The Effects of Boundary Tones on the f0 Scaling of Lexical Tones

Scott Myers

Department of Linguistics
University of Texas at Austin
s.myers@mail.utexas.edu

Abstract

Many languages display a pattern in which the f0 values of tones are higher in a phrase ending in a high boundary tone than in one ending with a low boundary tone. In some cases it is only the tones closest to the boundary tone that are affected, while in others it affects all tones in the phrase.

An account is suggested in this paper, according to which these effects are due to perceptually-based sound changes. Coarticulatory effects of the boundary tone on the f0 value of an immediately preceding tone can in shorter phrases be confused with a more global effect on phrasal scaling parameters. Such confusions can lead to a sound change linking pitch range and downtrends to final boundary tone. Such an account expresses the attested patterns without special representational machinery.

1. The effects

Boundary tones [25] are distinguished from other tones by their distribution. Unlike stress-dependent intonational pitch accents, boundary tones only occur on the edgemost syllable of a domain, such as the last syllable of an intonational phrase. Unlike lexical tones, boundary tones have a phrasal distribution and function to distinguish sentence types.

Boundary tones must also be distinguished from other sorts of tone with regard to their phonetic implementation. In Hungarian, Gósy and Terken [5] find that a boundary tone H% marking questions had a higher f0 value in the same position than a H* marking focus. In Chichewa, a Bantu language spoken in Malawi, questions end in a final rise, represented with the boundary tone H%, and statements in a final fall (L%) [18]. Lexical high tones are subject to **downdrift** – the second H in a HLH sequence has a lower f0 value than the first one. But H% neither conditions nor undergoes downdrift. Indeed, a H% at the end of a subject phrase has no effect on the downdrift pattern of surrounding lexical high tones [18].

Boundary tones often have strong effects on the f0 scaling of other tones in the utterance. The f0 values of lexical high tones in Chichewa are markedly higher in questions than in statements – an average of 75 Hz higher among the 6 subjects in [18]. In addition, while the second of two lexical high tones was an average of 12% lower than the preceding one in statements, it was only 2% lower in questions. Thus the pitch range was higher and the downdrift trend less steep in questions than in statements.

The same results held in a further study of Chichewa with three more subjects [20], in which pitch range was varied by asking the subjects to speak loudly, normally, or softly. F0 values were higher for louder conditions, as in the English data of Liberman and Pierrehumbert [16]. Mean f0 values for lexical high tones in a sentence with four high tones are displayed in Figures 1-3 for three speakers. Each line connects the mean values for the four high tones in a given condition. The solid lines represent questions and the dashed lines represent statements. There are three solid lines and three dashed lines – one line for each loudness level. The loudest speech of the third subject (the only female of the group) typically went into f0 values beyond the measurement capabilities of our pitch tracker, so only two loudness levels are presented for her. It can be seen that for all three subjects the solid line is almost always above the corresponding dashed line, and it always has a flatter slope.

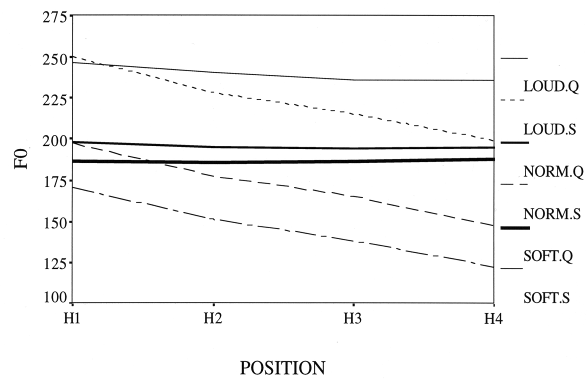

Figure 1: *F0 values for 4 successive high tones (H1-H4) in questions and statements (Subject #1)*

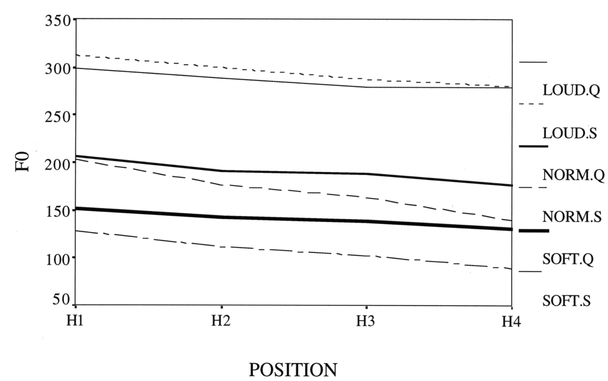

Figure 2: *F0 values for 4 successive high tones (H1-H4) in questions and statements (Subject #2)*

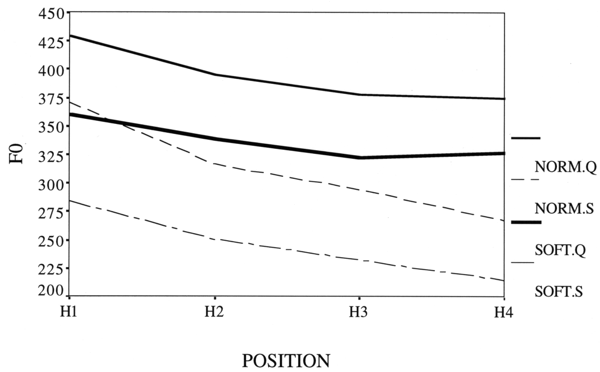

Figure 3: *F0 values for 4 successive high tones (H1-H4) in questions and statements (Subject #3)*

These are well-attested patterns. F0 peaks are higher in questions than in statements in languages as diverse as Swedish [3], Russian [29], Mandarin [10, 28], Vietnamese [22], Oneida [6], Brazilian Portuguese [17], Lingala [7], Jita [2], Kikuyu [1], and Hausa [14]. In perception tests, high values for f0 peaks contribute to identification of utterances as questions in American English and Swedish [8], Danish [31] and Vietnamese [22].

F0 downdrift is less in questions than in statements in English [24], Danish [30], Swedish [3], Russian [29], Kikuyu [1], Zulu [27], and Hausa [14]. The difference in downdrift has been showed to serve as a cue distinguishing questions and statements in Danish [31], Brazilian Portuguese [17], and Dutch [9].

Final lowering is a pattern in which a final f0 peak is lower than would be expected from other scaling effects such as downdrift. Liberman and Pierrehumbert [16] provide evidence for final lowering in English statements. Pierrehumbert and Beckman [26] show that final lowering occurs in statements but not in questions in Japanese. The same is found in Lomongo [12] and appears to be the case in Vietnamese, judging from the displays in Nguyên and Boulakia [22].

All three of these differences in f0 scaling between questions and statements are assimilatory, since in all three f0 values in a phrase are higher if there is a final high boundary tone than if there is a final low boundary tone. The effects are stronger the closer a tone is to the end of the phrase. Final lowering only affects the last high tone (or the last few [26]) before the boundary tone. Differences in downdrift affect all non-initial tones including the last one, and each such tone is lower the closer it is to the final boundary. This orientation to the end of the phrase suggests that the effects are conditioned, if only indirectly, by the final boundary tone.

Another fact that implicates the final boundary tones in these scaling effects is that the same patterns can occur in non-questions that end in a final rise. One Chichewa speaker participating in the experiment in [18] consistently had a final rise in statements that were non-final in the utterance. The f0 peaks in these non-final statements had higher values than in the final statements, and had a less steep downdrift trend. In both measures, the non-final statements were intermediate between statements and questions. The other subjects, who did not have a final rise in non-final statements, also did not have a distinction in f0 scaling between final and non-final statements. Thorsen [30] found that non-terminal declarative clauses have less downdrift that in final declarative clauses, and her sample pitch tracks suggest that these non-terminal clauses also have a final rise. Ho [10] similarly reports that both questions and exclamations have a final rise in Mandarin, and that both have a higher mean f0 values throughout than in statements. Inkelas and Leben [14] report that the f0 range is raised after a phrase-peripheral focus marker, as well as before a phrase-peripheral question marker.

2. An approach

One puzzle about the effects of boundary tones on f0 scaling is their globality. One tone generally only affects the neighboring tones immediately before or after it. This is true of coarticulatory effects [28, 32], as well as tonal phonological patterns [13, 33]. Yet in some of these cases, a phrase-final tone is affecting all the tones in a whole intonational phrase.

It is also remarkable that the effects seem to be uniformly assimilatory. Phonological tone assimilation can be clearly represented through spread of a tone from one syllable to others [4]. But while a boundary tone is a real tone associated with a specific syllable, pitch range and downtrends are gradient quantitative patterns best expressed in terms of the phonetic interpretation of phonological tones. Pitch range varies continuously as a function of overall emphasis [16], with higher f0 values associated with louder speech. This kind of variation cannot be represented with any set of discrete intonational tones, so it is difficult to see how there can be any kind of spreading between a boundary tone and the representation of pitch range.

One approach to this has been to use the representational resources of autosegmental phonology. Pierrehumbert and Beckman [26] propose that the boundary tone is associated not only to a phrase-peripheral syllable, but also to the whole prosodic phrase. The tone is thus not just a property of the syllable it is realized on, but also of the whole phrase that contains it. Myers [18] exploits such a representation to express the long-distance effects of H% in Chichewa. Inkelas and Leben [14] pursue a similar strategy, attributing the broad scope of question marking H to its being on a register tier distinct from the basic tone tier. As the only element on this tier, it is free to sweep across the phrase without encountering anything to block it.

But such a representational approach to the effects of H% does not ultimately answer the questions about them. Instead of asking why only boundary tones have global effects, we would now have to ask why only boundary tones are associated with a phrase node or a register tier. The representation allows us to restate the generalization in a new way, without gaining any particular insight into it.

Moreover, such an approach does not explain why the effects are assimilatory. A lexical high tone juxtaposed with a high tone is often turned into a low tone [19]. If pitch range and downdrift are built into the representation in the form of tones, why don't we see a parallel dissimilatory effect, in which a high boundary tone conditions lower pitch range and/or steeper downdrift?

An alternative account of these facts takes a diachronic perspective. It begins with the one direct physical effect of boundary tones – the coarticulatory raising of f0 before H% and lowering of f0 before L% [16, 26]. This is a local effect that is entirely analogous to the kinds of coarticulatory effects we see among adjacent lexical tones [28, 32].

Most intonational phrases are short, with only one or two f0 peaks. This is particularly true of speech directed to children learning a language. In such short phrases, the coarticulatory effect, moving the last f0 peak closer in f0 value to the following boundary tone, will be very difficult to distinguish from the global effects on downdrift and pitch range. If we find in a phrase with one peak close to the final boundary that the peak is higher before H% than before L%, that is compatible with a global raising of lexical H before boundary H%, or with a local anticipatory coarticulatory effect conditioned by that boundary tone. If we find in a phrase with two peaks that the second one is markedly lower before L% than before H%, that could be due to a difference in the downdrift slopes for H%-final compared to L%-final phrases, or it could be due to local coarticulatory lowering of H before L%. To distinguish the global effects from the local effects in a phonetic investigation, we would just have to look at longer phrases. To distinguish downdrift from final lowering, for example, Liberman and Pierrehumbert [16] have to turn to intonational phrases with 3 to 5 peaks.

But the point is that the sample of intonational phrases with 3 to 5 peaks is relatively small in natural speech, and in particular in speech directed to language learners. It would therefore be natural for listeners and language learners to overgeneralize from short phrases, concluding that the speaker's purely local coarticulatory effect of H% and L% on an immediately preceding tone is a global effect on pitch range or downtrends. The result is a mismatch between the speaker's intended pattern and the listener's learned pattern – the beginning of a sound change [23].

The change will be in an assimilatory direction because it starts with tonal coarticulation. Pitch range will be raised before H% because the effect on pitch range is based on a generalization of raising of a tone immediately before H%. Non-initial high tones will be lowered, leading to a downtrend, because this effect is based on the lowering of a tone immediately before L%.

According to this diachronic account, there is no direct representational connection between the boundary tones and the global raising effects. The global effects are represented in the setting of pitch range and downtrends for a phrase, and these are related to the final boundary tone only through a series of natural sound changes.

This account would allow us to make sense of the fact that the interaction of boundary tones with lexical tones is asymmetrical: boundary tones affect the f0 value of lexical tones, but not vice versa. The f0 value of a boundary tone seems to be completely independent of what lexical tones (or non-boundary intonational tones) precede it.

There doesn't seem to be a physical reason why laryngeal coarticulation should affect the boundary tone any less than a lexical tone, but there are functional factors that favor the boundary tone. The lexical tone usually comes in a morpheme with segments that help distinguish it from other morphemes, while the boundary tone is a floating tone without segmental content. Furthermore, the distinction between question and statement seems to be very basic in human language, and is very hard to predict from context, unlike lexical distinctions such as that between *gúrú* "big" and *guru* "burrow" in Shona. Such factors might be enough to counter any tendency for the boundary tone to change to match a preceding lexical tone.

There are certainly cases of languages of which it is said that they distinguish questions and statements just by means of differences in pitch range and downdrift, without a final rise or fall conditioning it. Downing [2], for example, describes questions in Jita (a Bantu language) as having a higher pitch range than statements, but does not transcribe any final rise in questions.

These cases could be interpreted as demonstrating that the pitch range and downdrift effects distinguishing sentence types are independent of the final boundary tone. But such effects could also arise diachronically if they were all originally conditioned by the boundary tones. Consider a language that has global effects on f0 scaling conditioned by the final boundary tone. If the contrast between H% and L% is lost over time due to the common pattern of final tone neutralization, avoiding final H [21], then the secondary cues of pitch range and downdrift could become the chief cues distinguishing the sentence types. This would be analogous to the way a vowel before a nasal can come to be distinctively nasalized as the nasal becomes weaker [15]. The conditioned effect of coarticulatory nasalization is reinterpreted with the loss of the nasal as unconditioned, distinctive nasalization. The emergence of tones from the secondary f0 effects of phonation classes is another parallel case [11].

This account is at this stage purely speculative (though not really any more so than the representational accounts). It would, however, be possible to test the fundamental premise of the account – that the connection between boundary tones and scaling effects arises out of a perceptual confusion.

The stimuli for this test could consist of series of pure tones, since the basis of the account is not specific to speech perception. Listeners would be given short series of tones in a training stage, divided explicitly into two classes. In the first class, the analogue of short questions, there would be three tones – the first one high, the second low, and the third in between the f0 values of the other two. The three tones would differ in f0, but not in their relative difference from each other. The tone series in the second class, the analogue of short statements, would differ from the first in that the second and third tones would both be at the same low level.

In the test stage of the experiment, subjects would be given longer series of tones varying in the trend of f0 values. They would vary in the downward slope of f0 values, and in the overall f0 range. The prediction is that subjects would generalize from the short forms, and include among the analogues of questions both those sequences with a flat slope and those with a high f0 level.

3. References

[1] Clements, G.; Ford, K., 1981. On the phonological status of downstep in Kikuyu. In *Phonology in the 1980's,* Goyvaerts, D., ed. Amsterdam: John Benjamins, 309-357.

[2] Downing, L., 1995. The metrical domain of register raising in Jita: An optimality approach. In *Bantu Phonology and Morphology,* Katamba, F., ed. München: Lincom Europa, 28-39.

[3] Gårding, E., 1979. Sentence intonation in Swedish. *Phonetica*, 36, 207-215.

[4] Goldsmith, J., 1976. *Autosegmental Phonology*. Bloomington (Indiana): Indiana University Linguistics Club.

[5] Gósy, M.; Terken, J., 1994. Question marking in Hungarian: Timing and height of pitch peaks. *Journal of Phonetics,* 22. 269-281.

[6] Grimm, C., 1997. *The Phonetic Realization of Pitch Accent in the Ontario Dialect of Oneida*. M.A. project report, SUNY-Buffalo.

[7] Guthrie, M., 1940. Tone ranges in a two-tone language (Lingala). *Bulletin of the School of Oriental and African Studies*, 10, 469-478.

[8] Hadding-Koch, K.; Studdert-Kennedy, M., 1964. An experimental study of some intonation contours. *Phonetica,* 11, 175-185.

[9] van Heuven, V.; Haan, J., 2002. Temporal distribution of interrogativity markers in Dutch: A perception study. In *Laboratory Phonology 7,* Gussenhoven, C.; Warner, N., eds. Berlin: Mouton de Gruyter, 61-86.

[10] Ho, A., 1977. Intonation variation in a Mandarin sentence for three expressions: Interrogative, exclamatory and declarative. *Phonetica,* 34, 446-457.

[11] Hombert, J.-M.; Ohala, J.; Ewan, W., 1979. Phonetic explanations for the development of tone. *Language*, 55, 37-58.

[12] Hulstaert, G., 1961. *Grammaire du Lomongo, Premiere Partie: Phonologie*. Tervuren: Musée Royal de l'Afrique Centrale.

[13] Hyman, L.; Schuh, R., 1974. Universals of tone rules: Evidence from West Africa. *Linguistic Inquiry,* 5, 81-115.

[14] Inkelas, S.; Leben, W., 1990. Where phonology and phonetics intersect: The case of Hausa intonation. In *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech*, Kingston, J.; Beckman, M., eds. Cambridge: Cambridge University Press, 17-34.

[15] Kawasaki, H., 1986. Phonetic explanation for phonological universals: The case of distinctive vowel nasalization. In *Experimental Phonology,* Ohala, J.; Jaeger, J., eds. Orlando: Academic Press, 81-103.

[16] Liberman, M.; Pierrehumbert, J., 1984. Intonational invariance under changes in pitch range and length. In *Language Sound Structure,* Aronoff, M.; Oehrle, R., eds. MIT Press, Cambridge. 157-233.

[17] de Moraes, J. A., 1998. Intonation in Brazilian Portuguese. In *Intonation Systems: A Survey of Twenty Languages,* Hirst, D.; Di Cristo, A., eds. Cambridge: Cambridge University Press, 179-194.

[18] Myers, S., 1996. Boundary tones and the phonetic implementation of tone in Chichewa. *Studies in African Linguistics,* 25, 29-60.

[19] Myers, S., 1997. OCP effects in Optimality Theory. *Natural Language and Linguistic Theory,* 15, 847-892.

[20] Myers, S., 1999. Downdrift and pitch range in Chichewa intonation. In *Proceedings of the XIVth International Congress of Phonetic Sciences,* Ohala, J.; Hasegawa, Y.; Ohala, M.; Granville, D.; Bailey, A., eds. Berkeley: Linguistics Department, University of California, 1981-1984.

[21] Myers, S., 1999. Tone association and f_0 timing in Chichewa. *Studies in African Linguistics*, 28, 215-239.

[22] Nguyên, T.; Boulakia, G., 1999. Another look at Vietnamese intonation. In *Proceedings of the XIVth International Congress of Phonetic Sciences,* Ohala, J.; Hasegawa, Y.; Ohala, M.; Granville, D.; Bailey, A., eds. Berkeley: Linguistics Department, University of California, 2399-2402.

[23] Ohala, J., 1993. The phonetics of sound change. In *Historical Linguistics: Problems and Perspectives*, Jones, C., ed. London: Longman, 237-278.

[24] O'Shaughnessy, D., 1979. Linguistic features in fundamental frequency patterns. *Journal of Phonetics,* 7, 119-145.

[25] Pierrehumbert, J., 1980. *The Phonetics and Phonology of English Intonation*. Bloomington, Indiana: Indiana University Linguistics Club.

[26] Pierrehumbert, J.; Beckman, M., 1988. *Japanese Tone Structure*. Cambridge: MIT Press.

[27] Rycroft, D., 1963. Tone in Zulu nouns. *African Language Studies,* 4, 43-68.

[28] Shih, C., 1987. *The Phonetics of the Chinese Tone System*. Technical Memorandum, AT&T Bell Laboratories, Murray Hill, New Jersey.

[29] Svetozarova, N., 1975. The inner structure of intonation contours in Russian. In *Auditory Analysis and Perception of Speech*, Fant, G.; Tatham, M., eds. London: Academic Press, 499-510.

[30] Thorsen, N., 1978. An acoustical investigation of Danish intonation. *Journal of Phonetics*, 6, 151-175.

[31] Thorsen, N., 1980. A study of the perception of sentence intonation - Evidence from Danish. *Journal of the Acoustical Society of America*, 67, 1014-1030.

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

[33] Yip, M., 2002. *Tone*. Cambridge: Cambridge University Press.