even override or push other marked tones. Tone spreading follows the crosslinguistic tendency for tones to prefer moving leftwards; Emihtazuu's just happens to move rightwards.

4.4 Tone shifting

In some languages, a morpheme's tone melody may attach not to the morpheme directly, but to the whole word. Here's an example from chiChewa.[19]

(7) a. -gula
-buy
'buy' (bare stem)

[17] Interestingly, it's been applied to explain non-tonal phenomena as well; see van Oostendorp (2005) for an interesting discussion.

[18] I assume that the low tones in the verb root are all underlyingly unmarked, as they seem to coexist just fine with each other and the flipped former high tone; or else the OCP in kiRundi only operates on high tones. The source doesn't give any explanation about this.

[19] Hat tip to Zev Brook for the example.

b. ku-sa-g**ú**la
INF-NEG-buy
'to not buy'

c. ku-sa-gul-**í**tsa
INF-NEG-buy-CAUS
'to not sell'

d. ku-sa-gul-its-**í**ra
INF-NEG-buy-CAUS-APPL
'to not sell something to/for/at'

In this case, the [HL] melody associated with *-sa-* 'NEG' attaches to the right edge of the word, regardless of wherever its source morpheme is.

4.5 Floating tones

In some languages, there are morphemes that have no segmental component at all—all they have is a tone melody. This phenomenon has led some researchers (mostly in the past) to describe languages as having 'grammatical tone'; this description mostly just means that the language has some morphemes that are nothing but a tone melody. These so-called 'floating tones' can often appear only as complications to existing tone allophony rules, as well. This is adapted from Paster (2003), a neat study on tones in Gã:[20]

(8) a. o-tʃa
2SG-dig
'You dig'

b. ó-tʃa
2SG[PERF]-dig
'You have dug'

c. e-tʃa
3SG-dig
'He digs'

d. é-tʃa
3SG[PERF]-dig
'He has dug'

You can see here how the perfect aspect is marked not by any segmental unit, but by a high tone that attaches to the subject prefix. You might think 'well, that's not a lot of perceptual difference, is it!', but keep in mind that your ear isn't tuned to it—other people struggle with English's stop voicing distinction. If you'd grown up speaking Gã, the difference would be just as clear as anything else.

4.6 Polar tones

We've talked about the distinction between marked tones and unmarked context-dependent tones, but some languages have something that seems to fall in between those categories. In these languages, some morphemes aren't themselves marked for a specific tone, but are marked to have the *opposite* tone from whatever they're next to. Hausa's copula *tʃeː/neː*, for example,[21] is high tone following a word ending in a low tone, and low tone following a word ending in a high tone. It has what's called a 'polar tone', so named for the flipped tonal 'polarity' it displays. It itself has no specific tone attached to it, but all the same, it's not entirely unmarked for tone—it's just marked as 'not whatever the last tone was'.

[20] Which is a really interesting paper, and comes highly recommended as a case study for further reading.
[21] Thanks again to Zev Brook for this example.

4.7 Downstep

Some languages exhibit a phenomenon where high tones become progressively lower in pitch over the course of a word or phrase—possibly even becoming indistinguishable from low tones near the end of a long sequence. (This is sometimes also called 'tone terracing'.) This isn't restricted to sequential high tones or high tones created through spreading; it can affect for example the second high tone in [HLH] sequences as well. Zerbian and Kügler (2015) give an example from seTswana, where high tones become lower after each word break. In this example, the high tone on *baná* is highest, *bátsʰábá* is level but somewhat lower, and the high tone on *maqʰóa* is slightly lower still.[22]

(9) ba-ná bá-tsʰábá ma-qʰóa
CLASS2.PL-child CLASS2.PL-fear CLASS6.PL-white.person
'The children fear the white people.'

Some languages have phonological processes that cause some high tones to downstep while others remain the same; see Paster (2003)—the paper on Gã from above—for a good description of a system where this happens.

4.8 Interactions with segments

As much as tones are best conceived as being on a separate phonological tier from segmental material, sometimes tones do interact with segments. Segmental features on consonants, typically phonation features like voicing or glottalisation, may restrict which tones can appear on a syllable ending with or beginning with such a segment. Segments may also count as boundaries for tone movement and spreading rules—in Emihtazuu, the few coda consonants seem to block some kinds of tone movement but not others (I haven't worked out the details yet). Conversely, tones may affect the segmental material—see the above section on STTH languages for how some tones can be accompanied by vowel phonation changes. Standard Japanese seems to have a complex and somewhat circular relationship between tone placement and vowel devoicing rules, where sometimes a high tone will block devoicing, and sometimes devoicing will move a high tone elsewhere. Glottal consonants can affect tone realisations sometimes; in some varieties of Mixtec, CVhCV words show a different surface tone pattern than CVCV words with the same underlying tone melody. It seems that overall, tonal realisations are most connected to phonation features rather than other things like place (excepting with glottal consonants), and tone movement rules are most likely to be affected by segmental timing concerns (such as the extra mora from a coda consonant).

4.9 Tone and allomorphs

Just as morphemes can have allomorphs that differ segmentally from each other, they can also have allomorphs that differ in tone. Sometimes this is a tone difference only—one allomorph has one tone melody; the other has a different one. Sometimes this is both segmental and tonal. Emihtazuu, for example, has two allomorphs for many verb stems—one which results from the loss of word-final consonants (which became tone), and one where the consonant wasn't word-final. Thus, the word *kémí* 'watch, listen to' has an allomorph *kémebi-* before certain suffixes. When the root-final consonant resurfaces, the tone pattern its loss created is removed.

5 Diachronic tone

Since any truly realistic conlang must take diachronics into account, it's important to have a sense of how tone works over time. Most of the principles of segmental diachronic change carry over into tonal change—such as allophonic variation becoming phonemic—but their application may or may not be straightforward.

[22] The source has a Praat diagram of the pitch height across the phrase, which is quite neat to look at.

As an overall note, tone seems to be strongly correlated with linguistic areas, often more strongly than with genetic language families. It seems that languages without tone contrasts will more often than not gain them if they're in contact with many languages that already have them, and languages with tone contrasts will more often than not lose them if they're in contact with many languages that don't have them. Phonemic tone can very well be a feature that characterises a family, but it's possibly more likely to be a feature that characterises an area.

5.1 Tonogenesis

Some languages seem to have had phonemic tone for as long as we can tell. Linguists working on Niger-Congo and Oto-Manguean, for example, have reconstructed tone contrasts all the way back to the proto-language. In some cases, though, we can see that a language may not have always had phonemic tone. This is most obvious when one language's tones correspond to another language's segments or segmental contrasts—the one with tones has turned segments into tones (since when tone goes away, it mostly leaves a suprasegmental trace, if anything at all). This seems to have happened in Southeast Asia, for example—many of the writing systems there don't account for tone, or account for a much reduced tonal system than the spoken language presents. Instead, they show a lot of extra segments or segmental contrasts that can be predictably turned into the spoken language's tones. In Athabaskan, one primary branch shows exactly the opposite tone patterns from another branch, which Kingston (2005) argues is due to different reflexes of a former consonant glottalisation contrast.[23]

Phonemic oppositions in general tend to produce oppositions in tone, i.e. where a two-way contrast in segments turns into a high-low contrast rather than a high-unmarked contrast. When the simple presence of something, e.g. a final segment, causes a tone contrast, it may or may not cause a potential 'opposite' situation to become marked for the opposite tone. Three-or-more-way contrasts will likely have one particular pair within them change to a tonal distinction, while retaining the rest of the segmental oppositions.

There are a number of possible sources for tonal contrasts. A very common one is phonation contrasts on an adjacent consonant—for example, a contrast between *ba* and *pa* can turn into a contrast between *pà* and *pá*. It tends to be the case that voicing produces lower tones and aspiration produces higher tones. Modern standard Korean is undergoing this kind of change right now, where an aspiration contrast in initial syllables is turning into a tone contrast instead—both what's written <pa> and what's written <pʰa> are pronounced as [pʰa], but the first produces an LH melody at the beginning of the word, and the second produces an HL melody. <kɛta> 'fold' is pronounced [kʰèdá], while <kʰɛta> 'dig' is pronounced [kʰédà].[24] As mentioned above, glottalisation in Proto-Athabaskan resulted in high tones in one branch and low tones in another branch—so it might not hurt to be aware of the specific phonetic details of whatever contrast you're trying to turn into tones.

Another common source is final consonant loss. A final voiceless stop may be lost to become a high tone; a final voiced stop may be lost to produce a low tone. A word without an original final consonant might become mid or low, or it might just become unmarked. Again, the 'more voiced' things tend to become lower tones, and the 'more aspirated' things tend to become higher tones. This can apply to all kinds of lost final consonants, not just stops—nasals might also produce low tones, and voiceless fricatives might produce high tones. In Emihtazuu, the loss of any continuant (nasals and fricatives) produces an HL melody at the end of the word; the loss of a stop results in an LH melody at the end of the word.

These two phenomena—loss of phonation contrasts on consonants and loss of final consonants—seem to account for the majority of tonogenesis cases. There's a whole range of other things that could quite conceivably create tone contrasts—vowel phonation, stress contrasts, loss of other syllables,[25] and so on—and if you'd like to turn one of these into tone, feel free to be creative! Just do the research on phonetic side effects of the articulations you'd like to turn into tone and see if pitch changes are or could be one of them.

[23] This is seemingly confirmed by the third (much smaller) branch of Athabaskan, which retains a corresponding glottalisation contrast.

[24] Initial /h/ and /sʰ/ also trigger an HL melody, while basically anything else—nasals, /l/ and the so-called 'tense' stops—triggers the LH melody. I'm not sure what /s/ does, as it's also a fricative, but otherwise tends to pattern with the 'tense' stops (in opposition to /sʰ/).

[25] Middle Korean tones are theorised to come from this. (Modern Korean had lost those tones before it started regaining them again.)

5.2 Tone over time

Tone diachrony seems to be very under-studied, and in doing research for both my own conlanging and for this article, I have found effectively no crosslinguistic studies of tone-only changes. Still, it's possible to infer some things based on how synchronic allophony in general results in diachronic change. For example, tone melodies that weren't valid before might become valid through multi-morphemic words becoming reanalysed as monomorphemic—if you have a compound word with the first part [LH] and the second part [L], you might soon find yourself with a single morpheme with an [LHL] melody. Don't forget analogical changes, as well—a marked tone that moves to a particular spot frequently may find itself attached there in the future, and if many words in a class share a particular tone melody, the melody might become a marker of that class.

There are some other tone-only changes attested in some languages; the few instances I've come across all involve tone melodies shifting one step left or right. For example, in both kiKuyu and (urban) Kansai Japanese, tone melodies have shifted one step leftwards. In both cases, this results in the odd situation of tone melodies attaching one syllable *before* the start of the word—in Kansai Japanese, for example, a word beginning with [HL] starts with a low tone, but a (toneless) determiner placed before that word has high tone: *hà꜀ì* 'chopsticks', but *kónó hà꜀ì* 'these chopsticks'. In some varieties around Japan, melodies have instead shifted rightwards one step, but interestingly, the shift is blocked if it would result in moving a high tone onto a high vowel—even though high tones already on high vowels don't seem to be dispreferred in any way.[26]

The same processes that cause tonogenesis can add tone contrasts to existing tone systems. In some cases, the original stimulus for a change may not go away, and you can end up with segment-based restrictions on tone patterns—you might, for example, find a language where all words that begin with a voiced stop also all begin with a low tone. In other cases, you might find that all [HL] words beginning with a voiced stop now have [LHL] melodies instead. If you have a tone contrast and a vowel nasalisation contrast, the nasalisation contrast might be lost to create even more tone contrasts.

Just as phonation contrasts can turn into tonal contrasts, tone contrasts can turn into phonation contrasts. An example is Danish stød, a kind of glottalisation / creaky voice phenomenon, which Riad (2000) argues is the modern reflex of an [HL] tone melody squeezed into a single syllable. It seems impressionistically that higher tones (or tones that *end* high) are more likely to develop a breathy voice component, and lower tones (or tones that end low) are more likely to develop a creaky voice component; I'm not sure if this guess is accurate. Tone contours as well seem more likely to develop into something like creaky voice.

Tone can also be lost over time fairly easily, often again due to areal effects. Lost tone contrasts may leave behind non-tonal contrasts, as in the case of Danish stød, or they may just vanish without a trace. If a language with tones is being learned as a second language by speakers of languages without tones, tone is likely to be lost entirely—see, for example, Hausa, where the home language version has tones and the trade language version has no evidence there ever were any.

6 Tone for conlangers

Conlangers specifically have some challenges when it comes to tone systems. Possibly the biggest is the fact that most conlangers aren't native speakers of languages with phonemic tone, and thus most conlangers haven't been trained to hear tone contrasts clearly. The easiest way to remedy this is by learning a natlang with tone. I discovered while learning Japanese that after hearing enough of it, I was automatically making the tone contrasts—even before I learned that Japanese had phonemic tone contrasts! Listen to as much audio in the language you're learning as you can, and try mimicking exactly the phrases that you hear.[27] Eventually, you may well find yourself having no problem at all with tones.

[26] I suspect the reason high vowels specifically are avoided is because they have slightly lower sonority than lower vowels, and thus the mismatch between a 'higher-prominence' high tone and a 'lower-prominence' lower-sonority vowel is enough to block the shift.

[27] Alternatively, spend a month or two specifically avoiding trying to reproduce what you hear—this helps your brain listen to the target language's contrasts without framing them in terms of your own language's phonemic categories.

The second problem comes with Roman-script orthography; and while it's a problem for natlang orthographies as well, the fact that most conlangs exist primarily in a written form using Roman letters makes it all the more important for conlangers. The question is this: how do you write a language morphophonemically when tones exist separately from the segments you have letters for? The general consensus seems to be that tones should be written however their attachment ultimately works out, regardless of how morphophonemic the rest of your spelling is. This means, though, that you might have to do some extra documentation work to describe underlying forms: if your citation form for verbs can't distinguish between [HL] and [HLH] melodies, you'll have to either state explicitly which melody a given verb root has, or give examples of inflected forms where you can tell the difference.[28]

You could, of course, create a conscript to get around this. If you have a truly morphophonemic conscript, you might find it easier to keep your master documentation in the script rather than with Roman letters! Keep in mind, though, that most natscripts that deal with tone tend to be more phonetic than strictly necessary—most speakers aren't aware of the autosegmental analysis of their language's tone system, and will just write whatever they hear on the surface. They might start to notice that different surface realisations sound 'the same' to them, and may look deeper to try and create a more morphophonemic system, but overall, a script is likely to show a more surface-based description of tone contrasts rather than a more phonemic one. This may not be the case if the script was developed for an earlier state of the language that had no tone contrasts—with this kind of script, you might inadvertently end up with a very morphophonemic spelling system, if the segmental source for tone contrasts turns into tone predictably enough.

Phonemic tone is a wonderful feature to play around with in conlangs, and it's a shame that more conlangers don't use phonemic tone systems. I hope this article has helped familiarise you with the basic concepts that underly tone systems, and has given you enough building blocks to start creating your own conlang tone systems. Don't be scared of tone—it's no more complicated than anything else in linguistics!

References

Corley, George and Britton Watkins (2015). *Thai (natlang)*. URL: http://media.blubrry.com/conlangery/p/content.blubrry.com/conlangery/Conlangery114.mp3.

Goldsmith, John A. (1976). "Autosegmental Phonology". PhD thesis. Massachusets Institute of Technology.

— (1990). *Autosegmental and Metrical Phonology*. Blackwell. ISBN: 978-0631136767.

Hyman, Larry M. (2009). "How (not) to do phonological typology: the case of pitch-accent". In: *Language Sciences* 31.2-3, pp. 213–238. DOI: 10.1016/j.langsci.2008.12.007.

Kingston, John (2005). "The Phonetics of Athabaskan Tonogenesis". In: *Athabaskan Prosody*. John Benjamins Publishing Company, pp. 137–184. DOI: 10.1075/cilt.269.09kin.

Paster, Mary (2003). "Floating Tones in Gã?" In: *Studies in African Linguistics* 32.1. URL: http://journals.linguisticsociety.org/elanguage/sal/article/download/1366/1366-2023-1-PB.pdf.

Riad, Tomas (2000). "The origin of Danish stød". In: *Analogy, Levelling and Markedness : Principles of change in phonology and morphology*. Berlin/New York, Mouton de Gruyter, pp. 261–300.

Van Oostendorp, Marc (2005). "Obligatory Contour Principle". URL: https://www.researchgate.net/profile/Marc_Oostendorp/publication/265119883_Obligatory_Contour_Principle/.

Zerbian, Sabine and Frank Kügler (2015). "Downstep in Tswana (Southern Bantu)". In: *Proceedings of 18th ICPhS*. URL: https://www.researchgate.net/profile/Frank_Kuegler/publication/280975887_Downstep_in_Tswana_Southern_Bantu/links/55ceeacc08ae6a881384a22e/Downstep-in-Tswana-Southern-Bantu.pdf.

[28] Which is actually a neat kind of way to use the Latin/Greek 'principle parts' system—you have to have several principle parts to see how the tones work out!