In *Texas linguistic forum 41: Exploring the boundaries between phonetics and phonology*. A. Doran, T. Majors, C. E. Mauk, and N. M. Goss (eds.), 57-69.

(H)igh-(H)igh, (L)ow-(L)ow, What's-What?[1]

Chip Gerfen
UNC-Chapel Hill

Paul Denisowski
UNC-Chapel Hill

1. Introduction

This paper examines the phonetic implementation of lexically contrastive tones in Coatzospan Mixtec (henceforth CM), an Otomanguean language spoken by roughly 2000 people (Small, 1990) in northern Oaxaca, Mexico. First described by Pike and Small (1974), CM is reported to have an extremely complex system of tone sandhi, together with a pattern of lexically induced downstep similar to the downstep systems of West African languages (see, for example, Welmers, 1959; Stewart, 1964; Clements, 1979). To date, however, there has been no acoustic investigation of the implementation of CM tones. One of the goals of this study is thus to provide a preliminary examination of how lexical tones are implemented in CM words that are uttered in isolation as a means of laying the groundwork for further study of the sandhi and downstep phenomena.

As with other Mixtec languages, tone is lexically contrastive in CM morphemes (known as *couplets* in the Mixtec literature). Couplets take two canonical shapes: (C)CVV and (C)CVCV and exhibit five basic tone patterns, as seen in (1).

(1)		**(C)CVV**			**(C)CVCV**	
	a.	H_H	skẃíí	'spotted'	kíní	'will see'
	b.	L_L	kẁìì	'green'	ʃkìnì	'last night'
	c.	L_H	ŋgẁìí	'remainder'	mìní	'lake'
	d.	L_LH	nɗìǐ	'sweat'	rkùnǔ	'board'
	e.	H_LH	ʃíǐ	'delicate'	ʃúmě	'wax'

[1] We are grateful to audiences at UNC-Chapel Hill and at the Texas Linguistics Society conference for feedback on the material under discussion here. In addition, we are especially grateful to Njau Piña for his friendship, expertise, and help, without which this research could not have been undertaken. All errors are, of course, ours alone.

Here, we focus on the three most prevalent patterns: (H)igh_(H)igh, (L)ow_(L)ow, and (L)ow_(H)igh. Besides being the most common, these offer an interesting point of departure for a study of CM tone, given Pike and Small's impressionistic claim that in isolation forms, H_H and L_L are phonetically indistinguishable, while a phonological distinction between H_H and L_L tone specifications is nevertheless crucial to understanding the behavior of these forms in context, i.e. in combination with other morphemes. In fact, the putative intricacy of the sandhi and downstep facts led Pike and Small to posit roughly 11 morpheme classes and 18 rules in an attempt to account for the data.

The organization of this paper is as follows. In §2, we discuss data collection, experimental set-up, and our choices regarding acoustic measurements. In §3, we discuss our results, focusing in particular on the evidence which supports Pike and Small's impression that phonological H_H and L_L are phonetically non-distinct when uttered in isolation. In §4, we conclude with a discussion of the relevance of our findings to the phonetics/phonology interface, arguing that the data here indicate that it remains useful to maintain a more abstract, categorical level of phonological representation that is interpreted by a phonetic component of the grammar.

2. Data Collection, Experimental Set-up, and Measurement

All of the data were collected in the field in San Juan Coatzospan by the first author during the summer of 1997. Experimental items were selected as near minimal pairs from the three tone classes H_H, L_L, and L_H for both (C)CVV and (C)CVCV couplets. We report here on data from six speakers (ages 24-42), four females and two males, all native speakers who use CM as their primary mode of communication. The experimental items are shown in (2). All were controlled for stress: (C)CVV couplets are stressed, and in (C)CVCV couplets, stress falls predictably on the first vowel.

(2) Experimental items

		(C)CVV		**(C)CVCV**	
a.	H_H	sk^{w}íí	'spotted'	kíní	'will see'
b.	L_L	k^{w}ìì	'green'	ʃkìnì	'last night'
c.	L_H	ŋg^{w}ìí	'remainder'	mìní	'lake'

These items were embedded in a list of 72 words that was randomized six times and presented to subjects in PowerPoint on a PowerBook. Though CM lacks a standard orthography, the words were transcribed in Mixtec, with tones marked. A smaller Spanish translation of each word was also provided below the CM word. Additionally, a highly literate CM speaker was present at all times to assist speakers whose reading skills were insufficient to carry out the task with ease. Speakers were first familiarized with each form in the list to guarantee that they accepted the form as a word. They were told that they would be reading words from their language out loud and were asked to speak in a natural but clear fashion. They pressed the mouse to proceed from one word to the next and were free to correct themselves. They were told only that they were participating in a study of how Mixtecs speak.

The data were recorded on a Marantz PMD 222 portable cassette recorder, with a Shure unidirectional, dynamic, head-worn, close-talking microphone, and later digitized at 22kHz, segmented, and analyzed in SoundScope. For each vowel (which we took to be the long V: in (C)CVV-type words and each V in (C)CVCV words), an f0 contour was generated via autocorrelation, with data points sampled at 5 msec intervals. Each f0 contour was fit to a second order polynomial via a least means squared method. From the curves, we extracted a series of measures (e.g. peak f0, min f0, mean f0, DN1 and so forth) for comparison.

Two issues arise here. First, why did we fit a polynomial to our f0 contours? And second, what measures should be taken when examining how tones are implemented? Regarding the first question, a fitted polynomial provides a mathematically sound way of smoothing the raw f0 contours. From our perspective, this has a number of advantages. First, it eliminates the possibility of outliers being taken at peak and minimum points, and secondly, it allows us to compare rates of change in f0 by comparing the first derivatives (DN1) of the curves derived from the different f0 contours. Finally, using fitted curves removes the experimenter from the process of selecting particular points for measurement and comparison, as we automated the entire process in MATLAB, thus ensuring the replicability of the study.

This leads to the second question: how should we measure tones? That is, what acoustic measure or measures afford the best characterization of how tones are implemented phonetically? Different researchers have adopted different strategies when conducting f0 studies. One strategy, for example, is to measure the peak f0 value of a high tone or a minimum value in the case of a low tone (e.g. Pierrehumbert and Beckman, 1988; Liberman et. al., 1993; Myers, 1998). Other possibilities include calculating the mean f0 across the span of a vowel and the standard deviation from the mean (see, for example, Kirk et. al., 1984 on laryngealization). Yet another is to measure the slope of the tone, i.e. the rate of change in f0 throughout the duration of a vowel. Finally, the question also arises of how many measurements to take. Is one f0 measurement per vowel sufficient (Pierrehumbert and Beckman, 1988; Liberman et. al., 1993; Myers, 1998), or should some arbitrary number of measurements be made at different points in a vowel (Laniran, 1992; Xu, 1997)?

We doubt there is a single answer to these questions, given that our own results here indicate that different languages will implement the "same" lexical tones in different ways, and given that the most appropriate measure or measures to take will depend on the research question at hand. In the case of CM, we part from a base in which there are no extant phonetic data regarding tone implementation. As a consequence, we used our curves to extract a number of potentially relevant dimensions in order to probe how lexical tones are implemented in the language. Bearing this in mind, we turn in §3 to a discussion of our results.

3. Experimental Results

For simplicity, we divide §3 into two experiments--one on (C)CVV forms and another on (C)CVCV forms--although all of the data were collected in a single recording session, as described in §2 above.

3.1. Experiment 1: (C)CVV forms

As noted, we treat the long vowel in (C)CVV forms as a single vowel gesture. For each vowel, an f0 contour was generated and fit to a second order polynomial. We then automated the extraction in MATLAB of nine potentially relevant measures: peak f0; valley (minimum) f0, mean f0, standard deviation of the mean, the first derivative (DN1) of the curve, the duration in data points of the curve, the % of the duration of the vowel where the peak is reached, the % of the vowel where the valley is reached, and the maximum deflection, i.e. the difference between the peak and the minimum f0.

These measures were used to compare the curves resulting from different lexical tone specifications. To afford consistency with our examination of (C)CVCV forms, we coded the three lexical tone patterns (under the variable name *word-type*) as either H_H, L_L, or L_H. Note, however, that this does not reflect assumptions regarding phonological representation in that we do not take H_H forms such as sk^{w}íí 'spotted' to consist of two adjacent tokens of H tones or of two adjacent /i/ vowels. Rather, we assume that such forms consist phonologically of a single long vowel associated to a single H tone, in accordance with standard assumptions regarding the representation of long vowels as bimoraic (Hyman, 1985) and with the OCP (Leben, 1993). Referring, however, to H_H, L_L, and L_H as word-types simply provides us with a convenient way of encoding the tone pattern of both (C)CVV and (C)CVCV forms.

Using the codes for the three word-types, H_H, L_L, and L_H, we hypothesized that the lexical tone pattern of a couplet should significantly affect a number of our phonetic measures. For example, we expected that H_H forms would show a higher peak f0 value than that of L_L and that L_H would have a greater maximum deflection than either L_L or H_H. Generally speaking, we assumed the null hypothesis in that we hypothesized that our phonologically distinct tone patterns would emerge as phonetically distinct along some relevant phonetic dimensions. Anticipating our results, our hypotheses were only partially confirmed. While L_H forms emerged as clearly phonetically distinct from both their L_L or H_H counterparts, we found no evidence that L_L and H_H forms are phonetically distinct in (C)CVV forms. This said, we turn to the particulars of our comparisons.

To determine whether the lexical tone pattern significantly affected our various measures, we ran a battery of 2 Factor ANOVAs ($p = .05$). We treated our subject pool as the population for the study and treated each repeated token as a sample of a particular tone pattern. In the ANOVAs, each measure (e.g. peak f0, valley f0, etc.) served as the dependent variable, with word-type (H_H, L_L, and L_H) and subject used as the independent variables. When we found a significant effect for word-type, we ran Sheffe's post-hoc tests to further explore how the word-type variable was producing the main effect. The table in (3) provides a summary of the ANOVA tests for (C)CVV forms by word-type.

(3) ANOVA results for (C)CVV forms by word-type

	P-VAL	**F-VAL**	**Sum of Sq.**	**Mean Sq.**
peak	.0001	10.895	2029.735	1014.868
valley	.6936	.367	60.371	30.185
mean	.649	.434	63.908	31.954
sd	.0001	17.119	123.637	61.818
DN1	.0001	38.554	0.039	0.019
dur	.0001	47.286	4100.056	2050.028
% to peak	.0001	690.199	19.25	9.625
% to min	.0001	1489.045	14.948	7.474
deflection	.0001	17.402	1415.055	707.527

(df: word-type (2), subject (5), word-type * subject (10), residual (90))

Only the mean and minimum f0 measures show no significant effect for word-type. For the other variables, we conducted post-hocs, which are summarized in (4).

(4) Summary of p values for Scheffe's tests (p < .05)

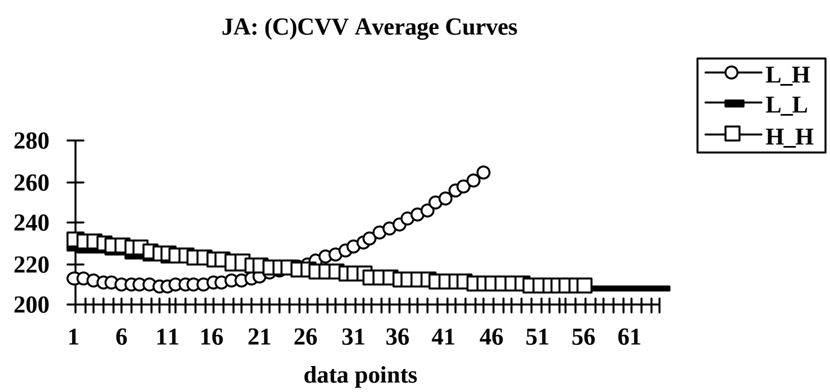
	Vs.	**peak**	**sd**	**DN1**
L_L	**H_H**	.6004	.6502	.9848
	L_H	.0001	.0001	.0001
H_H	**L_H**	.0039	.0001	.0001

	Vs.	**dur**	**% to peak**	**%to min**	**defl**
L_L	**H_H**	.0811	.9996	.7315	.6305
	L_H	.0001	.0001	.0001	.0001
H_H	**L_H**	.0001	.0001	.0001	.0001

Strikingly, the tests indicate that the lexical L_H specification alone drives the significant results. This is noteworthy for two reasons. First, there are no measures in which L_L and H_H emerge as significantly different from one another. Second, L_H is significantly different from both H_H and L_L for every measure.

On examining the shape of the curves, three generalizations emerge. First, the L_H contour exhibits an overall f0 rise, while the L_L and H_H contours fall. This difference underlies the post-hoc results for the %-to-peak and %-to-min variables. Second, L_H's f0 peak is higher than those of either H_H or L_L forms. And third, L_H exhibits a greater rate of change in f0 than that found in L_L and H_H forms, a fact captured by the results of the post-hocs on the standard deviation of the mean f0, the deflection, and the first derivative of the fitted curves. Note as well that the sharpness of the rise is also enhanced by the shorter duration of the vowel in the L_H word-type. The graph in (5) for the female speaker JA exemplifies these generalizations. An average curve of the six repetitions of each lexical tone pattern is provided. Note the overlap of the H_H and L_L patterns, which both decline gradually. By contrast, the L_H contour begins at a lower f0 and curves sharply upwards.

(5) Comparison of the three patterns



In short, the three lexical tone patterns are implemented phonetically in (C)CVV forms as if there were a distinction between only two: 1) a longer, flatter pattern exhibiting a gradual f0 decline (H_H or L_L) and 2) a shorter, rising pattern (L_H).

3.2. Experiment 2: (C)CVCV forms

We divide our discussion of (C)CVCV forms into two parts, focusing first on a comparison of V1 across the different lexical tone patterns and then turning to a consideration of the f0 contours for V2. The results of the ANOVAs reveal a different pattern for V1 in (C)CVCV than for the long vowel in (C)CVV forms, as seen in (6).

(6) ANOVA results for V1 of (C)CVCV forms

	P-Val	**F-Val**	**Sum of Sq.**	**Mean Sq.**
peak	0.0001	43.621	5150.016	2575.008
valley	0.0001	17.38	1775.941	887.97
mean	0.0001	37.569	3952.098	1976.049
sd	0.0001	27.133	87.101	43.55
DN1	0.0001	15.79	0.013	0.007
dur	0.8665	0.143	5.13	2.565
% to peak	0.1426	1.99	0.123	0.062
% to min	0.2205	1.537	0.122	0.061
deflection	0.0001	27.703	946.632	473.316

(df: word-type (2), subject (5), word-type * subject (10), residual (90))

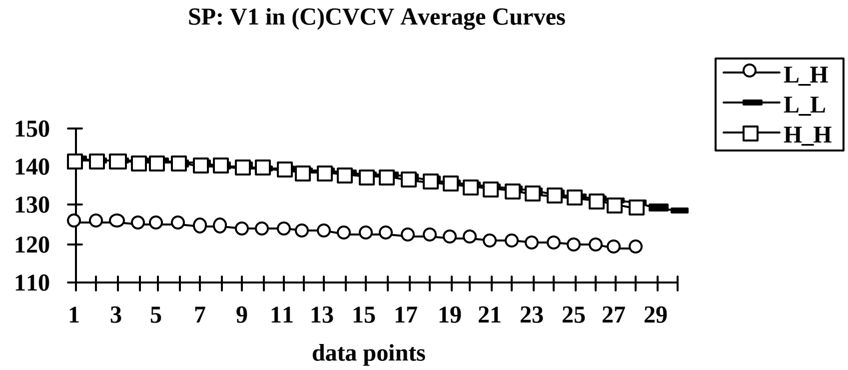
In particular, the duration, %-to-peak, and %-to-min variables were not significantly different across word-types, though all of the other measures are significantly affected by word-type. For those measures significantly affected by word-type, however, the same general pattern emerges as in the (C)CVV forms. That is, the post-hoc tests reveal that L_L and H_H are never significantly different, while L_H is always significantly different from both H_H and L_L.

(7) Summary of p values for Scheffe's tests for V1 ($p < .05$)

	Vs.	**peak**	**valley**	**mean**	**s.d.**	**DN1**	**deflection**
L_L	**H_H**	.1918	.1369	.1027	.9690	.210	.9970
	L_H	.0001	.0013	.0001	.0001	.0001	.0001
H_H	**L_H**	.0001	.0001	.0001	.0001	.0016	.0001

More specifically, we are left with the following observations. The L tone of V1 in the L_H forms has the lowest f0 peak, valley, mean, and standard deviation/deflection, and it has the flattest contour. At the same time, the H of the H_H forms is not significantly different from the L of L_L forms for any of these measures, and both H in H_H and L in L_L exhibit more of a decline in f0 throughout the duration of V1. This is exemplified in (8), which provides an average curve for the six tokens produced by the male speaker SP for each tone pattern.

(8) Comparison of the three patterns for V1



Let us now consider the f0 contours for V2 across the three tone classes. A table summarizing the results of the ANOVAs by word-type for V2 is provided in (9). As can be seen in the table, the standard deviation of the mean f0, the first derivative, and the f0 deflection are not significantly affected by the lexical tone pattern of the form.

(9) ANOVA results for V2 of (C)CVCV forms

	P-Val	**F-Val**	**Sum of Sq.**	**Mean Sq.**
peak	**0.0001**	73.412	10417.563	5208.782
min	**0.0001**	53.119	8877.405	4438.702
mean	**0.0001**	74.556	9950.141	4975.07
s.d.	**0.4164**	0.885	4.503	2.251
DN1	**0.6369**	0.453	0.046	0.023
dur	**0.004**	5.877	69.5	34.75
%to peak	**0.0001**	37.909	4.945	2.472
%to min	**0.0001**	49.116	9.727	4.864
deflection	**0.2627**	1.357	66.05	33.025

(df: word-type (2), subject (5), word-type * subject (10), residual (90))

Though the rate of f0 change is not significantly different across word-types, the peak and minimum f0 values, as well as %-to-peak and %-to-min are significantly affected by the lexical tone pattern of the word. The post-hocs again reveal the same general story. With the exception of the duration variable, L_H is systematically different from H_H and L_L for each measure, while H_H and L_L are again statistically non-distinct.[2] This is shown in (10).

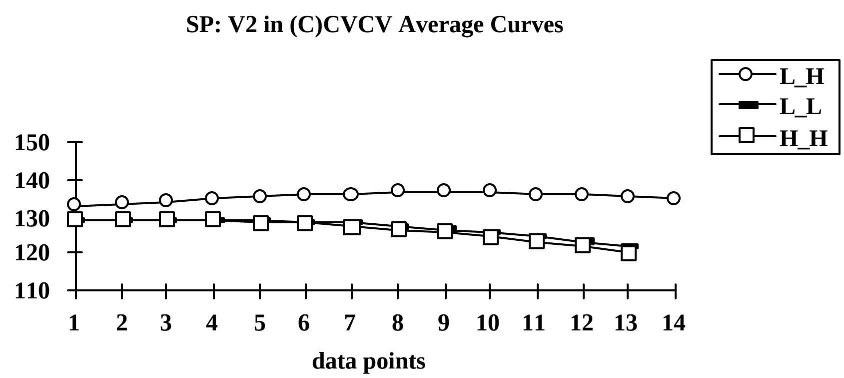
[2] Note in (10) that H_H and L_H are significantly different in the duration of V2. However, it is useful to note that the mean V2 duration (across subjects) in L_H is 14.222 data points, while that of V2 in H_H forms is 12.306. This difference is thus on the order of 10 milliseconds for two unstressed vowels which are themselves much

(10) Summary of p values for Scheffe's tests for V2 ($p < .05$)

	Vs.	peak	valley	mean	dur	%to peak	%to min
L_L	**H_H**	0.0928	0.2572	0.1397	0.5975	0.3875	0.9828
	L_H	0.0001	0.0001	0.0001	0.0722	0.0001	0.0001
H_H	**L_H**	0.0001	0.0001	0.0001	0.0051	0.0001	0.0001

The overall picture that emerges for V2 is thus the following. In the L_H word-type, V2 has higher peak, mean, and minimum f0 values than those found for V2 in H_H and L_L couplets. In L_H forms, there is a general rise to the peak f0 value, while in H_H and L_L, the peak is reached earlier and is followed by an f0 decline. The lack of a significant effect for the first derivative indicates that in V2, f0 targets, together with the direction of f0 changes, are marking phonological differences more than the mere rate of change itself in the f0 contour. To illustrate this, average f0 curves of V2 are provided for SP in (11) for each word-type.

(11) Comparison of the three patterns for V2



4. Conclusions: Implications for Phonetics and Phonology

Our statistical results confirm Pike and Small's (1974) original observation that H_H and L_L forms are not distinguishable in isolation, i.e., that the motivation for these lexical tone patterns only emerges in context, where tone sandhi and downstep can be observed.[3] For our purposes here, what is notable is that this situation gives rise to a somewhat unexpected mismatch between the phonological specification of

---

shorter than their stressed V1 counterparts. We are thus hesitant to view this result as meaningful.

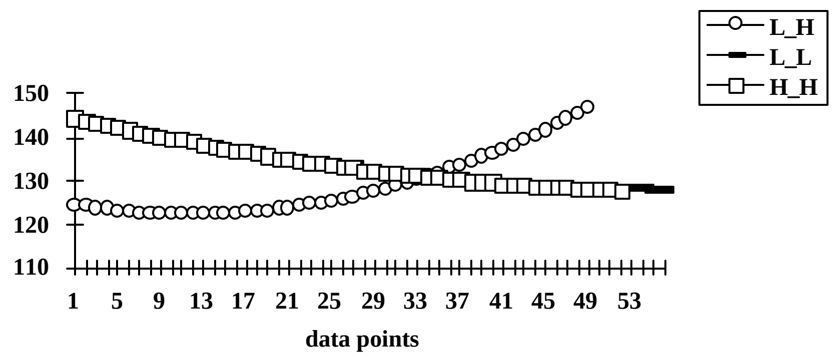
[3] Given the range of measurements compared as dependent variables, we feel confident that our results do indicate that L_L and H_H are implemented phonetically in the same fashion. However, there is, of course, always the possibility that we have missed some phonetic dimension that distinguishes between these tone patterns. Corroboration of our results should be sought via a perception study, which we hope to pursue in future research.

tones in CM and their phonetic implementation. As we note earlier, naively, one would expect H_H and L_L to be distinguishable in that H_H forms should at least be realized with a higher f0 peak. This is, in fact, the case in languages such as Yoruba (Laniran, 1992) and Igbo (cf. Liberman et. al., 1993; Gerfen and Laniran, 1997), where both patterns also display relatively flat contours but in which H_H forms are systematically realized as higher, even in disyllabic isolation forms.

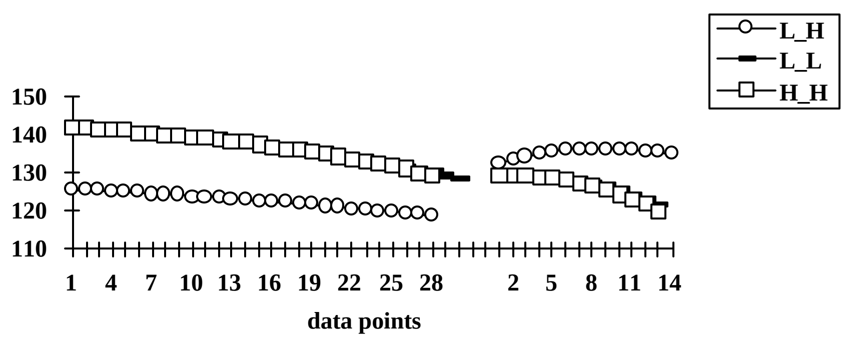
But if H_H and L_L contours are motivated phonologically, we are faced with the question of how these lexical specifications are to be phonetically realized. Consider the graphs in (12), with average curves for all three word-types for both (C)CVV and (C)CVV forms for SP.

(12) Comparison of the three patterns for both couplet types

SP: (C)CVV Average Curves

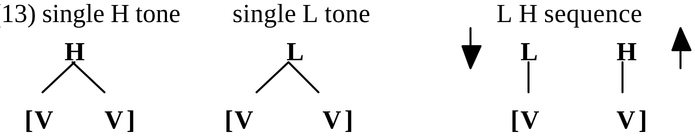


SP: (C)CVCV Average Curves



These contours suggest that in CM, H and L tones only receive distinct targets in cases in which speakers are required to implement a change in phonological tone value. One way to view this is to say that there is a phonetic neutralization of H and L in isolation forms containing only an H or L specification. Alternatively, we can say that extreme phonetic targets are only assigned to H and L when speakers must implement a contrast between the two lexically specified tones. In our data, only L_H forms provide such a context; i.e., only L_H forms contain the requisite shift in phonological tone. This generalization receives a straightforward characterization in terms of traditional autosegmental representations, such as those in (13).

(13) single H tone single L tone L H sequence



Note that two of the representations contain a single token of H and L, while the third contains adjacent tokens of L and H. We assume that this latter case triggers the dispersion of f0 targets for H and L in order to maximize the phonetic distance and thus express the phonological contrast between the two tones (see Liljencrants and Lindblom, 1972). When there is no sequence of distinct tones, more extreme targets are avoided, resulting in a phonetic mid tone for both the H and L specifications.

Two points are worth noting regarding this view. First, it assumes the existence of a categorical, symbolic level of phonological representation at which information regarding the distribution of lexical contrast is encoded. Secondly, it assumes a language particular phonetics which implements this symbolic information in time and space (cf. Pierrehumbert, 1980; Keating, 1988; Huffman, 1989; Cohn, 1993; Kingston and Diehl, 1994; Gerfen (forthcoming)). In this case, the interest lies in the unexpected fact that the phonetics does not implement L and H tones differently in couplets with a single lexical tone specification. That L and H do receive distinct targets only when they are adjacent is also interesting in that it provides further evidence for the local determination of f0 targets in context (cf. Liberman and Pierrehumbert ,1984 for English; Pierrehumbert and Beckman, 1988 for Japanese; Prieto et. al., 1995 for Spanish).

This rather traditional view of phonological representation provides insight into not only the dispersion of phonetic targets, but also, into the number of phonetic targets in each word-type. If we consider the examples for SP in (12) above, we notice important similarities between (C)CVV and (C)CVCV forms. In H_H and L_L contours, for example, regardless of couplet shape, the f0 contour reaches an early peak and gradually falls thereafter, either throughout the duration of the single long vowel in (C)CVV, or throughout both the first and second vowels in (C)CVCV.

In neither case do we see clear evidence for any more than one f0 peak target in the word. If we take our autosegmental representations seriously, this is arguably predicted by a representation containing only a single token of H or L, regardless of whether the lexical tone is linked to one long vowel or two distinct vowels. In simple terms, a single lexical tone should be interpreted as a single f0 target in the phonetics. By contrast, in L_H forms there is clear evidence of two targets in both (C)CVCV and (C)CVV forms: a low target corresponding to the L and a high target corresponding to the H. In (C)CVCV forms, each target is implemented on a distinct vowel, while in (C)CVV forms, both targets are implemented over the course of a single vowel gesture, thus generating the rising contour.

Finally, besides predicting the number of f0 targets per couplet, it is worth noting that our autosegmental representations provide potential insight into the relation between phonological representation and the general shapes of the f0 contours. In making these comparisons, we will focus on the contours of (C)CVV couplets and of V1 of (C)CVCV couplets, i.e. on stressed vowels implementing a lexically specified tone target. Obviously, (C)CVV forms specified for L_H exhibit the sharpest upward curve. But upon looking at the data, we noticed that there appeared to be an additional split between 1) flatter f0 contours for H_H and L_L in V1 of (C)CVV forms and for the L in (C)CVCV forms specified for L_H on the one hand and 2) a slightly more curved contour for H_H and L_L in V1 of (C)CVCV forms.

To examine whether our impressions were accurate, we recoded the data to compare V1 (defined as VV in (C)CVV forms and V1 in (C)CVCV forms) by creating an independent variable called tone&position. This yielded the six variables.

(14) V1 coded by tone&position

H1CVCV	= H1 of CVCV in H_H	HCVV	= H in CVV
L1CVCV	= L1 of CVCV in L_L	LCVV	= L in CVV
L_CVCV	= L of CVCV in L_H	LHCVV	= LH in CVV

An ANOVA on the first derivatives of V1 by tone&position and subject was significant for tone&position (p=.0001; df. tone&pos (5), subject (5, t&p*subj (25), residual (180)). And a Sheffe's test on the main effect yielded the following results.

(15) Summary of p values for Scheffe's tests for V1 by tone&position (p < .05)

	Vs.	Diff.	Crit. diff.	P-Value	
H1CVCV	**L1CVCV**	.009	.017	**.7150**	
	L_CVCV	.026	.017	**.0001**	S
	LCVV	.030	.017	**.0001**	S
	HCVV	.030	.017	**.0001**	S
	LHCVV	.070	.017	**.0001**	S
L1CVCV	**L_CVCV**	.018	.017	**.0339**	S
	LCVV	.021	.017	**.0060**	S
	HCVV	.022	.017	**.0034**	S
	LHCVV	.062	.017	**.0001**	S
L_CVCV	**LCVV**	.003	.017	**.9964**	
	HCVV	.004	.017	**.9880**	
	LHCVV	.044	.017	**.0001**	S
LCVV	**HCVV**	.001	.017	**1.0000**	
	LHCVV	.041	.017	**.0001**	S
HCVV	**LHCVV**	.040	.017	**.0001**	S

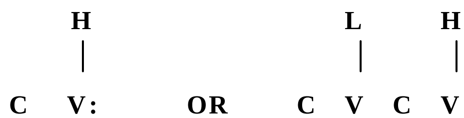
A more intuitive, composite view of the results is provided in (16), which contains pooled means (with standard deviations) of the first derivatives. The table reveals a three-way split among the first derivatives. The L_H contour in (C)CVV couplets is in a class by itself. The smallest first derivatives pertain to H_H and L_L implemented in CVV couplets as well as the L of V1 in (C)CVCV couplets specified for L_H. The third group contains the f0 contours for H_H or L_L in (C)CVCV.

(16) Pooled means for DN1

	CVV		V1 of CVCV	
H_H	**0.002**	**(0.017)**	**-.029**	**.(028)**
L_L	**0.001**	**(0.013)**	**-.020**	**(.019)**
L_H	**0.042**	**(0.054)**	**-.002**	**(.021)**

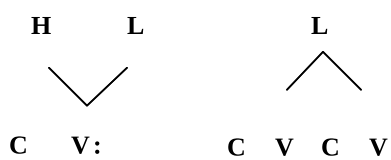
Note that autosegmental representations provide a suggestive account of this three-way split. In particular, the smallest first derivatives, i.e. the flattest curves, pertain to cases of a single vowel gesture bearing a single lexical tone, that is, to cases of the traditional one-to-one relation between vowels and tones, as seen in (17).

(17) One vowel to one tone



By contrast, the largest rate of change involves a single vowel gesture that bears two lexical tones--the traditional one-to-many relation between vowels and tones--found for (C)CVV forms with the L_H tone pattern, as seen in (18a). And the third group represents cases in which two vowels are linked to one lexical tone, i.e. cases of the many-to-one relation between anchors and tones. This occurs in (C)CVCV forms specified for what we have been calling the H_H or L_L tone pattern, but which we assume are represented as in (18b):

(18) a. b.



In sum, by abstracting away from the quantitative details of phonetic implementation and maintaining a symbolic level of phonological representation, we profit in at least three ways with respect to our understanding of how lexical tones are implemented in CM. First, such representations neatly encode the fact that the distinct H and L targets are only assigned when H and L autosegments are adjacent in the lexical representation of a form. Second, such representations predict that each token of a lexical tone will be interpreted in the phonetics as having one f0 target, and this prediction is supported by our results. Third, the consequences of the nature of autosegmental relations between anchors and targets for phonetic implementation--an issue to which phonologists have paid scant attention--are highlighted in an intriguing manner. Specifically, the rate of change of the f0 contours, at least for the data under examination here, correlates with the logical range of associative possibilities between anchors and features, suggesting that such relations are also interpreted by the phonetic component of the grammar.

References

Clements, G. N. 1979. The description of terraced-level tone languages. *Language* 55. 536-58.

Cohn, A. 1993. Nasalization in English: Phonology or Phonetics. *Phonology* 10. 43-81.

Gerfen, C. forthcoming. Coatzospan Mixtec: a study in phonology and phonetics. *Studies in Natural Language and Linguistic Theory*. Dordrecht: Kluwer.

Gerfen, C. and Y. Laniran. 1997. High raising in Igbo. UNC-Chapel Hill ms.

Huffman, M. K. 1989. Implementation of nasal: timing and articulatory landmarks. Ph.D. dissertation, UCLA.

Hyman, L. 1985. A theory of phonological weight. Dordrecht: Foris.

Keating, P. 1988. Underspecification in phonetics. *Phonology* 5. 275-292.

Kingston, J. and R. Diehl. 1994. Phonetic knowledge. *Language* 70. 419-454.

Kirk, P., P. Ladefoged, and J. Ladefoged. 1984. Using a spectrograph for measures of phonation types in a natural language. *UCLA Working Papers in Phonetics* 59. 102-113.

Laniran, Y. O. 1992. Intonation in tone languages: the phonetic implementation of tones in Yoruba. Ph.D. dissertation, Cornell U.

Leben, W. 1973. Suprasegmental phonology. Ph.D. dissertation, MIT.

Liberman, M. and J. B. Pierrehumbert. 1984. Intonational invariance under changes in pitch range and length. In *Lanaguage and Sound Structure*. M. Aronoff and R. Oehrle (eds.) Cambridge: MIT Press. 157-233.

Liberman, M., J. M. Schultz, S. Hong, and V. Okeke. 1993. The phonetic interpretation of tone in Igbo. *Phonetica* 50. 147-160.

Liljencrants, J. and B. Lindblom. 1972. Numerical simulation of vowel quality systems: the tole of perceptual contrast. *Language* 48. 839-862.

Myers, S. 1998. Surface underspecification of tone in Chichewa. U. of Texas at Austin ms.

Pierrehumbert, J. B. 1980. The Phonetics and phonology of English intonation. Ph.D. dissertation, MIT.

Pierrehumbert, J. B. and M. Beckman. 1988. Japanese tone structure. Cambridge: MIT Press.

Pike, E. and P. Small. 1974. Downstepping terrace tone in Coatzospan Mixtec. In *Advances in Tagmemics*. R. Brend (ed.) Amsterdam: North Holland. 105-134.

Prieto, P., J. van Santen, and J. Hirschberg. 1995. Tonal alignment patterns in Spanish. *Journal of Phonetics* 23. 429-451.

Small, P. 1990. A syntactic sketch of Coatzospan Mixtec. In *Studies in the Syntax of Mixtecan Languages* 2. C. H. Bradley and B. E. Hollenbach (eds.) Arlington, TX: The Summer Institute of Linguistics and U. of Texas at Arlington. 261-479.

Stewart, J. M. 1964. The typology of the Twi tone system. *Bulletin of the Institute of African Studies* 1. 1-67.

Welmers, W. E. 1959. Tonemics, morphotonemics, and tonal morphemes. *General Linguistics* 4. 1-9.

Xu, Y. 1997. Contextual tonal variations in Mandarin. *Journal of Phonetics* 25. 61-83.

Department of Linguistics
UNC-Chapel Hill
Chapel Hill, NC 27599
gerfen@email.unc.edu