Testing a Model for Hausa Intonation

Mona Lindau

This study reports a test of a formal model of the relationship between tone and intonation, using Hausa, a Chadic language spoken in Nigeria. This language has a relatively simple tone system of basically two tones, High and Low. The language also has a falling tone that can be analyzed as a combination of a High and a Low, and will not be considered further here (Meyers 1975, Cowan and Schuh 1976). The goal of this study is to test a model of intonation by matching actual fundamental frequency curves with curves generated by rule for Hausa statements and questions on two different tone patterns. The model is based on that developed at Lund in Sweden (Gårding and Bruce 1981, Gårding 1981). The results will also be useful in speech synthesis of Hausa utterances.

Figure 1 shows a model Hausa sentence. The input is a phonetic transcription with boundaries and tones marked. The model describes intonation as grids of baselines and toplines. In this description the grids are stored lexically as a basic slope for a particular number of syllables for each type of utterance, and are generated by rule. The boundary marks, and the Highs and Lows are inserted onto the grid lines. Local rules may modify the underlying patterns. The last step is to concatenate the maxima and minima into a smooth curve.

The model can be tested by the degree to which it will account for both tonal and intonational patterns used in common by different speakers. An analysis was made of Hausa utterances. The data consist of tape recordings from nine male speakers of Kano Hausa, recorded in Kano. The speech material was designed to illustrate intonation in statements and questions in appropriate contexts, and consisted of sentences on alternating High and Low tones, and on High tones only. Each tone pattern occurred in a short and a long version. There were two short basic sentences on High tones only, "Muudii yaa zoo gidaa" (Muudii came home), and "Muudii yaa ga zoomoo" (Muudii saw a hare), and one long sentence, "Uwargidan Muudii taa ga zoomoo bayan gidaanaa" (Muudii's senior wife saw a hare behind my house). The basic short sentence on strictly alternating High and Low tones was "Maalam yaa auni leemoo" (The teacher weighed the oranges), and the long sentence on this tone pattern was "Maalam yaa auni leemoo gaban garinmuu" (The teacher weighed the oranges in front of our town). In addition there was one sentence of Highs and Lows with two adjacent High tones, "Maalam yaa raba naamaa", HLHHLHL (The teacher divided the meat). Basic sentences on Low tones only were excluded by the fact that Hausa does not have any low tone verbs. The sentences were organized in blocks of one theme. Each block began with an introductory statement about a given topic; there then followed various question versions of the statement, each followed by a proper reply. This study examines a subset of the data, consisting of the introductory statements, the yes/no questions, and a question beginning with the question-word "waa...?" (who?).

The sentences were read by the speakers for the recording. The style of speech is fairly formal, but a non-colloquial style was intended, so as to maximize contrasts at this exploratory stage of the investigation. The speakers did not always produce the intended prosodic patterns on each utterance, so the results may reflect less than nine speakers. The number of speakers analyzed is specified in each case in describing the results.

Particularly the long sentences were affected in that some speakers seemed to prefer breaking them up into smaller phrases with pauses. Only the long sentences without such pauses were included in the analysis.

The recorded material was analyzed from narrow band spectrograms made on a digital spectrograph. Wide band spectrograms were used to aid in segmentation and for durational measurements. The fundamental frequency curves were traced from the 10th harmonic. One advantage of tracing pitch on spectrograms is that a pitch curve is visible even when the intensity drops at the end of an utterance, or at creaky phonation. Several of the speakers had quite creaky phonation, particularly at the end of utterances.

In order to test the model, measurements for all speakers were made of the fundamental frequency at the beginning points (BEG), the end points (END), and at all turning points (H,L) in statements and questions (figure 1). The model predicts that it should be possible to specify tones in terms of an intonational grid. Accordingly, straight lines were fitted by eye through the High and Low turning points to form a grid. Each line was drawn so as to fit as many maxima and minima as possible in the utterance. Occasional turning points may fall outside the grid, but they can usually be accounted for by local rules. For example, in utterances with alternating Highs and Lows the last High is often lowered, and in the long utterances on alternating Highs and Lows the penultimate High is often raised. In yes/no questions the last High is always considerably raised and is considered to be outside the grid. The rates of slope of these grid lines were measured and calculated in % per second. In addition the total duration for each statement and corresponding yes/no question was measured.

RESULTS AND DISCUSSION

Consider the statements first. Figure 2 shows one speaker's pitch curves of three statements with different tone patterns and lengths. In the utterances with alternating Highs and Lows the two grid lines are be in a fairly constant relationship to each other throughout the utterances, so it is possible to derive one line from the other. As there are no sentences on Low tones only, the topline is taken as basic, and the bottom line as derived.

	short	long
H only	13.89 (4.48)	7.86 (2.91)
	n=18	n=7
	v=32.3 %	v=37 %
HL	32.71 (13.28)	16.29 (2.98)
	n=7	n=7
	v=40.6 %	v=18.3 %
HLHHLHL	28.75 (6.69)	
	n=8	
	v=23.3 %	

Table I

Mean slopes in % per second with standard deviations.
v = coefficient of variation (SD/x).

The statement intonation describes a downward slope, even in utterances on High tones only. The parameters that need to be set in order to generate pitch curves for statements include:

1. the slope of the topline for short and long sentences.

2. the frequency of the first High.

Table I shows the mean % per second rate of slope with standard deviations of the top line slopes. The different means show that the slope is affected by both sentence length and tone patterns. The longer sentences have a slower rate of descent of the topline, and paired t-tests show that the differences between short and long sentences are significant ($p<0.001$). As the data only include two sentence lengths for each tone pattern, the precise relationship between sentence length and topline slope cannot be determined here. Very approximately for these data, doubling the number of syllables will halve the rate of the slope. Sentence length will thus have to be part of the specification of the topline. The rate of slope also varies with different tone patterns. The utterances with alternating Highs and Lows have about twice as steep a slope as those with Highs only in utterances of comparable length. Paired t-tests show that the differences in slope between utterances with different tone patterns are significant ($p<0.001$). The slope of the utterance on the HLHHLHL pattern with two adjacent Highs is also significantly different from that of the strictly alternating sentence ($p<0.05$). The downward slope in statements with alternating Highs and Lows is wellknown as Downdrift, a phenomenon where Highs and Lows are lower than in other positions. Downdrift has sometimes been regarded as an intonational phenomenon (e.g. Hombert 1974, Ladefoged 1975, Meyers 1976), sometimes as an effect of local tone assimilations of Highs to preceding Lows (Abraham 1941, Hyman and Schuh 1974). Both views are partly correct. The downward slope in the sentences on High tones only can of course not be an effect of local tone assimilations, but must be considered to be a global statement intonation. The much steeper slope in the sentences with alternating High and Low tones is here interpreted as an effect of combining the global statement intonation from Highs only with a fairly slow rate of slope with local tone assimilation rules that lower Highs

following Lows substantially. This view is supported by the fact that in the sentence with two adjacent Highs in the middle ("Maalam yaa raba naamaa"), where only two Highs follow Lows, the slope is less than in the strictly alternating "Maalam yaa auni leemoo", where three Highs follow Lows. The more Low-High sequences there are, the steeper the slope becomes. In this view then, Hausa has both a global statement intonation with the slope of that on High tones only, and local tone assimilations that will change the slope depending on the number of Low-High sequences in the derivation of the final slope. The final slope can be described as an effect of both an almost universal downward slope in statement intonations, and more language specific modification of that declination.

The topline thus varies with sentence length and tone patterns. It does not seem to be predictable from any other parameter. It is possible that the slope could depend on the frequency of some point at the beginning or end of the utterance. But in these data there is no significant correlation between the slope, on the one hand, and the beginning or endpoints, or the initial or final Highs, on the other. The topline slope in statement intonation is a parameter that functions like a primitive. Barring a more precise knowledge of the relationship between sentence length and slope, it has to be set for each length in generating pitch curves.

The topline has to be anchored at some point, presumably at the beginning or end of a statement. Paired t-tests were used to determine which point differed least between the various types of statements. The first High was the only point that showed nonsignificant differences between most sentences of different tone patterns and different lengths. In the relatively few cases where the first High did differ between sentences, the difference was in the direction of this point being higher for longer sentences than for shorter sentences. The first High is also the point with the smallest coefficient of variation

for different types of sentences within speakers. The first High will thus be used as a further set value to anchor the topline.

This is a different situation than in Swedish and in English, where the first High is higher the longer the utterances are (Bruce 1982, Pierrehumbert 1980). The end point tends to be invariant, as well, so the slope in these languages is in a sense a function of these two points. In Hausa, however, the first High, as well as the starting point (BEG), are points that, within a speaker, show very little variation. It remains an open question how typical the differing results for Hausa are in this respect. It is possible that the differences in the behaviour of the first High could be an effect of different type of speakers rather than a language difference. Many studies of intonation have used speakers that are highly trained in the task of reading in front of a microphone, while the speakers used in this study were not used to this.

Within each type of statement there is a great deal of variation between speakers in the topline slope, as evidenced by the fairly large standard deviations in table I. Although a speaker will vary the topline slope with sentence length and tone patterns, the precise amount of slope appears to be idiosyncratic. The variation between speakers is not correlated with the speakers' mean fundamental frequency or range. In generating the grid for a statement, the mean slope will be taken to represent a typical speaker.

In the sentence with alternating Highs and Lows speakers also vary to some extent in their range. This variation probably reflects a non-linguistic factor of the attitude of the speaker. An involved speaker uses a larger range than a detached speaker (Hadding-Koch 1961, Bruce 1982). The range can be expressed as the mean ratio between Highs and following Lows throughout the utterance. Speakers tend to keep this ratio fairly constant throughout an

utterance. The mean ratio between Highs and following Lows for the short "Maalam yaa auni leemoo." is 1.25 (SD=0.05, n=7), and for the longer "Maalam yaa auni leemoo gaban garinmuu." this mean ratio is 1.2 (SD=0.06, n=5). The mean ratio will be used in deriving the bottom line from the topline in the grid.

The timing of turning points in the fundamental frequency curve in relation to the segmental structure of the utterance is also a necessary part of the specification. In sentences with alternating Highs and Lows there is a strong tendency for the turning points to occur right at the syllable boundaries. In sequences with open syllables, like CVV-CVV, and CV-CVV the turning point coincides with the V-C boundary, but in sequences with closed syllables, CVC-CVV, the turning point occurs in the middle of the double consonant. The fundamental frequency curve associated with a High tone syllable will thus describe a movement towards a maximum during the syllable, and reach its maximum at the border to the next syllable. Similarly, the fundamental frequency curve associated with a Low tone syllable will reach its minimum at the very end of the syllable. The turning points in the fundamental frequency curve thus serve as a phonetic correlate to syllable boundaries. This pattern in the timing of the turning points is not much influenced by position in the sentence, but it is relatively stable throughout an utterance.

Figure 3 shows how a pitch curve for the High/Low sentence "Maalam yaa auni leemoo" can be generated.

First, set the duration for the given number of syllables (here, the seven syllables take an average of 130 centiseconds). Set the value for the first High (the average speaker's value is 160 Herz). Construct the intonational grid, anchored on an arbitrary (speaker-specific) value for the first High. The basic topline slope for a statement of this length is taken to be the slope for High tones only, about 14% per second.

Construct the bottom line as a ratio (here 1.25) of the topline. Secondly, insert start and end points. As a convention, every statement begins and ends with the pitch for a Low, even if the first (or last) tone is a High to (or from) which the pitch must immediately rise (or fall). Thirdly, insert High and Low tones on the grid lines. The Highs and Lows will constitute turning points on the grid lines. Align the sentence so that the Highs and Lows occur at the syllable boundaries. The fourth step involves applying the tone assimilation rules and this will result in a steeper slope, the more Low-High sequences the sentence has. For this sentence the rate increases from about 14% per second to 33% per second. At this stage optional rules may apply that move specific tones away from the grid lines. These are processes that are found for some speakers. They include lowering the final High tone when it occurs after a Low, raising a Low after the first High, and raising a penultimate High in longer sentences. Lastly, concatenate all the turning points into a smooth curve. For these simple statements the resulting pitch curve approximates the curves in the data well.

Now the intonation in questions can be considered. In Hausa, yes/no questions are signalled by intonational means, not by syntactic, or morphological means. Cowan and Schuh (1976), and Miller and Tench (1980) characterize this type of question intonation as similar to intonation in statements but with less of a slope, and a local rise of the final High. Newman and Newman (1981) posit a separate question morpheme at the end of all questions. This morpheme consists of a Low tone with length. Short final vowels lengthen in questions. The low tone is justified by a claim that in questions the final raised High is followed by a fall.

Typical fundamental frequency curves of yes/no questions are illustrated in figure 4. The salient features of yes/no questions are a suspension of the statement downward slope to zero slope, and a considerably raised

pitch of the last High tone. Yes/no questions on High tones never exhibit any downward slope, and in the questions with alternating High and Low tones only two out of nine speakers follow the pattern described in the references with somewhat of a downward slope of the grid before the final raised High. The rest have zero slope. The width of the grid is not significantly different from that in statements. The frequency of the first High tone does not differ significantly between statements and corresponding yes/no questions, so Hausa questions are not characterized by a raised register, as questions may be in other languages, for example in Swedish (Hadding-Koch 1961, Gårding 1979, Bredvad-Jensen 1983).

The raised last High is sometimes, but not always followed by a fall. Figure 5 shows the fundamental frequency curves of the last two syllables of all the speakers superimposed for each yes/no question. Unbroken lines indicate a final rise that is not followed by a fall, and dashed lines indicate a rise that is followed by a fall. Many speakers do not end a yes/no questions with a fall. In fact, even when the last syllable is on a Low tone, as in the utterances ending with "...naamaa?", three of the nine speakers do not end in a fall. A final Low tone question morpheme is not part of a general pattern in the yes/no questions. The final falls that do occur can be viewed as instances of the general convention of starting and ending low, rather than as a manifestation of an abstract low tone question morpheme.

In addition, statements and corresponding yes/no questions differ significantly in total duration, The questions are about 10% shorter than the corresponding statements. The difference is due to an overall shortening of the question utterance as compared to the statement, not to shortening of any particular part of the utterance. This shortening is not found in all languages, for example not in German (Bannert 1983). If it is language specific, this type of overall shortening will have to be part of the rules

for question formation in Hausa.

Figure 6 shows how a yes/no question "Maalam yaa auni leemoo?" can be generated. First, set the duration of the question at 10% less than the corresponding statement (here = 117 csec.). Construct the grid with zero slope, anchored on the first High as for statements. Secondly, insert a low starting point, and optionally a low endpoint. Thirdly, insert the tones, and fourthly, raise the last High tone. The amount of increase of the last High varies considerably, but is typically around 5-10%. Lastly, concatenate the turning points into a smooth curve.

Typical pitch patterns of question word questions are illustrated in figure 7. They display characteristics that place them as something between the pitch patterns in statements and yes/no questions. A question-word question of a sentence with alternating High and Low tones will have a downward slope, but it is significantly less than that in statements. The mean slope of "Waa ya auni leemoo?" is 23% per second (SD=3.1, n=8). But "waa"-questions of the basically High tone sentences have a zero slope, just like the yes/no questions. Also like the yes/no questions the last High tone may be raised, but only for two out of the nine speakers. Mostly the last High tone is not raised. There is however always a fall after the final High tone. This fall is most simply accounted for by assuming it to be part of the general convention of starting and ending Low.

SUMMARY

The intonation of simple sentences in Hausa can be represented as grids of (near)parallel lines. The rate of slope of the grid is at least related to sentence type (statement and question), sentence length, and tone pattern. Pitch curves are generated by rules from underlying grids. The base form of a grid is a topline from which the bottom line is derived as a ratio. The topline is anchored on the

first High tone. The slope of the base intonational grid is that which appears from High tones only, when there is no influence from tonal assimilations. All questions are specified with zero slope, and statements with a slope that depends on the length of the sentence. In this study only two slopes appear as base toplines, one for short, and one for long sentences, but obviously a function can be worked out from more data on different sentence lengths. Lexical tones and boundary marks are mapped onto the grid lines. In sentences with alternating High and Low tones assimilatory processes apply that lower Highs following Lows, so that the slope of the grid becomes steeper, thus claiming that downdrift is an effect of both intonation and tonal assimilations. These assimilations apply in statements and question-word questions, but not usually in yes/no questions. Local tone rules may move tones away from the grid lines. Lastly, a smooth curve is generated through the boundaries and tones. At least for the simple sentences dealt with above, the resulting pitch curves approximate the data well.

Acknowledgments: Many thanks to Will Leben, Nicolas Faraclas, and Brian McHugh for recording the speech material in Kano. I am much indebted to Peter Ladefoged, Ian Maddieson, and the UCLA Phonetics Lab for comments on earlier versions. Eva Gårding provided much inspiration. Financial support came from the National Science Foundation and the Swedish Council for Research in the Humanities and Social Sciences.

REFERENCES

Abraham, R. 1941. Modern Grammar of Spoken Hausa. Kaduna: Government Printer.

Bannert, R. 1983. Some phonetic characteristics of a model for German prosody. Working Papers. 25:1-34. Department of Linguistics, Lund University.

Bredvad-Jensen, A.-C. 1983. Perception studies of Swedish interrogative intonation. Abstracts of the Tenth International Congress of Phonetic Sciences. 578.

Bruce, G. 1982. Experiments with the Swedish intonational model. In H. Fujisaki and E. Gårding (eds.) Preprints of Papers for the Working Group on Intonation. The XIIth International Congress of Linguists,Tokyo 1982.

Cowan, R. and R.Schuh. 1976. Spoken Hausa. Ithaca: Language Services, Inc.

Gårding, E. 1979. Sentence intonation in Swedish. Phonetica 36.207-215.

Gårding, E. 1981. Contrastive prosody: a model and its application. Studia Linguistica 35:146-165.

Gårding, E. and G. Bruce. 1981. A presentation of the Lund model for Swedish intonation. In T.Fretheim (ed.) Nordic Prosody II: 33-39. Trondheim: Tapir.

Hadding-Koch, K. 1961. Acoustico-phonetic studies in the intonation of Southern Swedish. Lund: Gleerups.

Hombert, J.-M. 1974. Universals of downdrift: their phonetic basis and significance for a theory of tone. In W. Leben (ed.) Studies in African Linguistics Supplement 5:169-183.

Hyman, L. and R.Schuh. 1974. Universals of tone rules: evidence from West Africa. Linguistic Inquiry 5.81:115.

Ladefoged, P. 1975. A Course in Phonetics. New York: Harcourt, Brace, Jovanovich.

Meyers, L. 1976. Aspects of Hausa tone. UCLA Working Papers in Phonetics 32.

Miller, J. and P.Tench.1980. Aspects of Hausa intonation: utterances in isolation. Journal of the International Phonetic Association 10,1-2: 45-63.

Newman, P. and R. Ma Newman. 1981. The question morpheme q in Hausa. Afrika und Ubersee 64,1:35-46.

Pierrehumbert, J. 1980. The Phonology and Phonetics of English intonation. PhD thesis, MIT.

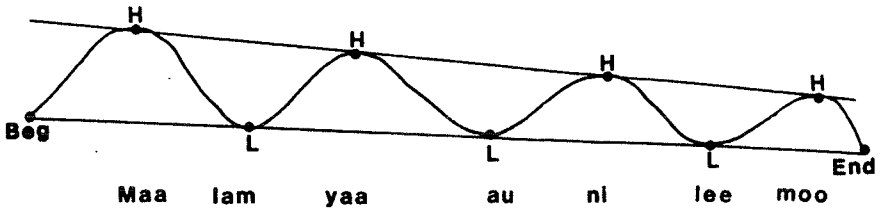

Figure 1

A model Hausa utterance.

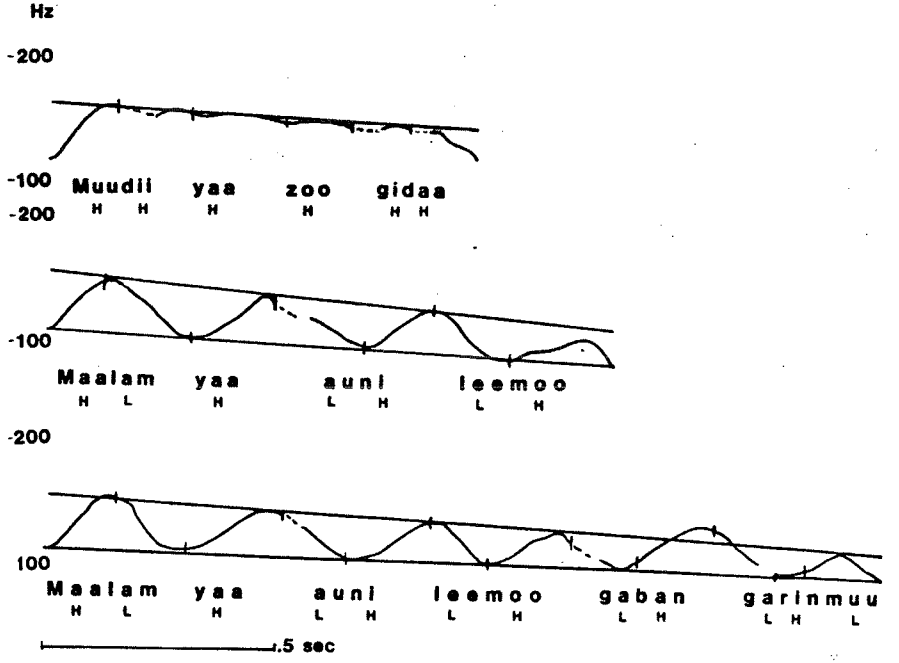

Figure 2

Pitch curves of statements on different tone patterns and different lengths of one speaker.

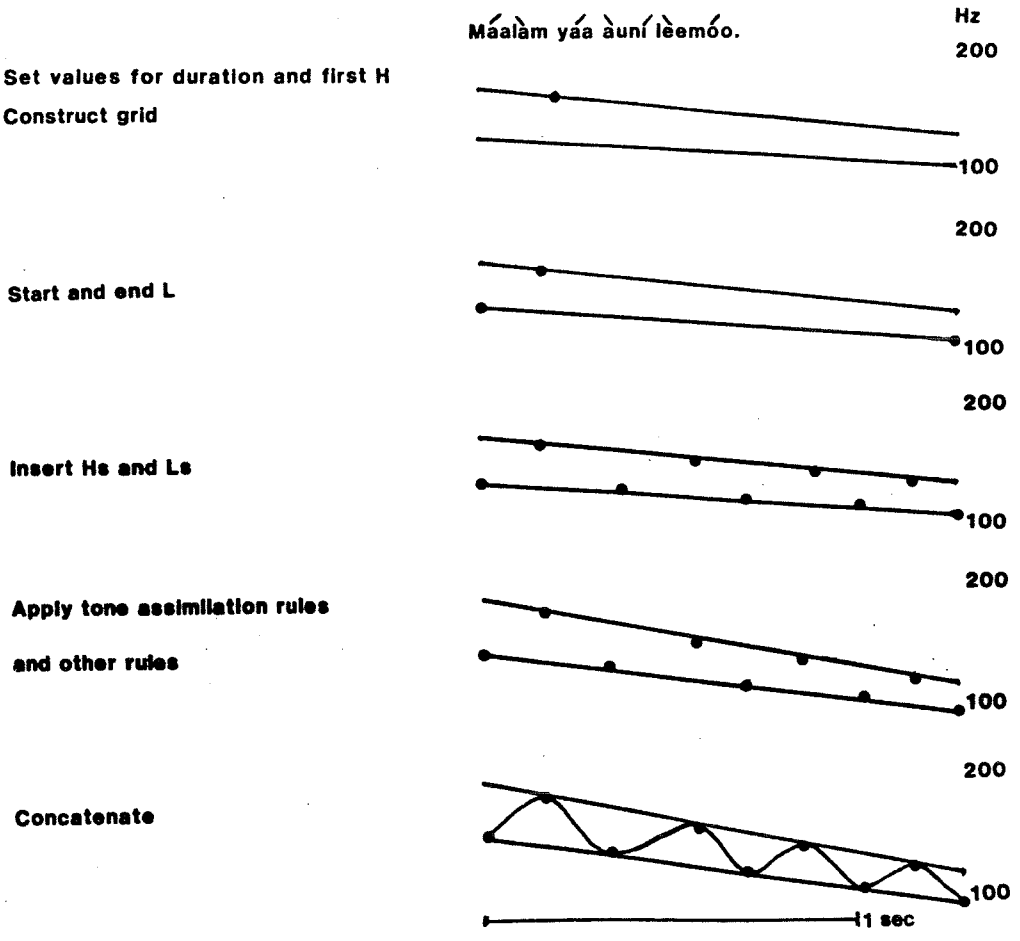

Figure 3

Rules for generating the pitch curve associated with a sentence on alternating High and Low tones.

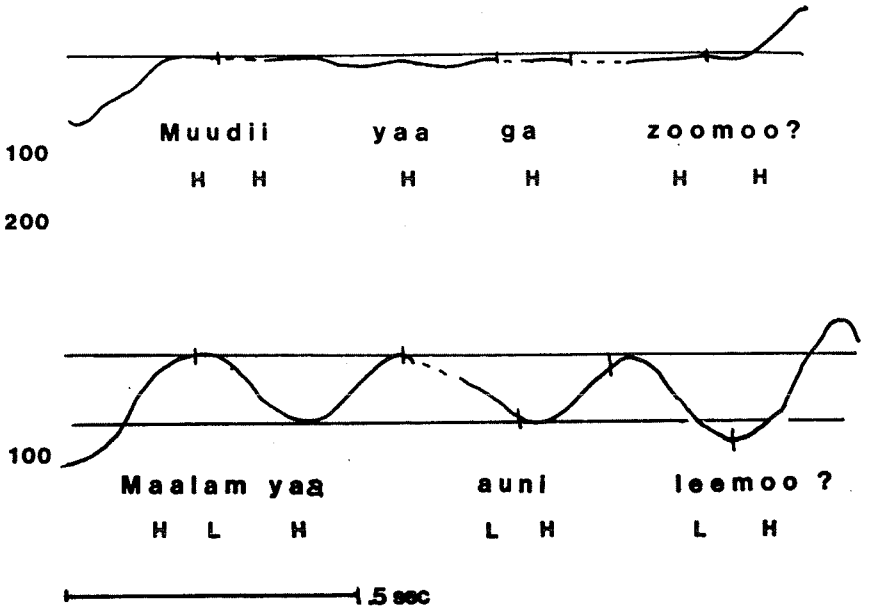

Figure 4

Pitch curves of yes/no questions on different tone patterns of one speaker.

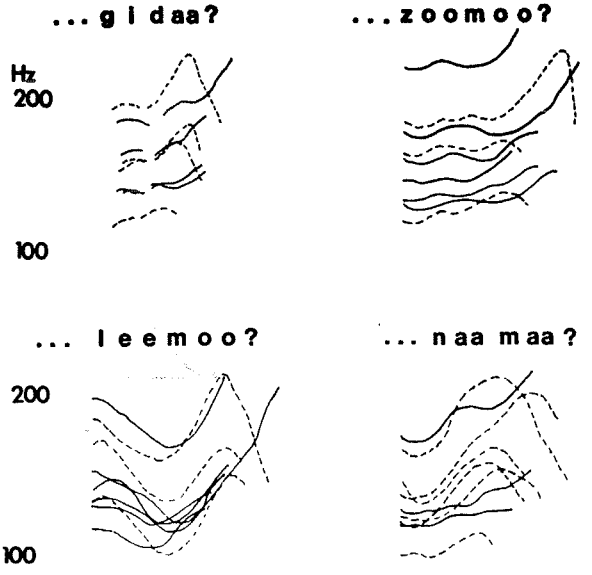

Figure 5

Pitch curves of the last two syllables in four questions for 8 or nine speakers.

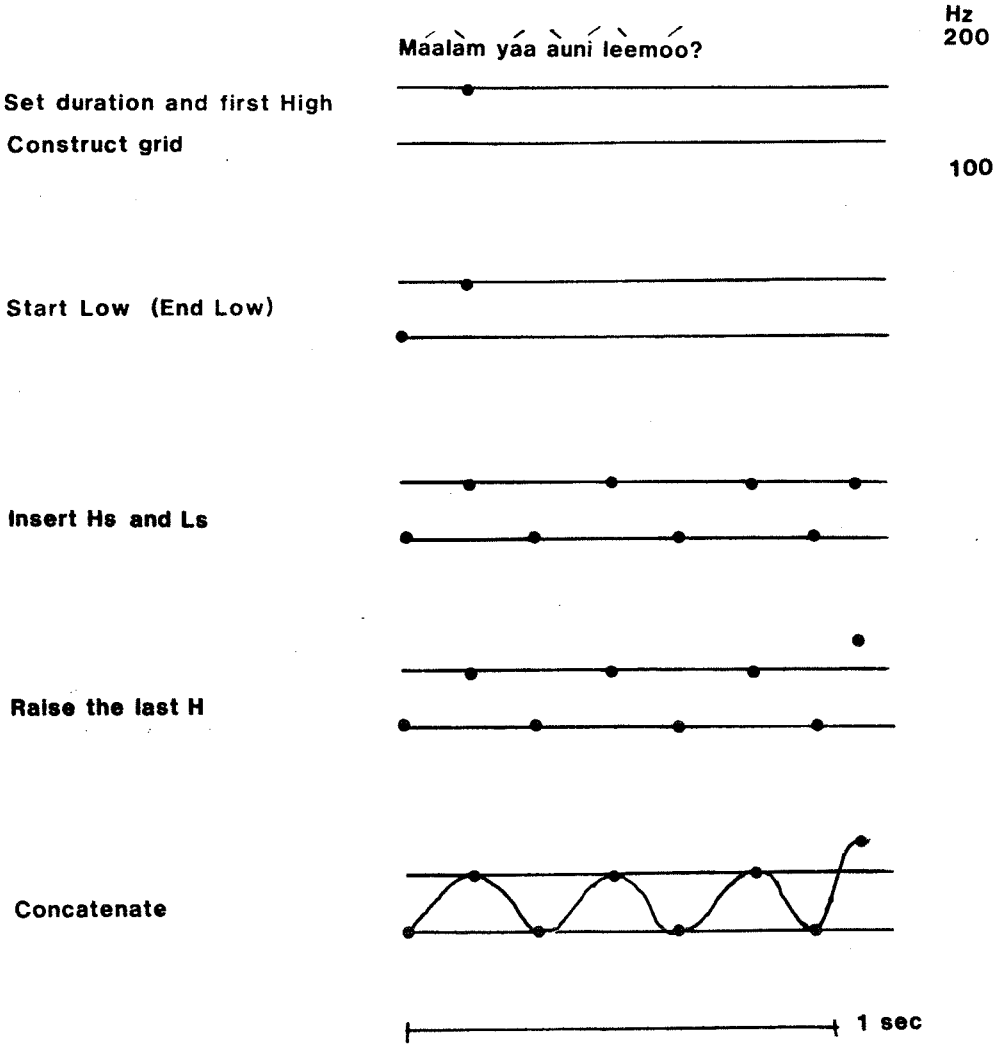

Figure 6

Rules for generating the pitch curve associated with a yes-no question on alternating High and Low tones.

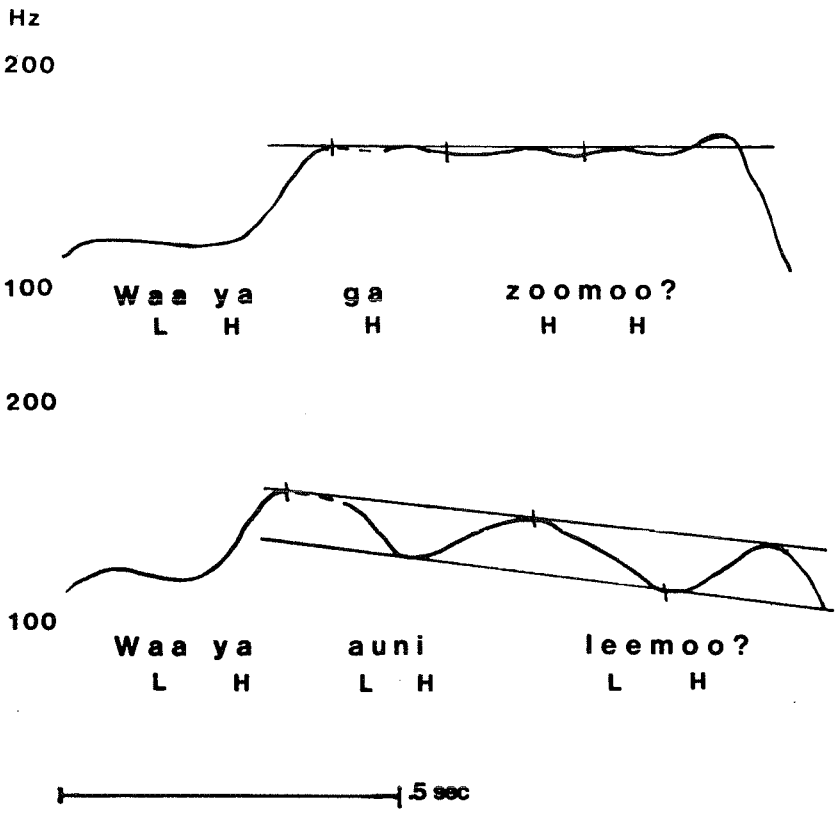

Figure 7

Pitch curves of question-word questions on different tone patterns of one speaker.