

On Establishing Underlying Tonal Contrast

Keith Snider
SIL International

Phonological field work is largely about establishing contrast in comparable environments. The notion of phonological contrast, however, can be confusing, particularly in its application to tone analysis. Does it mean phonemic contrast in the structuralist sense, or does it mean underlying contrast in the generative sense? Many linguists, in publications otherwise written from a generative perspective, support underlying tonal contrasts with minimal pairs and other data that are based on structuralist criteria. This paper critiques how tonal contrast is often supported in the literature and demonstrates that many supposed minimal pairs are invalid from a generative perspective. It further demonstrates that because many morphemes in tone languages consist solely of floating tones, the potential for these cannot be ignored when establishing comparable phonological environments.

1. INTRODUCTION¹. The notion of phonological contrast can be confusing, particularly in its application to tone analysis. While most phonologists today assume a generative theory of some sort that derives surface forms from underlying forms, when it comes to tone analysis, they often establish and support their underlying contrasts with phonemic forms that are based on structuralist notions. Establishing phonemic tone contrasts is neither a necessary first step to establishing underlying contrasts nor even a particularly helpful one, and it can lead to wrong conclusions.

Snider 2013 describes the object of tone analysis as twofold: a) to discover the different underlying tone patterns that are potentially possible for each category of morphemes (e.g., verb roots, noun roots, subject markers, etc.), and b) to explain the different surface realizations of the underlying patterns. In order to achieve these goals, Hyman (this volume) identifies three different tasks, referred to as stages, that one must accomplish in order to conduct tone analysis from scratch. Stage I: determine the surface tonal contrasts by comparing words in isolation; Stage II: identify any tonal alternations by eliciting paradigms and/or short phrases; and Stage III: analyze the data obtained in the first two stages, drawing on theoretical constructs and formal devices such as autosegmental theory in order to express one's insights.

The present work assumes the goals of Snider 2013, but mainly limits itself to helping to identify the surface tonal contrasts of words compared in isolation in a manner that lends itself to later establishing the underlying patterns. Put another way, this is how Snider

¹ I wish to thank the following people, listed in alphabetical order, for their written comments on previous versions of this paper: John Alderete, Rod Casali, Kathleen Hall, Larry Hyman, Will Leben, David Odden, and three anonymous reviewers. These comments fueled many changes. I further wish to thank Phil Davison and Rebecca Ouwehand for their excellent and extensive editorial suggestions. The paper is much easier to read as a result of their help. This work also benefitted from the discussion that followed its presentation to the Vancouver Phonology Group, at Simon Fraser University, March 19, 2013. I take full responsibility for any remaining shortcomings.

carries out Hyman's stage I.² Beginning tone analysts will find the goal of identifying underlying forms more easily accomplished if, when they compare words in isolation, they take care to ensure that the data upon which they are basing their contrasts are as comparable as possible in ways that can affect the outcome. It is hoped that this paper will help researchers better understand how to evaluate the phonological comparability of the data upon which they are basing tonal contrasts.

The "gold standard" for establishing tonal contrast has long been the minimal pair (cf. Mazaudon (this volume) and Coupe (this volume)). If general usage in the tonal literature is anything to go by, the terms "minimal pairs/triplets, etc." refer to words or utterances that: a) differ in meaning, b) differ in tones or tone patterns, and c) are segmentally identical. And what could be more convincing than a set of forms like those in (1)?³

(1) Chiquihuitlan Mazatec [maq]⁴ (Jamieson 1977)

[-]	[-]	[-]	[-]
čha	čha	čha	čha
‘I talk’	‘difficult’	‘his hand’	‘he talks’

In this quadruplet, the meaning of each word is different, the tone on each word is unique, and the segments of each word are identical. Moreover, each word is pronounced in isolation. Ergo, tonal contrast.

Some linguists even go so far as to say (at least off the record) that tonal contrast is not proven *unless* there are minimal pairs. For example, the editor of a reputable journal recently criticized a paper, written by one of my colleagues, on the grounds that my colleague hadn't adequately established that the language was tonal because he hadn't provided examples of minimal pairs. Most linguists, however, acknowledge that segments do not actually need to be identical in order to prove tonal contrast, provided that any segmental differences do not influence the surface realizations of the tones in any way. Accordingly, examples like those in (2) are also common in the literature.

(2) Buli [bwu] (Akanlig-Pare & Kenstowicz 2003)

	H		M		L	
CV	lé	‘spinster’	lō	‘fall’	mà	‘mother’
CV:	mí má:	‘I helped’	mā:	‘help!’	wà mà: mǐ	‘he helped me’

² One minor difference with Hyman's work is that whereas Hyman focusses on identifying underlying tones, the present work focusses on identifying underlying tone patterns. The term "tone pattern" in this article refers to a complete tonal pattern associated with the morpheme or word in focus. An underlying pattern may be something as simple as /H/, /L/, or /Ø/ (toneless), or more complex like /HL/ or /HØ/.

³ Throughout this article, pitch is often represented using the following graphic notation: level pitches are represented with level bars that correspond to the pitch heights they represent (e.g., LH [- ¯]), and contour pitches are represented with angled bars that slope from higher to lower in the case of falling pitches (e.g., falling [\]), or from lower to higher in the case of rising pitches (e.g., rising [/]).

⁴ At the first mention of each language name, the appropriate ISO 639-3 code for the language is provided.

CVC	zúk	‘head’	bāŋ	‘bangle’	bàŋ	‘lizard’
CV:C	bí:k	‘child’	bū:k ^w	‘goat’	nà:b	‘chief’

Although the segments of these contrast sets are not identical (with the exception of ‘bangle’ and ‘lizard’), to the extent that they differ, they differ in a manner that is presumed not to affect the surface realization of the tones. Assuming that all other things are equal, these data demonstrate that the tonal differences are contrastive. Exactly what is meant by “contrastive” and by “all other things are equal,” however, is not always properly understood and is the central focus of this paper.

Section 2 of the paper discusses the topic of phonological contrast and seeks to clarify the difference between phonemic contrast and underlying contrast. Next, factors that can affect the comparability of tone data are explored: section 3 identifies phonological factors, and section 4 looks at grammatical ones. Section 5 critiques examples of data that are not phonologically comparable from a generative perspective, including those in (1) and (2), above; and section 6 presents and discusses a paradigm of data that are genuinely comparable. Finally, there is an appendix of phonological and grammatical factors that must be borne in mind when establishing underlying tonal contrasts.

2. PHONOLOGICAL CONTRAST. In everyday usage, the term “contrast” is only meaningful if the things in contrast are comparable, that is, they are the same with respect to some agreed upon basis for comparison. For example, one can say that “red” contrasts with “orange” but not with “moon” if the basis for comparison is *colour*, because “red” and “orange” are both colours and “moon” is not. However, “red” does not contrast with “orange” if the basis for comparison is *primary colour* because while “red” is a primary colour that contrasts with blue and yellow, the other primary colours, “orange” is only a secondary colour, one derived from the coalescence of red and yellow.

When discussing sounds that are phonologically contrastive, which sounds are in contrast depends upon what the agreed upon basis for comparison is. For some linguists, phonological contrast means phonemic contrast in the structuralist sense, while for others, it means underlying contrast in the generative sense. Neither interpretation is necessarily right nor wrong, but since the basis for comparison is different in each case, it is important to understand which interpretation is assumed when an author employs the term.

In order to better understand the difference between phonemic contrast and underlying contrast, imagine the following scenario. A hypothetical language has three underlying vowels: /i, u, a/, confirmed by a minimal triplet of monomorphemic verbs: *bi* ‘run’, *bu* ‘walk’, and *ba* ‘swim’. Based only on these data, a phonemic analysis would also conclude a three-way contrast and would say there were three vowel phonemes. However, there are more data.

(3) More data from hypothetical language

bi-i	→	[bi]	‘he runs’
bu-i	→	[bwi]	‘he walks’
ba-i	→	[be]	‘he swims’

bi-u	→	[bj <u>u</u>]	‘he ran’
bu-u	→	[b <u>u</u>]	‘he walked’
ba-u	→	[b <u>o</u>]	‘he swam’

Despite the additional data, there is still no more than a three-way surface contrast in any given grammatical environment (isolation, present tense, or past tense), so a generative analysis would still conclude no more than three underlying vowels. A structuralist analysis, however, would now conclude five vowel phonemes, and would support them with the minimal quintuplet in (4).

(4) Hypothetical minimal quintuplet

bi	‘run’
be	‘he swims’
ba	‘swim’
bu	‘walk’
bo	‘he swam’

No one would dispute the fact that these five forms contrast phonologically. However, one must be careful not to mistake the five phonemic vowels of a structuralist analysis for five underlying vowels of a generative analysis, because the basis for comparison differs between the two analyses. In the case of the structuralist one, the five vowels are comparable because the basis for comparison is solely the phonological similarity of the *surface* environments. In the case of the generative analysis, however, the five vowels are not comparable because [e] and [o] are derived vowels (secondary colours, as it were) while [i], [u], and [a] are underlying vowels (primary colours, as it were). Equating phonemic contrasts with underlying ones can be misleading, as the data in (3) and (4) (and the discussion in the remainder of this paper) hopefully demonstrate.

Contesting this claim, an anonymous reviewer has stated that all words, whether they consist of derived forms or underived forms, “are ‘spelled’ with the same finite set of phonemes,” so establishing underlying contrasts with data from derived forms should not pose a problem for tone analysis. According to this reviewer, one *could* argue against this position “if one found that a tone resulting from coalescence ALWAYS results from coalescence—but at least in segmental phonology, a derived-only segment is a relative rarity.” I would counter this by saying that derived-only segments are only as rare as segmental coalescence itself is rare, and, in fact, derived-only segments may not be all that rare.⁵ Regardless of their rarity (or lack thereof), given the prevalence of floating tones⁶ in many languages, the coalescence of floating tones with non-floating tones is not rare at all, and tonal coalescence often does

⁵ Larry Hyman (personal communication) also questions the rarity of derived-only segments, citing the front rounded vowels of some languages as an example. To this one could also add the labialized consonants of many African languages.

⁶ A floating tone is a tone that is underlyingly unassociated to any tone-bearing unit (TBU), but which is nevertheless postulated to exist because of its surface tonal effect on neighboring TBUs.

result in derived-only tones. For example, rising and falling contours are derived-only tones in many languages, as are high tones that have been downstepped by preceding floating low tones. Since the existence of floating tones can only come to light through tone analysis, linguists in the early stages of that analysis cannot assume that phonemic contrasts and underlying contrasts are always one and the same. Basing underlying contrasts on data from mixed grammatical categories is therefore *not* a best practice.

The above is not to say that one does not need to investigate surface contrasts. Indeed, both structuralists and generativists need to do this. However, as just demonstrated in (3) and (4), the criteria for deciding which surface forms are in contrast are different for each type of analysis. In the case of phonemic analyses, it is only necessary to control for the phonological environments of the surface data being compared. In the case of generative analyses, however, it is necessary to control for *all* factors, including grammatical ones, that can potentially affect the comparability of the surface data being compared. The present work identifies those factors that must be the same in order to establish surface tone contrasts that reveal underlying contrasts.

Before discussing what this means in the context of tone analysis, let's continue to consider how it applies to segmental phonology. Imagine that you are reading an article that claims that English [eng] has a phonological contrast between the consonants *t* and *d* and supports this claim with the data in (5).

(5) Contrast between /t/ and /d/ in English

hɔt	'hot'
bæd-li	'badly'

While it is not unusual for a language to contrast /t/ and /d/, few linguists would accept these data as supporting this claim. The reason, of course, is that these data are not phonologically comparable; there are at least two reasons why one or the other of these consonants could be derived. With regard to the *t* (supported by 'hot'), it is voiceless and occurs word finally. Consonant devoicing is not uncommon in word final environments, so [hɔt] could easily be derived from /hod/. In the case of the *d* (supported by 'badly'), it is voiced and occurs between a vowel and *l*, both of which are voiced. This is an environment where voicing might be expected, so [bæd-li] could easily be derived from /bat-li/.

In order to demonstrate that these consonants truly contrast phonologically, they need to be presented in comparable phonological environments.

(6) More English

a. Adj.		b. Adv.	
hɔt	'hot'	hɔt-li	'hotly'
bæd	'bad'	bæd-li	'badly'

Such data could show both consonants occurring word finally, as in (6a), or both being followed by the adverbial suffix *-li*, as in (6b), but not the mix of both in (5) that compromises the phonological comparability of the data.

While few linguists compromise phonological comparability when dealing with seg-

mental data, they sometimes unwittingly do so when supporting tonal contrasts, due to the presence of floating tones. In this case, the analyst is totally unaware of the presence of a floating tone in the environment and wrongly assumes that the environments are comparable. For examples of this and further discussion, see section 4.1.

The principles followed for establishing tonal contrast are essentially the same as those followed for establishing segmental contrast. *All factors that can potentially affect the surface realization of the tones being compared must be the same for all data that are being compared.* Such factors can be divided into two major groups: a) phonological and b) grammatical.

3. PHONOLOGICAL FACTORS THAT AFFECT TONAL CONTRAST. Phonological factors that can affect the comparability of tone data obviously include any tones (floating or otherwise) that are adjacent to the tone patterns being compared. Other (perhaps less obvious) phonological factors that can affect the comparability of tone data include the number of TBUs in morphemes, the sonorancy of codas, and in some languages, the laryngeal states of consonants and vowels (e.g., voiced or voiceless obstruents in the case of consonants, and breathiness, creakiness, or glottalization in the case of vowels).⁷ The general tendency for linguists to prefer tonal contrasts that are based on minimal pairs/triplets, etc., has, of course, the benefit of avoiding data compromised by segmental factors. But in any case, it's important to look at a number of examples of how segments can influence tonal contrast.

Citing *Jiangsusheng he Shanghai Shi Fangyan Gaikuang* 1960, Bao 1990: 20-21 describes the Songjiang dialect of Wu Chinese [wuu], spoken in urban Shanghai, China, as having six different possible pitch patterns on CV words. Some consonants, however, interact with tones, with voiceless obstruents causing tones to be realized on the higher yin-register, and voiced obstruents and sonorants causing tones to be realized on the lower yang-register. For CV words that begin with voiceless obstruents, one finds only three contrastive pitch patterns: level, rising, and falling. Similarly, for CV words that begin with voiced obstruents or sonorants, one finds these same three patterns, but realized on the lower yang-register. This is displayed graphically in (7).

(7) Songjiang consonant-tone interaction

	Yin-register	Yang-register
Level	$\left[\begin{array}{c} - \\ ti \end{array} \right]$ 'bottom'	$\left[\begin{array}{c} - \\ di \end{array} \right]$ 'brother'
Rise	$\left[\begin{array}{c} / \\ ti \end{array} \right]$ 'emperor'	$\left[\begin{array}{c} / \\ di \end{array} \right]$ 'field'

⁷ See also Mazaudon (this volume).

	$\left[\begin{array}{c} \backslash \\ \text{ti} \end{array} \right]$	$\left[\begin{array}{c} \backslash \\ \text{di} \end{array} \right]$
Fall	‘low’	‘lift’

Examining data like these, we conclude that there are only three underlying pitch patterns in these data, and that it is the voicing quality of the consonants that multiplies this number by two.⁸

Next, let’s look at how the sonorancy of codas can affect the realization of underlying tone patterns. Controlling for sonorancy has long been a tradition in Asian linguistics where the term “checked” refers to syllables that end in an “occlusive coda” (e.g., *p, t, k*) and the terms “smooth” or “slack” refer to both open (e.g., CV) syllables (possibly with an offglide) and those ending in a nasal consonant (Chen 2000: 5). The need to control for coda sonorancy, however, is not restricted to Asia. To take an example from Africa, compare the surface patterns of these four words in Chumburung [ncu], a Guang language spoken in Ghana (personal field notes).

(8) Sonorant vs. non-sonorant codas in Chumburung

	Root /HL/		Root /H/
	$\left[\begin{array}{c} - \\ \backslash \end{array} \right]$		$\left[\begin{array}{c} - \\ - \end{array} \right]$
a.	ki-baŋ ‘paddle’	b.	ki-laŋ ‘jug’
	$\left[\begin{array}{c} - \\ - \end{array} \right]$		$\left[\begin{array}{c} - \\ - \end{array} \right]$
c.	ki-teʔ ‘feather’	d.	ki-brʔ ‘hill’

All of the words in (8) are nouns, drawn from the C3 noun class, and all have stems with the syllable profile CVC. In Chumburung, the voicing quality of onset consonants does not influence tone, but all codas are moraic and can have tones associated to them autosegmentally. Given these facts, since ‘paddle’ and ‘jug’ in (8a) and (8b) have different surface patterns but are otherwise identical in ways that can affect tone, one can reasonably conclude that they have different underlying root tone patterns. Following this same line of reasoning, since the words for ‘paddle’ and ‘feather’ in (8a) and (8c) also have different surface patterns, one might be tempted to think that they also have different underlying patterns. This, however, would not be a safe assumption because while both words have CVC stems, their codas are different with regard to sonorancy. In Chumburung, this factor plays a role in whether tones associated to codas are pronounced.

Phonologically, the roots of both ‘paddle’ and ‘feather’ have the *same* tone pattern, analyzable as /HL/ (see below). In each case, the high tone associates to the nucleus of the syllable and the low tone to the coda. This is demonstrated with autosegmental representations in (9).

⁸ Dave Odden (personal communication) has suggested that rather than consonants affecting tone, as Bao claims, these data are equally amenable to an analysis whereby tones affect consonants. Either way, when analyzing tone systems like this, it is still necessary to factor into the analysis the voicing quality of consonants.

(9) Autosegmental representations

a. $\begin{array}{c} \text{L HL} \\ | | | \\ \mu \mu \mu \\ \text{k}\bar{\text{i}}\text{-ba}\eta \end{array} \rightarrow \begin{array}{c} \left[- \ \backslash \right] \\ \text{k}\bar{\text{i}}\text{-ba}\eta \end{array} \text{ ‘paddle’}$
 C3-paddle

b. $\begin{array}{c} \text{L HL} \\ | | | \\ \mu \mu \mu \\ \text{k}\bar{\text{i}}\text{-t}\acute{\text{e}}? \end{array} \rightarrow \begin{array}{c} \left[- \ - \right] \\ \text{k}\bar{\text{i}}\text{-t}\acute{\text{e}}? \end{array} \text{ ‘feather’}$
 C3-feather

With respect to $\text{k}\bar{\text{i}}\text{-b}\acute{\text{a}}\eta$ in (9a), because the final mora (η) is sonorant, the low tone associated to it can be heard as the lower part of the falling contour. With respect to $\text{k}\bar{\text{i}}\text{-t}\acute{\text{e}}?$ in (9b), because the final mora (?) is not sonorant, the low tone associated to it cannot be heard, with the result that only the high tone on the (sonorant) nucleus is heard. As a result, the surface pattern of $\text{k}\bar{\text{i}}\text{-t}\acute{\text{e}}?$ in (8c) and (9b), whose root is /HL/, is identical with that of $\text{k}\bar{\text{i}}\text{-b}\acute{\text{i}}?$ in (8d), whose root is /H/. Evidence for the contrast between $\text{k}\bar{\text{i}}\text{-t}\acute{\text{e}}?$ and $\text{k}\bar{\text{i}}\text{-b}\acute{\text{i}}?$ may be seen in (10) by comparing the surface forms of all four words when pronounced before $\text{s}\acute{\text{o}}$ ‘scent’, itself underlyingly /H/.⁹

(10) Root /H/ contrasted with root /HL/ before /H/

Root /H/	Root /HL/
$\left[- \ - \ - \right]$	$\left[- \ - \ - \right]$
a. $\text{k}\bar{\text{i}}\text{-lan } \text{s}\acute{\text{o}}$ ‘jug’s scent’	b. $\text{k}\bar{\text{i}}\text{-ban } \text{s}\acute{\text{o}}$ ‘paddle’s scent’
$\left[- \ - \ - \right]$	$\left[- \ - \ - \right]$
c. $\text{k}\bar{\text{i}}\text{-b}\text{II} \text{s}\acute{\text{o}}$ ‘hill’s scent’	d. $\text{k}\bar{\text{i}}\text{-tee } \text{s}\acute{\text{o}}$ ‘feather’s scent’

Unsurprisingly, when high-toned $\text{s}\acute{\text{o}}$ occurs after $\text{k}\bar{\text{i}}\text{-l}\acute{\text{a}}\eta$ ‘jug’ in (10a) and $\text{k}\bar{\text{i}}\text{-b}\acute{\text{i}}?$ ‘hill’ in (10c), it is realized with a phonetically high pitch, confirming underlying /H/ for the roots of these two words. In the case of $\text{k}\bar{\text{i}}\text{-b}\acute{\text{a}}\eta$ ‘paddle’ in (10b) and $\text{k}\bar{\text{i}}\text{-t}\acute{\text{e}}?$ ‘feather’ in (10d), however, $\text{s}\acute{\text{o}}$ is downstepped, which in Chumburung confirms the presence of a final low

⁹ Glottal stops in Chumburung are only pronounced phrase finally; elsewhere their moras are realized with the vowel quality of the preceding vowel. When nasal consonants occur before other consonants within words and across word boundaries, they assimilate to the place of articulation of the following consonant.

tone in each of these words. *kî-báîŋ* itself ends with a level pitch in this medial environment due to a constraint in Chumbrung that prohibits the occurrence of phrase-medial falling pitches. What may initially be surprising, though, given the level nature of the final pitch of *kî-té?* when it is pronounced in isolation, is the surface realization of *só* in (10d), which shows it being downstepped after *kî-té?*. The logical conclusion is that the underlying patterns of *kî-bí?* and *kî-té?* are *not* the same since at this point it is only an underlying difference that could result in the different surface realizations of *kî-bí* *só* and *kî-tée* '*só*' in (10c) and (10d), respectively. We further conclude that the underlying patterns of *kî-báîŋ* and *kî-té?* are the same, since they produce similar results in the case of the following high-toned *só* in (10b) and (10d). The reason the isolation patterns of *kî-báîŋ* and *kî-té?* are different when they are pronounced in isolation is therefore due to the difference in sonorancy of their final TBUs.

The foregoing discussions demonstrate that in order for tonal differences to be contrastive, any segmental differences must not affect tones differently. As stated elsewhere, this principle is generally well accepted, and controlling for segmental differences is usually carried out carefully, if not to a fault. This brings us to the matter of controlling for tonal differences in the environments by regulating grammatical factors.

4. GRAMMATICAL FACTORS THAT AFFECT TONAL CONTRAST. Since grammatical factors play a significant role in determining the juxtaposition of morphemes, and since the juxtaposition of morphemes plays a significant role in determining the comparability of phonological data, it is also necessary to take grammatical factors into consideration when establishing tonal contrasts.¹⁰

In principle, controlling for comparable adjacent tones when conducting tone analysis is little different from controlling for comparable adjacent segments when conducting segmental analysis. In practice, however, the latter is easier. When conducting segmental analysis, non-comparable phonological environments are usually readily apparent because all morphemes present normally have segmental representation. That being said, morphemes that consist solely of floating segmental features have nevertheless been reported. Roberts 1994, for example, describes the incompletive aspect marker in Kanembu [kbl], a Nilo-Saharan language spoken in Chad, as consisting solely of a floating +ATR feature. Thanks, however, to the rarity of floating segments, linguists seldom base segmental contrasts on non-comparable data. Tone languages, on the other hand, often have morphemes that either: a) end in floating tones, or b) consist solely of floating tones, and this often causes linguists to inadvertently base tonal contrasts on non-comparable data.

Here is an example of a morpheme in Kenyang [ken], a Bantoid language spoken in southwestern Cameroon (personal field notes), that consists solely of a floating low tone. In Kenyang, floating low tones cause following high tones to be downstepped relative to preceding high tones. They also block high tones from spreading in environments in which they would otherwise do so. The examples in (11) demonstrate contrast between high and low-toned verbs pronounced in imperative utterances.

¹⁰ See also Mazaudon (this volume) and Rice (this volume).

(11) High and low-toned imperative verbs in Kenyang

/H/	$\left[\begin{array}{c} - \\ \text{pa} \end{array} \right]$	/L/	$\left[\begin{array}{c} - \\ \text{k}\text{ɔ} \end{array} \right]$
	‘Spit!’		‘Walk!’

The examples in (12) show these same two verbs as they are pronounced in perfective and imperfective aspects when preceded by the high-toned 3P subject prefix *bá-*.

(12) Perfective and imperfective forms in Kenyang

		Perfective			Imperfective		
a.	/H/	/bá-Ø-pá/	→	$\left[\begin{array}{cc} - & - \end{array} \right]$	/bá-`-pá/	→	$\left[\begin{array}{cc} - & - \end{array} \right]$
		3P-PERF-spit		ba-pa	3P-IMPERF-spit		ba-pa
				‘they spat’			‘they are spitting’
b.	/L/	/bá-Ø-kɔ/	→	$\left[\begin{array}{c} - \\ \backslash \\ \text{ba-k}\text{ɔ} \end{array} \right]$	/bá-`-kɔ/	→	$\left[\begin{array}{cc} - & - \end{array} \right]$
		3P-PERF-walk		ba-kɔ	3P-IMPERF-walk		ba-kɔ
				‘they walked’			‘they are walking’

The perfective aspect in Kenyang is morphologically unmarked, so in (12a), when the high-toned verb *pá* is in the perfective aspect, it is unaffected by any other tone and is phonetically realized at the same high pitch level as the preceding high-toned prefix *bá-*. However, in the imperfective aspect, the floating low-toned imperfective marker downsteps the high tone of *pá*, causing it to be realized at a significantly lower pitch than the preceding high-toned prefix. In (12b), when the low-toned verb *kɔ* is in the perfective aspect, it undergoes high-tone spreading from the preceding high-toned prefix. This causes *kɔ* to be phonetically realized with a high-falling pitch, a concatenation of the high and low tones realized on the single TBU. However, in the imperfective aspect, the high tone of the preceding prefix does not spread, due to the presence of the intervening floating low tone. This causes *kɔ* to be phonetically realized with a low pitch. Although one never hears the floating low tone of the imperfective aspect marker, its presence is nevertheless confirmed by its effects of downstepping the following high-toned *pá* in (12a) and blocking the spreading of the high tone from the subject prefix onto the following low-toned *kɔ* in (12b).

Short of carrying out significant tone analysis, it is often impossible to be sure whether a floating tone is present or not. Given this uncertainty, any time tone analysis is carried out on data from mixed grammatical environments, the potential exists for those data *not* to be phonologically comparable. When grammatical marking employs segments, words that are not phonologically comparable are usually quickly spotted and eliminated from unfair comparisons. However, when grammatical marking consists solely of floating tones, it is not necessarily immediately clear when the comparability of phonological environments has been compromised. This makes it much harder for linguists to base tonal contrasts on

phonologically comparable data, which, in turn, creates something of a *Catch-22* situation. On the one hand, one cannot carry out good tone analysis if it is not based on comparable data. On the other hand, one often cannot evaluate the phonological comparability of data without first carrying out a significant amount of tone analysis. Researchers in the early stages of tone analysis are therefore in no position to assume that tone data from non-homogenous grammatical environments are phonologically comparable. In order to ensure that all tonal comparisons are based on phonologically comparable data, they must ensure that those same data are grammatically homogenous.

Due, then, to the pervasive nature of floating tone morphemes, all tonal comparisons need to be carried out in identical grammatical environments to ensure that the data are phonologically comparable. My experience suggests that the four grammatical environments listed in (13) are especially problematic.

- (13) Grammatical environments that can potentially affect phonological comparability
- a) lexical category of stem (e.g., noun, verb)
 - b) grammatical categories of affixation (e.g., person, number, tense, aspect)
 - c) word class (e.g., noun class, verb class)
 - d) syntactic environment (e.g., isolation, verb-object construction)

Given the breadth of linguistic phenomena that occur across languages, lists like these cannot be definitive. However the diversity of the environments here will hopefully help the reader to recognize the types of environments that need to be considered when analyzing the tone of a particular language.

4.1 LEXICAL CATEGORY OF STEM. Hopefully, the preceding discussions make it clear that tonal contrasts can only be established after controlling for *all* of the factors that can affect how tone is realized. Judging, however, from the prevalence in the literature of examples that mix nouns with verbs with adjectives, etc. (cf. the examples in (1) and (2), above), the reason why *lexical category* should be one of those factors is clearly not well understood.

The problem with basing tonal contrasts on words from different lexical categories is that their grammatical environments can never be identical. Words are always more than just lexemes. Verbs pronounced in isolation, for example, are often imperatives or infinitives, and nouns pronounced in isolation are often nominatives or other cases. While distinctions like these may or may not be marked morphologically, they often are, and if that marking consists solely of floating tone affixes, this will not necessarily be immediately obvious. In fact, it may not become obvious until the tone analysis is fairly well advanced. For this reason, tonal contrasts should not be based on data from mixed lexical categories.

Consider the following apparent minimal pair in Chumburung.

- (14) Apparent minimal pair in Chumburung

k̀p̀à	‘want’
k̀p̀á	‘path’

At first glance, this is an ideal minimal pair, appearing to provide clear evidence of a lexical tone contrast. However, there is both a lexical difference *and* a lexical category difference, and without further data, it is impossible to know which of these is responsible for the tonal differences. As it happens, all verbs in the language with the syllable profile CV are pronounced with low pitch in isolation, and all nouns with the same syllable profile are pronounced with high pitch in isolation. This may be seen with the words in (15), transcribed as they are pronounced in isolation.

(15) More apparent minimal pairs in Chumburung

k̀pà	k̀pá
‘want’	‘path’
fè	fé
‘sell’	‘rope’
ẁù	ẁú
‘die’	‘thorn’
ɲ̀ì	ɲ̀í
‘know’	‘mother’
kà	ká
‘advise’	‘wife’

Chumburung verbs pronounced in isolation are always in the imperative mood, which consists solely of a floating low-tone prefix; this tone docks onto the first syllable of the verb and completely obliterates the underlying tone of that syllable. When the syllable profile is CV, any underlying tonal contrasts between verbs are neutralized. As for Chumburung nouns, most are polysyllabic, with the first syllable consisting of a noun class prefix (cf. the nouns in (8)). Although lacking a *segmental* prefix, many nouns with the word shape CV can nevertheless be analyzed as having a *floating* high tone prefix. As is the case with verbs, this floating tone prefix neutralizes the underlying patterns of CV noun stems. In order to establish lexical contrast, it is therefore necessary to compare nouns only with nouns, and verbs only with verbs (i.e., words from the same lexical category).

These examples from Chumburung clearly show why lexical categories should be kept separate in tone analysis. At the same time, they could also create the false impression that non-comparable data from mixed lexical categories is relatively easy to spot. But this is not always the case. Complete complementarity of this kind is somewhat unusual in languages, but it occurs in Chumburung due to two separate factors. The first is that the words of both lexical categories have floating tone prefixes. The most common case, of course, is for only one or the other of two lexical categories to have a floating tone affix. In that case, complementarity between the surface tones of words from the two different categories is greatly reduced, making it much harder to recognize data that are not phonologically comparable.

The other factor responsible for this complementarity in Chumburung is that the floating tone prefixes have the same surface effect on *all* words of that particular category: all

nouns are always realized with high pitch, and all verbs with low pitch, regardless of their underlying patterns. More common would be for at least one, if not both, of the floating tone prefixes to *interact* with the underlying stem tones, rather than to *replace* them, as happens in Chumburung. When floating tone affixes interact with underlying stem tones, the result is often not a single consistent surface pattern, but rather a number of different surface patterns, depending on the number of different underlying stem patterns. Again, this more normal behaviour significantly reduces complementarity between the surface tones of mixed lexical categories, making it that much harder to identify phonologically non-comparable data.

Returning to the discussion of Chumburung tone, one of the simplest ways to establish lexical tone contrast between CV verbs is to compare their nominalized forms. One way to nominalize verbs in Chumburung is to prefix the verb stem with the *ki-* noun class prefix, which creates a form similar to a gerund in English.

(16) Nominalized verbs compared

/H/	kpà	‘want’	kì-kpá	‘wanting’
/L/	ǰǰ	‘know’	kí-ǰí	‘knowing’
/H/	kà	‘advise’	kì-ká	‘advising’
/H/	fè	‘sell’	kì-fè	‘selling’
/H/	wù	‘die’	kù-wú	‘dieing’

Segmentally, the vowel of the *ki-* prefix agrees with the first stem vowel with respect to ATR and Back/Round values (Snider 1988). Tonally, all noun class prefixes in Chumburung behave identically and are phonetically realized in a “polar” manner: when the stem tone is high, the prefix is low, and when the stem tone is low, the prefix is high. A high prefix tone spreads rightwards onto the stem and completely obliterates the low tone of the stem (e.g., *kí-ǰí* ‘knowing’).¹¹ In (16), since all of the nominalized verbs have the same prefix, contrast is therefore clearly established between ‘know’ (underlyingly /L/) and the other verbs (underlyingly /H/).¹²

Moving to the nouns, one of the simplest ways to establish lexical tone contrast between CV nouns in Chumburung is to compare their plural forms. Before looking at some examples, it should be noted that the singular forms of the nouns in (15) belong to the same noun class (the ‘floating high-toned prefix’ class), which permits us to compare them with each other. However, their plurals belong to two different classes, the I-class and the A-class. As discussed below, when dealing with a noun-class language, it is important to ensure that nouns chosen for comparison all belong to the same class. Following this principle, we should compare I-class plural forms only with other I-class plurals, and the same

¹¹ Snider 2013 analyzes these stems as underlyingly toneless, which explains the lack of a low tone in this environment.

¹² The interested reader is referred to Snider 2013 for an in-depth discussion of the Chumburung tone system.

for A-class plurals, as indeed is done in (17) and (18) below. In Chumburung, however, all noun class prefixes behave identically with respect to tone, so we would obtain the same results even without following this strict principle. This of course is not necessarily the case with other noun class languages, as is shown below in section 4.3.

(17) I-class plural nouns compared

/H/	$\widehat{\text{kpá}}$	‘path’	$\grave{\text{i}}\widehat{\text{kpá}}$	‘paths’
/L/	fè	‘rope’	$\grave{\text{i}}\text{fè}$	‘ropes’
/H/	wú	‘thorn’	$\grave{\text{i}}\text{wú}$	‘thorns’

(18) A-class plural nouns compared

/L/	jí	‘mother’	$\acute{\text{ó}}\text{jí}$	‘mothers’
/L/	ká	‘wife’	$\acute{\text{á}}\text{ká}$	‘wives’

The tonal behaviour of the A- and I- class prefixes in (17) and (18) is exactly the same as that of the *kI*- class prefix of the nominalized verbs in (16), i.e., a polar tone on the prefixes combined with spreading from a high prefix. In (17), one can see that contrast is clearly established between ‘rope’, which is underlyingly /L/ and ‘path’ and ‘thorn’, which are underlyingly /H/.

Returning to the apparent minimal pairs in (15) above, we learn nothing about their underlying tones by comparing the isolation forms of verbs with the isolation forms of nouns. Those same words are repeated in (19), together with what we now know of their underlying tones.

(19) Actual underlying tones compared

Underlying	Surface	Underlying	Surface
/H/	$\widehat{\text{kpà}}$	/H/	$\widehat{\text{kpá}}$
	‘want’		‘path’
/H/	fè	/L/	fè
	‘sell’		‘rope’
/H/	wù	/H/	wú
	‘die’		‘thorn’
/L/	jì	/L/	jí
	‘know’		‘mother’
/H/	kà	/L/	ká
	‘advise’		‘wife’

Perusal of the data in (19) reveals that of the five apparent minimal pairs in (15), there is no underlying tonal contrast between the roots of three of them. Of the two that do demonstrate contrast, the underlying tones are actually *opposite* to what one might expect: the surface forms of ‘sell’ and ‘advise’ are [L], but their underlying forms are /H/. It is the exact opposite for ‘rope’ and ‘wife’. The surface tones of the apparent minimal pairs really tell us nothing beyond the fact that one member of each pair is a noun and the other is a verb.

While the tones of words from mixed lexical categories are not always influenced differently, they often are when floating tones are involved. A good presentation of lexical tone contrast therefore draws data from the same lexical category. Presenting tonal contrasts based on data from different lexical categories at best erodes confidence in what might otherwise be accurate analyses, and at worst supports wrong conclusions.

4.2 GRAMMATICAL CATEGORIES OF AFFIXES. As demonstrated above, it is often the case that certain morphemes are only revealed in the presence of other morphemes. It is therefore important to ensure that any obligatory affixation is identical for all data that are being compared. If one were comparing the tones of verb roots, for example, it would be important that all of the verbs be identical morphologically (e.g., same tense and person marking, etc.), as opposed to being a collection of mixed morphologies.

Compare the Kenyang data in (20).

(20) Kenyang (personal field notes)

bǎ-té	‘you (pl.) have drilled’
bǎ-tē	‘you (pl.) stood’
bá-tê	‘they have stood’
bá-tè	‘they drilled’

Since these four words are identical segmentally, there is a four-way phonemic contrast between high, mid, falling, and low tones.¹³ However, from the discussion of Kenyang above in section 2, the reader will know that these phonemic contrasts do not accurately reflect the underlying tones. As beautiful and convincing as these data might initially appear to be, they do not establish the actual underlying contrasts because for any two words, there is always more than one morpheme that is different between them. For example, look at the roots for ‘stand’ and ‘drill’. Do their underlying tones contrast? If we compare ‘you (pl.) have drilled’ with ‘you (pl.) stood’, the root for ‘drill’ is realized with a high tone while that for ‘stand’ is realized with a mid. Although they are different, one cannot be sure why. The differences could be due to: a) ‘drill’ and ‘stand’ having different underlying tones, b) ‘drill’ being influenced by its floating tense/aspect marker, c) ‘stand’ being influenced by a different floating tense/aspect marker, d) both ‘drill’ and ‘stand’ being influenced by different respective tense/aspect markers, or e) ‘drill’ and ‘stand’ having different underlying tones together with any one or more of the other possibilities. When the morphological structures of data that are being compared are not identical, it is impossible to establish

¹³ In actual fact, the “mid” tone is a downstepped high tone, and the “falling” tone is a concatenation of high and low tones realized on a single TBU.

underlying tonal contrasts between them if some of the morphemes consist solely of floating tones.

On the other hand, as may be seen in (21), underlying tonal contrast between any two morphemes can readily be established when their morphological structures are identical, despite the fact that there are multiple morphemes present in each word.

(21) Morphologically comparable data in Kenyang

Tense/aspect	2P	3P	Gloss
Imperfective	bǎ-tè	bá-tè	'you (pl.)/they are standing'
Perfective	bǎ-tê	bá-tê	'you (pl.)/they stood'
Imperfective	bǎ-tē	bá-tē	'you (pl.)/they are drilling'
Perfective	bǎ-té	bá-té	'you (pl.)/they drilled'

This data set allows us to compare the roots 'drill' and 'stand' in four different identical environments: 'you (pl.) are drilling/standing', 'they are drilling/standing', 'you (pl.) drilled/stood', and 'they drilled/stood'. The same can be said for any other given morpheme pair in (21). The true contrasts are clearly revealed when one controls for all of the grammatical categories of the data.

4.3 WORD CLASS. In many languages, words from the same lexical category are grouped into classes, often based, at least partially, on semantic grounds. The Niger-Congo languages of Africa, for instance, are noted for their complex noun class systems, and many languages also have verb classes. Classes are distinguished from each other by means of their morphology; in the case of tone languages, these morphological differences may be signaled solely by differences in tone. As is the case for other grammatical morphemes, word class markers can consist solely of floating tones.

As demonstrated above in other contexts, a tonal contrast between two lexemes can only be established if any affixes involved are tonally identical. In languages with word classes, this is more difficult because of the extensive morphological agreement between words (e.g., nouns, in the case of a noun class system) and other related words in the utterance that relate to them (e.g., adjectives, pronouns, etc.).

Compare the following nouns from Mada [mda], a Benue-Congo language spoken in Nigeria.¹⁴

(22) Selected words in Mada

d̄ɔ́í	'mortar'
kā	'compound'
kì	'thing'
dǎ	'cutlass'

¹⁴ Most of these data were elicited by Norman Price during a phonology workshop that I conducted in Jos for the Nigeria Bible Translation Trust in 1994. I am very grateful to Norm for making these data available to me. During another workshop that I conducted in Jos, in September 2012, I elicited additional Mada data from a close relative of the first speaker.

Although these words do not constitute a minimal quadruplet, if assured that the tones of Mada are unaffected by vowel and consonant quality (as is the case), many linguists would consider the nouns in (22) as providing clear evidence of an underlying four-way lexical contrast between high, mid, low, and rising tones because each word is a singular noun, each has a CV syllable profile, and each is pronounced in isolation. But such a conclusion would be wrong.

Mada has noun classes, and a careful analysis of Mada tone needs to take this into consideration. My work on Mada to this point suggests that over 70% of CV nouns form their singulars and plurals according to one or other of the pairings in (23).

(23) Singular/plural noun class prefix pairings in Mada

	Singular pfx.	Plural pfx.
Pairing 1	Floating H	mə-
Pairing 2	Floating L	Floating H

Of these, over 65% have one or the other of the underlying tone patterns: low, high, or toneless. These underlying stem patterns interact with the floating tone prefixes as shown in (24).

(24) Interaction of prefixal floating tones with stem tones in Mada

Floating pfx. tone	Stem tone	Surface realization	Floating pfx. tone	Stem tone	Surface realization
/L/	/H/	[Rising]	/H/	/H/	[H]
/L/	/L/	[L]	/H/	/L/	[M]
/L/	/∅/	[L]	/H/	/∅/	[H]

Here are some examples of Pairing 1 CV nouns.

(25) Pairing 1 CV nouns in Mada

Underlying Root tones	Singular Floating H		Plural mə-		
/H/	/´-d̄ɜ́í/	→	[d̄ɜ́í]	/mə-d̄ɜ́í/	→ [məd̄ɜ́í] ‘mortar’
/L/	/´-kə/	→	[kə]	/mə-kə/	→ [məkə] ‘compound’

In (25), the roots for ‘mortar’ and ‘compound’ are underlyingly /H/ and /L/, respectively. When these tones interact with the preceding low tone of the plural prefix mə-, they are realized as surface high and surface low, respectively. When they interact with the preceding floating high singular prefix, the underlying high-toned ‘mortar’ is, not unexpectedly, realized with surface high, while the underlying low-toned ‘compound’ is realized with surface mid, a merger of underlying high and low.

Here are some examples of Pairing 2 CV nouns.

(26) Pairing 2 CV nouns in Mada

Underlying Root tones	Singular Floating L		Plural Floating H		
/H/	/`-dá/	→	[dǎ]	/´-dá/	→ [dá] ‘cutlass’
/L/	/`-kì/	→	[kì]	/´-kì/	→ [kī] ‘thing’
/Ø/	/`-wu/	→	[wù]	/´-wu/	→ [wú] ‘mosquito’

In these examples, the lexemes ‘cutlass’, ‘thing’, and ‘mosquito’ are underlyingly /H/, /L/, and /Ø/, respectively. The surface tones of ‘cutlass’ and ‘thing’ are realized in a manner identical to that of the plural forms for ‘mortar’ and ‘compound’, respectively, in (25). The surface tone of ‘mosquito’, underlyingly toneless, is realized with the tone of whichever prefix is assigned to the word. Given this knowledge of the noun class system, although noun roots have three underlying patterns in the language, the seemingly four underlying tones of the words in (22) can actually be reduced to two: /H/ and /L/. The mid and rising pitches in (22) are the result of interaction between the stem tones and the floating high and low singular noun class prefixes.

(27) Underlying tones of selected words in Mada

Class	Pfx	Stem	Surface	
C1	Ⓜ	/H/	ḍǔí	‘mortar’
C1	Ⓜ	/L/	kǎ	‘compound’
C2	Ⓛ	/L/	kì	‘thing’
C2	Ⓛ	/H/	dǎ	‘cutlass’

The analysis of tone in Mada nouns is more difficult than for many other tone languages, even other ones with noun classes, because for all nouns, the singular prefix is a floating tone, and for many of them, the plural prefix is also a floating tone. This again presents something of a *Catch-22* type situation with respect to morphology and tone. When analyzing tone languages, it is therefore essential to interleave tone analysis with morphological and syntactic analysis.

In many noun class systems, there is extensive agreement between nouns, their modifiers, and any pronominal references. It may therefore be necessary to pay attention not only to the affixation of nouns, but also to that of other, concordant, elements in the utterance, to ensure that they can genuinely be compared.

The following is an example from Bamileke-Dschang [ybb], a Grassfields Bantu language spoken in Cameroon (data from Hyman & Tadadjeu 1976 and Hyman 1985).

(28) Same underlying patterns, different noun classes in Bamileke-Dschang

a. Class 1	b. Class 7
è-fǎ → èfǎ ‘chief’	à-pà → àpà ‘bag’
C1-chief	C7-bag

The two words in (28) have identical surface patterns. However, this is only one environment (viz., isolation). In order to be sure that these words have the same underlyingly tones, we need to compare their behaviour in other environments. One such environment is the associative (i.e., possessive) construction.

(29) Nouns compared in associative constructions

ɛ̀fɔ̃ ñḁḁḁ^wɪ̃] ‘chief of leopard’
 àpà ñ^hḁḁḁ^wɪ̃ ‘bag of leopard’

A cursory comparison of ‘chief’ and ‘bag’ in (29) suggests that although their tones are identical in their isolation forms, their underlying tone patterns must be different because of the different surface effects they have on the forms for ‘leopard’ when they are adjacent to it. But this comparison is misleading. As in most, if not all, Bantu languages, there is an associative marker (AM) between the “possessee” noun and the “possessor” noun that is concordant with the class of the head (possessee) noun. A phrase translated as ‘leopard’s chief’ therefore has the word order ‘chief of leopard’, with ‘of’ being the translation for the associative marker. In actual fact, ‘chief’ and ‘bag’, the words in (29), not only have identical isolation patterns, but they do, in fact, have identical *underlying* patterns. This fact is obscured, however, when one compares these words in the associative construction. This is because the associative marker for Class 1 nouns, to which ‘chief’ belongs, is underlyingly low-toned while that for Class 7 nouns, to which ‘bag’ belongs, is underlyingly high-toned (viz., /è-/ for AM.C1 and /á-/ for AM.C7), and this makes it impossible to directly compare these words in this construction. A further complicating factor is that the vowels of the associative markers are usually elided in surface forms, leaving only the interactions of their tones with other tones to signal their presence. In (30), we take a closer look at what’s going on in (29).

(30) Possessive construction in Bamileke-Dschang

a. Class 1

ɛ̀-fɔ̃ è ñ-ḁḁḁ^wɪ̃ → [ɛ̀fɔ̃ ñḁḁḁ^wɪ̃] ‘chief of leopard’
 C1-chief AM.C1 C9-leopard

b. Class 7

à-pà á ñ-ḁḁḁ^wɪ̃ → [àpà ñ^hḁḁḁ^wɪ̃] ‘bag of leopard’
 C7-bag AM.C7 C9-leopard

In (30), all underlying tones in ‘chief of leopard’ are low, with the result that all four TBUs in this phrase are realized with low pitch. In ‘bag of leopard’, however, the final low-toned TBU is phonetically downstepped in relation to the low tones of the preceding three TBUs.¹⁵ This difference in surface patterns can only be due to the underlying tonal difference in the associative markers. If one did not take into account the facts of elision

¹⁵ A detailed explanation of the Bamileke-Dschang data is beyond the scope of the present work. For further information, the interested reader is referred to Hyman 1985 and Chapter 7 of Snider 1999 for an analysis of how underlying tones are phonetically realized in the Bamileke-Dschang associative construction.

and the noun class system when comparing ‘chief’ and ‘bag’ in otherwise identical possessive phrases, one could erroneously conclude that their underlying patterns were different. Only by carefully comparing words from one class with words from that same class can problems of this nature be avoided.

Problems like this are not limited to just a few languages. Most languages of the Niger-Congo family (the largest in Africa) are tonal and have noun classes. Here are further examples from Niger-Congo languages, showing different grammatical affixes that are concordant with the class of the head noun and distinguished solely by tone.

(31) Noun class affixes differentiated solely by tone

Bakweri [bri] (Bessem 1987)	Mankon [nge] (Leroy 1980)	Akoose [bss] (Hedinger 2008)	Mundani [mnf] (Parker 1989)
C10 Noun í-	C1 Number ì-	C9 Subj. è-	C8 Genitive é-
C19 Noun ì-	C3 Number í-	C10 Subj. é-	C9 Genitive è-

In Bakweri, a Bantu language spoken in Cameroon, the noun class prefixes of Classes 10 and 19 are segmentally identical. Tonally, however, the Class 10 prefix is high and the Class 19 one low. Accordingly, if one were to compare a noun from Class 10 with another from Class 19, any surface contrast that arose would not necessarily represent a phonological contrast between the stem tones; such a difference could just as easily be attributed to the tonal difference between the prefixes.

The other languages presented in (31) show that similar tonal differences can distinguish the affixes of words that are concordant with the head noun: this means that when determining (or presenting) tonal contrasts in these languages, one must ensure that: a) any tonal comparisons between nouns are between nouns from the same class, and b) any tonal comparisons between words that are concordant with head nouns are also between words whose head nouns belong to the same class.

It is hoped that the foregoing discussions illustrate why one cannot ignore word classes when conducting tone analysis, and (once again) why it is so important to ensure that one compares like with like.

4.4 SYNTACTIC ENVIRONMENT. It is generally well understood that the surface realization of underlying patterns can be significantly affected by tones on adjacent words. Given this understanding, controlling for syntactic environments might seem to be an obvious prerequisite to carrying out tonal comparisons. The fact remains, however, that examples like (32), repeated from (2), continue to appear in the tonal literature.

(32) High, mid, and low pitches in Buli (Akanlig-Pare & Kenstowicz 2003)

H
M
L

CV: mǐ má: ‘I helped’ mā: ‘help!’ wà mà: mǐ ‘he helped me’

Some, or possibly all, of the three pitch levels of this triplet are derived, and this is attributable to two possible sources: a) the influence of floating tone grammatical affixes (see section 4.2), and/or b) the influence of tones from adjacent words (see discussion of this

example further below). Examples like this underscore how essential it is that tonal comparisons be made in comparable syntactic environments.

Next, consider the surface realizations of the Chumburung words for ‘hawk’ and ‘skin’ in (33) with a view to discerning whether their underlying patterns are the same or different.

(33) ‘hawk’ compared with ‘skin’

[- -]

dapu ‘hawk’

[- - - -]

owure wuri ‘chief’s skin’
chief skin

There is certainly a phonetic difference between the high pitches of ‘hawk’ and the mid pitches of ‘skin’, but their syntactic environments are different. The fact that there is another word adjacent to ‘skin’ creates the potential for the mid pitch in ‘skin’ to be derived, created by interaction of the underlying tones of the two words. Before one can conclude that the underlying tones of ‘hawk’ and ‘skin’ are different, one needs to see these words compared in identical environments. Whether or not any given environment includes more than one word is not necessarily relevant as long as the environments are truly identical.

(34) a. ‘hawk’ compared with ‘skin’ in multi-word environments

[- - - -]

owure dapu ‘chief’s hawk’

[- - - -]

owure wuri ‘chief’s skin’

[- - \]

cf. owure ‘chief’

b. ‘hawk’ compared with ‘skin’ in single-word environments

[- -]

dapu ‘hawk’

[- -]

wuri ‘skin’

Once one compares these words in identical environments, it becomes readily apparent that they behave identically in each of these environments, and there is no contrast between their underlying tone patterns. On the other hand, if one compares ‘hawk’ with another word, ‘monkey’, in the same multi-word environment, it is just as clear that these words *do* have an underlying tonal contrast, evidenced by their differing surface forms. This is demonstrated in (35).

(35) ‘hawk’ compared with ‘monkey’ in multi-word environments

$\left[\begin{array}{cccc} - & - & - & - \end{array} \right]$	
owure dapu	‘chief’s hawk’
$\left[\begin{array}{cccc} - & - & - & - \end{array} \right]$	
owure kɔti	‘chief’s monkey’

In these examples, the word ‘chief’ precedes both ‘hawk’ and ‘monkey’. The surface tones of both ‘hawk’ and ‘monkey’ therefore have the potential to be derived, due to influence from the tones of the preceding word. Given this possibility, without more data, these examples cannot tell us the complete story regarding the underlying tones of these two words. But no further data are needed to establish the tonal contrast itself, because the data are syntactically identical (unlike the situation in (32)).

While differences in grammatical environments do not automatically translate into differences in phonological environments, they certainly can, and, as we have seen, they often do. Due to the abundance of floating tone morphemes in tone languages, as well as the abundance of segmentally identical morphemes distinguished only by tone, the comparability of phonological environments is potentially compromised when the grammatical environments of the forms being compared are not identical. When analyzing (or presenting) data with a view to establishing tonal contrast, it is therefore vital to ensure that: a) the segmental environments of the data being compared are comparable in ways that affect tone, and b) the grammatical environments of the data being compared are identical in order to help ensure comparable tonal environments.

5 CRITIQUE OF NON-COMPARABLE DATA. We turn now to a critique of the data of (1), (7), and (2), repeated here as (36), (37), and (38), respectively.

(36) Chiquihuitlan Mazatec (Jamieson 1977)

$\left[\begin{array}{c} - \\ \hline \end{array} \right]$	$\left[\begin{array}{c} - \\ \hline \end{array} \right]$	$\left[\begin{array}{c} - \\ \hline \end{array} \right]$	$\left[\begin{array}{c} - \\ \hline \end{array} \right]$
čha	čha	čha	čha
‘I talk’	‘difficult’	‘his hand’	‘he talks’

Since the Mazatec data in (36) were originally intended to support phonemic contrasts, one does not want to cast them in a negative light because they fail to pass muster when measured by generative criteria. That being said, these data nevertheless illustrate well the fallacy of trying to establish underlying contrasts based on structuralist criteria. In this data set, ‘his hand’ appears to be an inflected noun, so one concludes that there must be at least two morphemes present. Anytime there is more than one morpheme present in the morphology of a word, the possibility exists that the surface pitch is derived. In this case, the (second-lowest) pitch of ‘his hand’ could easily be a derived level, resulting from the interaction of an underlying lower tone with an underlying higher tone. With no other grammatically related forms available for comparison, it is impossible to know what is going on. Similarly, ‘difficult’, realized with the second-highest pitch, is probably an adjective. While this word lets us know that there is indeed a phonemic pitch at this level in the language, we have no idea whether this level is underlying, assigned perhaps to the lexeme ‘difficult’, or whether it is derived, due perhaps to interaction between the tone of the lexeme ‘difficult’ and a floating tone affix peculiar to adjectives. Once again, in the absence of truly comparable data, one cannot be sure.

Finally, in the case of ‘I talk’ and ‘he talks’, we do find comparable environments: their roots are identical, their tenses are identical, and their segments are identical. The only non-phonetic difference between these words is the verbal person marking, so one can safely attribute the tone difference to this grammatical difference. This is therefore a true minimal pair whose tone contrast is rooted in genuinely comparable environments. But we need to go further: this comparison is only between the 1S and 3S tone affixes, and these are almost certainly floating tones interacting with a root tone and possibly a floating tone tense-aspect marker as well. Since there are multiple morphemes involved in the morphological make-up of both words, one, or possibly both, of the surface pitches of these words must be derived. Without additional grammatically comparable forms, the reader is unable to draw any firm conclusions about the underlying tones of either the root or the person markers. All one can say for sure is that the underlying tone patterns of the two person markers are different.

Clearly, the purpose of the data in (36) is to demonstrate the presence in the language of four *phonemic* tone heights using the very best type of data possible, namely a minimal quadruplet. There may very well be four *underlying* tone heights, but we would need other data than that of (36) to prove it. The only solid conclusion one can draw from these four words is that there are four phonemic levels of pitch. With the exception of ‘I talk’ and ‘he talks’, which are a minimal pair, these words do not form a minimal quadruplet in a generative analysis because the differences between them are not minimal.

This misunderstanding of what minimal really means within the context of evaluating comparable tone data causes many linguists to place undue confidence in the phonological comparability of segmentally identical data. A further example of this may be found in the Songjiang data in (7), repeated here as (37).

(37) Songjiang dialect of Wu (Bao 1990)

	Yin-register	Yang-register
	[-]	[-]
Level	ti 'bottom'	di 'brother'
	[/]	[/]
Rise	ti 'emperor'	di 'field'
	[\]	[\]
Fall	ti 'low'	di 'lift'

Again, these data were originally provided to support phonemic contrasts, so the above disclaimer for the Mazatec original analysis equally applies here. In this data set, of the three words in the Yin-register, the first two are nouns, and the last one is an adjective. And of the three words in the Yang-register, again, the first two are nouns, but the last one is a verb. This mismatch of lexical categories casts doubt on the claim that there is an underlying three-way tonal contrast in either consonant set. Such data are not ideal, and frankly there is no need for it. The discussion in Bao 1990, from which these data are drawn, claims that consonant voicing is the critical factor in determining the register on which the tones will be realized. Since nothing in this source claims that vowel quality or consonantal place of articulation are factors that influence Songjiang tone, there is no particular reason to limit the segments to *ti* and *di*.

Despite the grammatically mismatched data, in this particular case I do not actually doubt the original author's claims. I suspect that he could probably have easily presented comparable data, but perhaps did not do so due to the ideal of presenting data that are segmentally identical. All other things being equal, tonal contrasts that are supported by segmentally identical data truly are convincing because they eliminate the possibility of their pitch differences being partially or entirely caused by segmental differences. However, as stated above, the key lies in ensuring that all other things truly are equal. So while the data in (37) look nice, they are not nearly as convincing as they would be if they were all drawn from the same lexical category.

Next, let's revisit the data from (2), repeated in (38). These are taken from a section of the source work titled, "Tonal Contrasts." Although the authors (hereafter A-P&K) don't specifically state whether these contrasts are phonemic or underlying, the paper is written from a generative perspective, and it is clear that A-P&K are attempting to establish the underlying contrasts that serve as inputs to the generative rules they discuss later.

(38) Buli (Akanlig-Pare & Kenstowicz 2003)

	H		M		L	
CV	lé	spinster	lō	fall	mà	mother
CV:	mí má:	I helped	mā:	help!	wà mà: mǎ	he helped me
CVC	zúk	head	bāŋ	bangle	bàŋ	lizard
CV:C	bí:k	child	bū:k ^w	goat	nà:b	chief

First, let it be known that I am not contesting the conclusion that there are three underlying tone heights in Buli; I have no reason to suspect there are not. I am, however, contesting the choice of data employed to support these claims. From a generative perspective, the data in (38) are simply not convincing for a variety of reasons, and an underlying three-height contrast is not adequately established on any given row.

From the foregoing discussions, it should be clear why comparing nouns ('spinster' and 'mother') with the verb ('fall'), on the CV row, is unacceptable: there is a strong possibility that their phonological environments are not comparable because their grammatical environments are not comparable. The mid pitch on the verb 'fall' could easily be a level that is derived from interaction of the root tone with a floating tone verbal affix. On the CV: row, contrast is proposed between: a) a subject-verb (present tense) construction, b) an imperative verb, and c) a subject-verb-object (past tense) construction. Moreover, the three "contrastive" levels are all found *on the same root*, with the different pitch heights clearly attributable to influence from adjacent words. Logically, this means that at least two of these levels (and we can't know which two) must be postlexically derived. The CV and CV: rows therefore do not establish an underlying three-height contrast.

With regard to the remaining two rows, all of the forms are nouns spoken in isolation. Data like these are much better for establishing underlying contrast, but again there is a problem because Buli has noun classes. As demonstrated above, noun class affixation often consists of floating tones, meaning that one or more of these nouns may have (non-transparent) affixation that differs from that of the other nouns. In the CVC row, the nouns denote a) a body part, b) a body ornament, and c) an animal. It is very possible that these nouns belong to three different classes. With respect to the CV:C row, 'child' and 'chief' often belong to separate noun classes in Niger-Congo languages, and 'goat' and 'chief' most certainly belong to different classes. Again, differences like these raise the possibility that one or more of the levels in each of these rows could be a derived level. In short, the data in (38) support phonemic, as opposed to underlying, contrasts.

Since phonemic contrasts can be based on both derived and nonderived data (and one seldom knows which is which), what can we learn from the forms in (38)? Very little actually. We are unable to confirm that the language has three underlying tone heights, and we cannot assign any particular underlying tone pattern to any particular morpheme.

6. COMPARABLE DATA IN NJYEM [NJY]. Having critiqued different sets of surface tone contrasts based on data that are non-comparable from a generative perspective, let's turn our attention to some contrasts that are based on comparable data. The data presented in (39) are from Njyem, a Bantu A language spoken in eastern Cameroon and Congo (data from Keith Beavon, personal communication).

(39) Surface contrastive tone patterns of Njyem Class 7 nouns

CV		CVC		CVCV	
[\]		[\]		[- \]	
ba	'piece of bark'	ler	'bat-wing'	baɲɔ	'dry season'
[-]		[-]		[- -]	
li	'tree'	ṽjim	'cry (n.)'	baha	'overhanging rock'
[\]		[\]		[- -]	
ḍzo	'sleep (n.)'	lam	'trap (n.)'	lima	'dream (n.)'
[-]		[-]		[- -]	
gu	'madman'	ḍjim	'bad luck'	dila	'burial'

(39) is a truly beautiful example of comparable tone data. From a purely segmental point of view, although the consonants and vowels differ among the words, the segmental differences are such that they do not affect the surface pitches differently. We do know that different syllable profiles affect underlying tone patterns differently in tone languages, but this is controlled for in these data. All data in any one column have an identical syllable profile, i.e., they have the same number of TBUs to which tones can associate. Notice also that the codas of the CVC column share the same specification for sonorancy, in this case [+sonorant]. From a syntactic point of view, all data are single word noun phrases, pronounced in isolation. From a morphological point of view, all words are nouns with simple stems (single roots only) that belong to the same noun class, Class 7, a fact established by study of the language's concord system (cf. Beavon & Beavon 2006). Knowing this means that whatever floating affixes (if any) these nouns may have, they will all be the same. In other words, one can be sure that the nouns in (39) are all morphologically identical except for their roots. Since the syntactic and morphological environments of the data being compared are all identical, the researcher can be confident that all surface tone differences between comparable data (i.e., all data belonging to the same column) represent underlying phonological contrasts.

From these data, one can see that for each root syllable profile, there is a four-way surface contrast. A phonemic analysis would conclude that there were four phonemic tones: low-falling, high, high-falling, and low, but provide no explanation for why high-falling

was missing from CVCV forms or why low-falling was missing from the first syllable of CVCV forms. At the end of Hyman's stage 1, a generative analysis would conclude that there were four contrastive surface tone patterns. After eliciting short phrases and identifying tonal alternations (Hyman's stage 2), it would conclude (Hyman's stage 3) that the four contrastive surface patterns were /L/, /H/, /HL/, and /LH/ (see discussion immediately below).

Explaining the underlying patterns of the first three rows is reasonably transparent. In the case of the first two rows, we can posit /L/ and /H/ patterns, respectively; the respective tones associate to each TBU in the word, regardless of whether there are one or two TBUs present. When the tone is low, it is phonetically realized as low-falling pitch when it occurs before pause, and in this respect it is little different from that of most other African languages. In the case of the third row, there is no difficulty in proposing the underlying pattern /HL/. In this case, both tones associate to the single TBU of the CV column, and this results in a phonetically high-falling pitch. With respect to the CVC column, the two tones associate to the two moras of the syllable, the nucleus and the coda. This again results in a high-low falling pitch over the course of the syllable. When the two moras belong to two different syllables, as in the CVCV column, the high tone associates to the nucleus of the first syllable, and the low tone to that of the second syllable. This results in the high followed by low surface pitch pattern.

The underlying pattern for the fourth row is not as straightforward as the first three, since proper analysis of this row depends on looking at additional data (Hyman's stage 2). First, let's look at words from the first row of (39), which are assigned the underlying tone /L/, and compare them with their phrase-medial counterparts in (40) when they are followed by the demonstrative *jà* 'that'. Like the words of the fourth row, *jà* is also realized with a low tone that does not fall when it occurs in pre-pausal environments (discussed further below).

(40) /L/ when followed by demonstrative *jà*

CV	$\left[\begin{array}{c} \backslash \\ \end{array} \right]$		$\left[\begin{array}{c} - - \\ \end{array} \right]$	
	ba	'piece of bark'	[ba ja]	'that piece of bark'
CVC	$\left[\begin{array}{c} \backslash \\ \end{array} \right]$		$\left[\begin{array}{c} - - \\ \end{array} \right]$	
	ler	'bat-wing'	[ler ja]	'that bat-wing'
CVCV	$\left[\begin{array}{c} - \backslash \\ \end{array} \right]$		$\left[\begin{array}{c} - - - \\ \end{array} \right]$	
	baŋɔ	'dry season'	[baŋɔ ja]	'that dry season'

When words that are assigned an underlying low tone occur phrase medially before another low tone, they are realized with a low, level pitch. This is not totally unexpected if they are

indeed underlyingly /L/. Now compare these words with their fourth row counterparts in (41).

(41) Fourth row pattern, when followed by demonstrative ‘that’

CV	$\left[\begin{array}{c} - \\ \text{gʊ} \end{array} \right]$	‘madman’	$\left[\begin{array}{c} / - \\ \text{[gɔ ja]} \end{array} \right]$	‘that madman’
CVC	$\left[\begin{array}{c} - \\ \widehat{\text{dʒim}} \end{array} \right]$	‘bad luck’	$\left[\begin{array}{c} / - \\ \widehat{\text{[dʒim ja]}} \end{array} \right]$	‘that bad luck’
CVCV	$\left[\begin{array}{c} - - \\ \text{dɪla} \end{array} \right]$	‘burial’	$\left[\begin{array}{c} - - \\ \text{[dɪla ja]} \end{array} \right]$	‘that burial’

In (41), when ‘madman’ and ‘bad luck’ occur phrase medially before a low tone, they are each realized with a low-rising pitch, and when their disyllabic counterpart occurs before a low tone, it is realized with a low followed by high pitch. This leads us to conclude that the underlying pattern for the fourth row is /LH/. Following application of Hyman’s stages 2 and 3, the conclusions are summarized in the row headings of (42).

(42) Underlying tone patterns of Njyem Class 7 nouns

	CV		CVC		CVCV	
/L/	$\left[\begin{array}{c} \setminus \\ \text{ba} \end{array} \right]$	‘piece of bark’	$\left[\begin{array}{c} \setminus \\ \text{ler} \end{array} \right]$	‘bat-wing’	$\left[\begin{array}{c} - \setminus \\ \text{baɲɔ} \end{array} \right]$	‘dry season’
/H/	$\left[\begin{array}{c} - \\ \text{li} \end{array} \right]$	‘tree’	$\left[\begin{array}{c} - \\ \widehat{\text{tʃim}} \end{array} \right]$	‘cry (n.)’	$\left[\begin{array}{c} - - \\ \text{baha} \end{array} \right]$	‘overhanging rock’
/HL/	$\left[\begin{array}{c} \setminus \\ \widehat{\text{dzo}} \end{array} \right]$	‘sleep (n.)’	$\left[\begin{array}{c} \setminus \\ \text{lam} \end{array} \right]$	‘trap (n.)’	$\left[\begin{array}{c} - - \\ \text{lima} \end{array} \right]$	‘dream (n.)’
/LH/	$\left[\begin{array}{c} - \\ \text{gʊ} \end{array} \right]$	‘madman’	$\left[\begin{array}{c} - \\ \widehat{\text{dʒim}} \end{array} \right]$	‘bad luck’	$\left[\begin{array}{c} - - \\ \text{dɪla} \end{array} \right]$	‘burial’

Due to the lack of minimal pairs, the data in (42) may not be as esthetically pleasing as some might like, but because the contrasts are based on genuinely comparable data, they are truly beautiful and readily lend themselves to a convincing generative analysis.

Finding adequate comparable data to support underlying segmental contrasts is normally not a serious problem, and the same holds true for supporting tonal ones. When underlying tonal contrasts are genuine, there is normally no problem finding adequate comparable data to support them. So why do so many linguists fail when it comes to establishing support for them? Will Leben (personal communication) suggests this may be because linguists' notions of "what segments can and can't do are better developed [than their notions of what tones can and can't do], thanks to more exposure to both theory and practice." This paper tries to rectify this state of affairs by establishing better "best practices" for tone analysis.

APPENDIX:

FACTORS THAT AFFECT THE PHONOLOGICAL COMPARABILITY OF TONAL CONTRASTS

In order to help ensure that tonal contrasts are based on phonologically comparable data, some readers may find helpful the following list of phonological and grammatical factors that need to be the same for all tonal data being compared. I also highly recommend the phonology search and data management tool *Dekereke*, currently being developed by Rod Casali (rod_casali@sil.org), as one of the most useful software applications I am aware of to help control for these factors.

PHONOLOGICAL FACTORS

- a. prosodic domain (e.g., phonological word, phonological phrase), in all languages. This can be ensured by strictly controlling for grammatical factors.
- b. syllable profile of the morpheme, in all languages. The term syllable profile refers to both the number of TBUs in the morpheme as well as the sonorancy of any codas.
- c. adjacent tone patterns from other morphemes. This can be helped by strictly controlling for grammatical factors.
- d. depressor/raiser qualities of consonants, in language families where consonant-tone interaction is known to occur (e.g., Songjiang dialect of Wu, discussed above).
- e. phonation quality of vowels (e.g., breathy, creaky, glottalized, etc.), in language families where phonation quality is known to affect tone (e.g., Jingpho [kac], a Tibeto-Burman language spoken in Myanmar (Maddieson & Hess 1986).
- f. syllable stress patterns, in language families known to have mixed tone-stress systems (e.g., Iquito [iqu], a Zaparoan language spoken in northern Peru (Michael 2011).

GRAMMATICAL FACTORS

- a. lexical category of stem (e.g., noun, verb).
- b. stem type (simple, complex, compound, or borrowed).
- c. word class (e.g., noun class, verb class), in languages known to have these.
- d. any grammatical categories that involve affixation, whether segmental or otherwise (e.g., person, number, tense, aspect, derivation).
- e. syntactic environment (e.g., isolation, verb-object construction).

REFERENCES

- Akanlig-Pare, George & Michael Kenstowicz. 2003. Tone in Buli. *Studies in African Linguistics* 31. 55-96.
- Bao, Zhiming. 1990. *On the nature of tone*. Cambridge, MA: MIT dissertation.
- Beavon, Keith & Mary Beavon. 2006. *Une orthographe proposée pour la langue njiem*. SIL unpublished ms.
- Bessem, Bate Beatrice. 1987. *The noun class system of Bakwere*. Yaoundé, Cameroon: University of Yaoundé I thesis.
- Chen, Matthew. 2000. *Tone sandhi: Patterns across Chinese dialects*. Cambridge: Cambridge University Press.
- Coupe, Alexander R. This volume. Strategies for analyzing tone languages. *Language Documentation and Conservation* 8. 462-489
- Hedinger, Robert. 2008. *A grammar of Akoose: A northwest Bantu language*. Dallas: The Summer Institute of Linguistics and the University of Texas at Arlington.
- Hyman, Larry M. 1985. Word domains and downstep in Bamileke-Dschang. *Phonology Yearbook* 2. 85-138.
- Hyman, Larry M. This volume. How to study a tone language, with exemplification from Oku (Grassfields Bantu, Cameroon). *Language Documentation and Culture* 8. 525-562
- Hyman, Larry M. & Maurice Tadjadjeu. 1976. Floating tones in Mbam-Nkam. In Larry M. Hyman (ed.), *Studies in Bantu tonology* (Southern California Occasional Papers in Linguistics 3), 57-111. Los Angeles, CA: University of Southern California.
- Jamieson, Allan R. 1977. Chiquihuitlan Mazatec tone. In William Merrifield (ed.), *Studies in Otomanguean phonology*, 107-136. Dallas: Summer Institute of Linguistics.
- Jiangsusheng he Shanghai Shi Fangyan Gaikuang* [A survey of Jiangsu Province and Metropolitan Shanghai]. 1960. Nanjing: Jiangsu Renmin Chubanshe.
- Leroy, Jacqueline. 1980. The Ngemba group: Mankon, Bangangu, Mundum I, Bafut, Nkwen, Bambui, Pinyin, Awing. In Larry Hyman & Jan Voorhoeve (ed.), *L'expansion bantoue*, vol. 1, 111-141. Paris: SELAF.
- Maddieson, Ian & Susan Hess. 1986. 'Tense' and 'lax' revisited: more on phonation type and pitch in minority languages of China. *UCLA Working Papers in Phonetics* 63. 103-109.
- Mazaudon, Martine. This volume. Studying emergent tone systems in Nepal: Pitch, phonation and word-tone in Tamang. *Language Documentation and Conservation* 8. 587-612

- Michael, Lev. 2011. The Interaction of tone and stress in the prosodic system of Iquito (Zaparoan). *UC Berkeley Phonology Lab Annual Report (2010)*. 57-79.
- Parker, Elizabeth. 1989. Le nom et le syntagme nominal en mundani. In Daniel Barreteau and Robert Hedinger (eds.), *Descriptions de langues camerounaises*, 131-177. Paris: ACCT & ORSTOM.
- Rice, Keren. This volume. On beginning the study of the tone system of a Dene (Athabascan) language: Looking back. *Language Documentation and Conservation* 8. 690-706
- Roberts, James S. 1994. Nontonal floating features as grammatical morphemes. *Work Papers of the Summer Institute of Linguistics, University of North Dakota Session* 38. 87-100.
- Snider, Keith. 1988. The noun class system of proto-Guang and its implications for internal classification. *Journal of African Languages and Linguistics* 10(2). 137-163.
- Snider, Keith. 1999. *The geometry and features of tone*. Dallas: The Summer Institute of Linguistics & The University of Texas at Arlington.
- Snider, Keith. 2013. *Tone analysis for field linguists*. SIL unpublished ms.

Keith Snider
keith_snider1@sil.org