Phonetic Motivations for the Development of Tones from Postvocalic [h] and [ʔ]: Evidence from Contour Tone Perception[1]

Jean Marie Hombert

Abstract

The purpose of this paper is to explain that the development of contrastive tones on vowels from postvocalic [h] and [ʔ] is phonetically motivated. The intrinsic effect produced by these consonants on the fundamental frequency of the preceding vowel was measured from Arabic data (4 subjects). Ten American subjects participated in a perception experiment involving synthesized vowels with Fo contours. The results show that subjects were able to perceive Fo changes smaller in magnitude than the Fo perturbations caused by [h] and [ʔ] on the preceding vowel. These two sets of data (production and perception) validate and explain the development of rising and falling tones from the loss of [h] and [ʔ] postvocalically.

In this paper I will show that the development of contour tones from the loss of final [h] and [ʔ][2] is phonetically motivated. First, I will report historical cases related to these developments; second, I will provide production data indicating the extent of the perturbation caused by [h] and [ʔ] on the fundamental frequency of the preceding vowel. Third, results from a perception experiment using synthesized speech will be presented in order to show to what extent these instrumentally-detected Fo perturbations found in production can be perceived.

The theory of sound change (discussed in Ohala 1974, 1975) motivating this study can be summarized as follows: If a sound change is widely attested in unrelated languages, it is highly probable that the explanation of the origin of such a change can be found in the functioning (or malfunctioning) of our articulatory and/or auditory mechanisms. These physiological constraints lead to intrinsic variations in the speech signal, that is, the signal produced by the speaker is a distorted version of the intended signal. If these distortions are large enough to be perceived it is reasonable to assume that they can be used by the listener to facilitate his perception of the speech signal and, at a later stage, even become the primary cues. The

---

1. Paper delivered at 8th International Congress of Phonetic Sciences, Leeds, 17-23 August 1975.

2. The symbol [ʔ] refers to a segment which involves complete closure of the vocal cords for a duration of the order of 50 msec (and consequently excludes creakiness, for instance).

original intrinsic variations are then used for distinctive purposes. Historical development of tones or tonogenesis[3] illustrates clearly this theory of sound change (Hombert 1975b). The development of contrastive tones on vowels due to consonantal influences is the best documented type of tonogenesis. One of the clearest cases of tonogenesis is the bipartition of a tonal system resulting from the loss of a voicing contrast in prevocalic position (Haudricourt 1961, Matisoff 1973). In Hombert (1974) I have provided data accounting for this development both in articulatroy and perceptual terms. In this paper I shall address myself to a less widely discussed type of tonal development, namely the development of contour tones (falling and rising) from the loss of [h] and [ʔ] in postvocalic position.

The effect of a glottal stop on the pitch of the preceding vowel is widely attested. By the 6th Century glottal stop had disappeared in Vietnamese and was replaced by a rising tone (Haudricourt 1954, Matisoff 1973). In the Lolo-Burmese family, Burmese high tone corresponds to Jingpho glottal stop (Maran 1971) and Lahu high rising tone developed through glottal dissimilation (Matisoff 1970). Mei (1970) has shown that Middle Chinese shang sheng (rising tone) comes from a final glottal stop. Another development probably related to the pitch-raising effect of glottal stops is the upstep phenomenon observed in Acatlan Mixtec (Pike and Wistrand 1974). In this language it seems that there was historically a glottal stop at the points where there is now a break in the register level, the following syllable establishing a higher register (H. Javkin, personal communication). The development of a falling tone from a post-vocalic [h] has been observed in two cases. Middle Chinese qu sheng (falling tone) comes from postvocalic [h] (Pulleyblank 1962; Baron, personal communication). The same origin has been reported for Vietnamese falling tone[4] (Haudricourt 1954, Matisoff 1973).

In order to evaluate the phonetic processes underlying these developments, I will first present data showing the intrinsic effects of [h] and [ʔ] on the fundamental frequency of the preceding vowel. The subjects in this experiment were four Arabic speakers without speech disorders or history of hearing pathology. They spoke six C_1VC_2 nonsense words placed within the frame 'ulu C_1VC_2 liyya' (say C_1VC_2 to me) (C_1 = {m}, V = {i,a,u} and C_2 = {h,ʔ}). The word list consisted of ten tokens of each test word arranged in random order and was read twice. Measurements of Fo were made on a PDP-12 computer after hardware pitch extraction. Starting from a reference point at the offset of the vowel, Fo values were measured every 10 msec back to a point 100 msec before the offset. The data obtained from the four subjects are presented in Fig. 1 and the results are summarized in Table 1. As seen from this table, Fo goes up a minimum of 9 Hz before [ʔ] and down 25 Hz before [h]. The two sets

---

3. Although the term tonogenesis was first used by Matisoff (c. 1970), the widespread distribution of the phenomenon was discussed in Haudricourt (1961).

4. Purcell (1974) and Purcell et al (1975) claim that post vocalic [h] can give rise to a high tone. I have discussed this proposition elsewhere (Hombert 1975b).

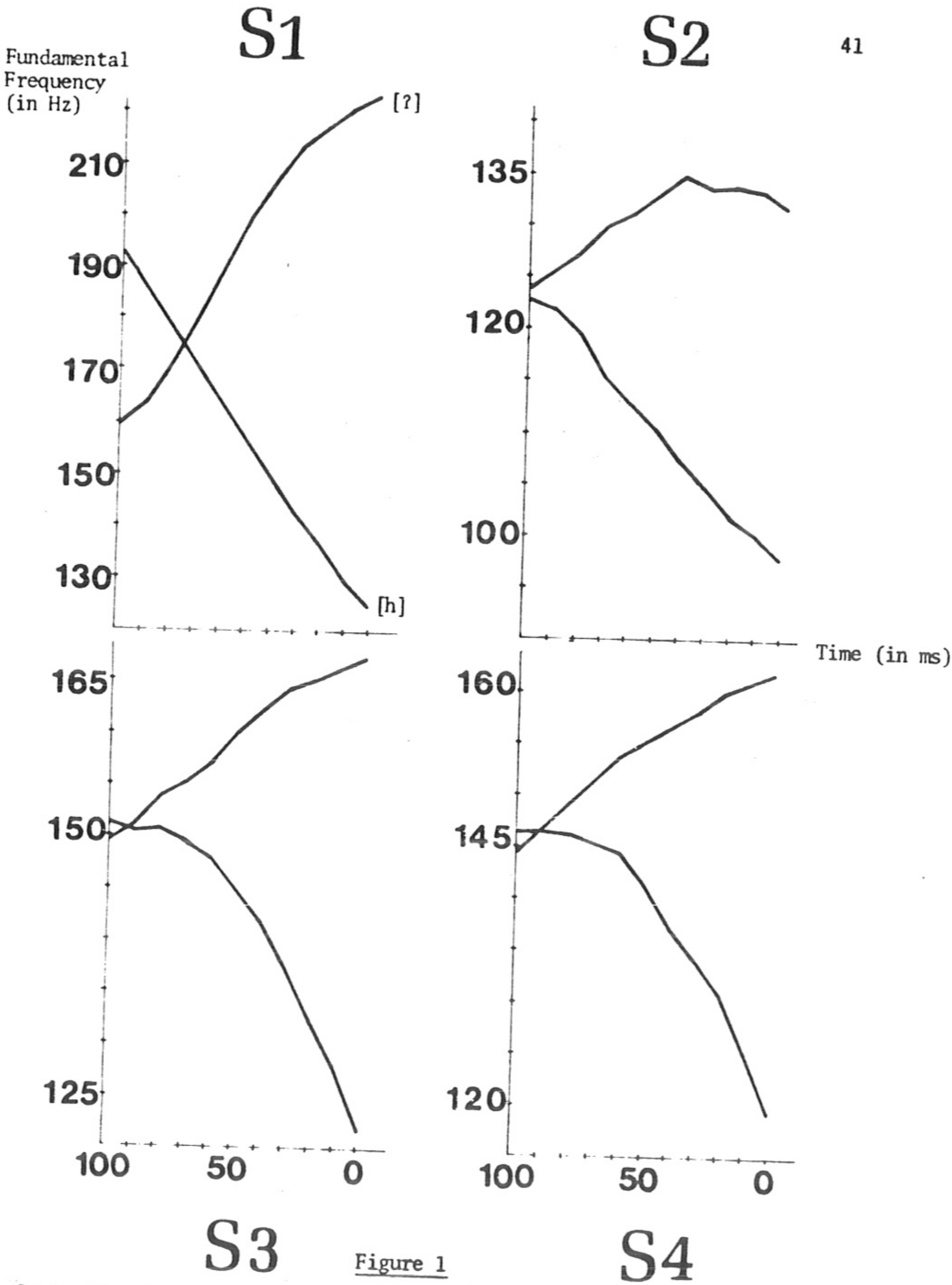

Figure 1

Production data: effect of [h] and [ʔ] on the Fo of the preceding vowel. Four Arabic speakers. For all four subjects the rising slope corresponds to the Fo perturbations caused by [ʔ]. The falling slope represents the effect of a following [h].

TABLE 1

	Δt ms	Δt_s ms	Δf_h Hz	$\Delta f_ʔ$ Hz
Subject 1	70	70	-50	+48
Subject 2	100	80	-25	+25
Subject 3	90	80	-29	+29
Subject 4	90	30	-27	+27

Δt = Time interval between the points of intersection of the two Fo curves and the vowel offset.

Δt_s = The smallest time interval at which the two curves are significantly different.

Δf_h = Fo difference between Fo offset of vowel preceding [h] and Fo value at point of intersection.

$\Delta f_ʔ$ = Fo difference between Fo offset of vowel preceding [ʔ] and Fo value at point of intersection.

of curves are significantly different at least 70 ms before vowel offset. The parameter chosen for the statistical analysis (i.e. the interval between vowel offset and intersection of Fo values) represents the most conservative estimate of statistical significance. If slope[5] of the curves would have been chosen as the parameter, the statistical significance would have been improved.

In order to determine the extent to which such Fo variations can be perceived the following experiment was conducted.[6] Ten subjects, native speakers of American English with normal hearing, participated in the experiment. The stimuli consisted of 30 instances of the vowel [i] synthesized with different fundamental frequency patterns. As shown in Figure 2, the onset frequency was 120 Hz, the frequency difference ΔF between offset and onset frequency was varied taking the following values: -50, -20, -10, +10, +20, and +70 Hz. The overall duration of the stimulus was fixed at 250 ms. The duration of the rising or falling portion of the stimulus, Δt, was varied between 40, 60, 100, 150, and 250 ms. In other words, 15 stimuli (3 Fo offset x 5 slope durations) had a falling fundamental frequency and 15 stimuli had a rising fundamental frequency. Each time the stimulus was presented, it was followed by a 500 ms pause and a second 250 ms vowel [i] with a steady-state fundamental frequency. The level of the fundamental frequency of this second vowel was adjustable by a knob controlled by the subject. Each of the 30 stimuli was presented three times in a randomized order. The stimuli were presented over earphones at a comfortable level (about 70 dB). Subjects were asked to match the steady-state Fo of the second vowel with the Fo offset of the first vowel. A statistical analysis of the results shows that the set of synthesized Fo contours simulating the effect of following [ʔ] are significantly different ($P<0.01$ for paired data and $P<0.05$ for grouped data, 10) from the Fo contours simulating the effect of following [h] when $\Delta F_ʔ = +10$ Hz and $\Delta F_h = -10$ Hz and $\Delta t = 40$ ms. For all other greater values of ΔF and Δt the two Fo patterns are perceived as significantly different ($P < 0.01$, 10).

From the Arabic data presented earlier, we concluded that an [h] produces a drop in Fo (varying from 25 to 50 Hz) on the preceding vowel while [ʔ] produces a rise in Fo (from 9 to 48 Hz). It was also shown that these two curves became significantly different at least 70 ms before vowel offset. The perceptual data presented in the preceding paragraph show that Fo perturbations similar to those caused by [h] and [ʔ] can be perceived even when these perturbations are very small ($\Delta = \pm 10$ Hz and $\Delta t = 40$ ms).

---

5. In fact, this would have been a very reasonable choice since we have neurophysiological evidence of the importance of slope detectors (Whitfield and Evans 1965, Møller 1973).

6. See Hombert (1975a) for a more detailed description of this experiment.

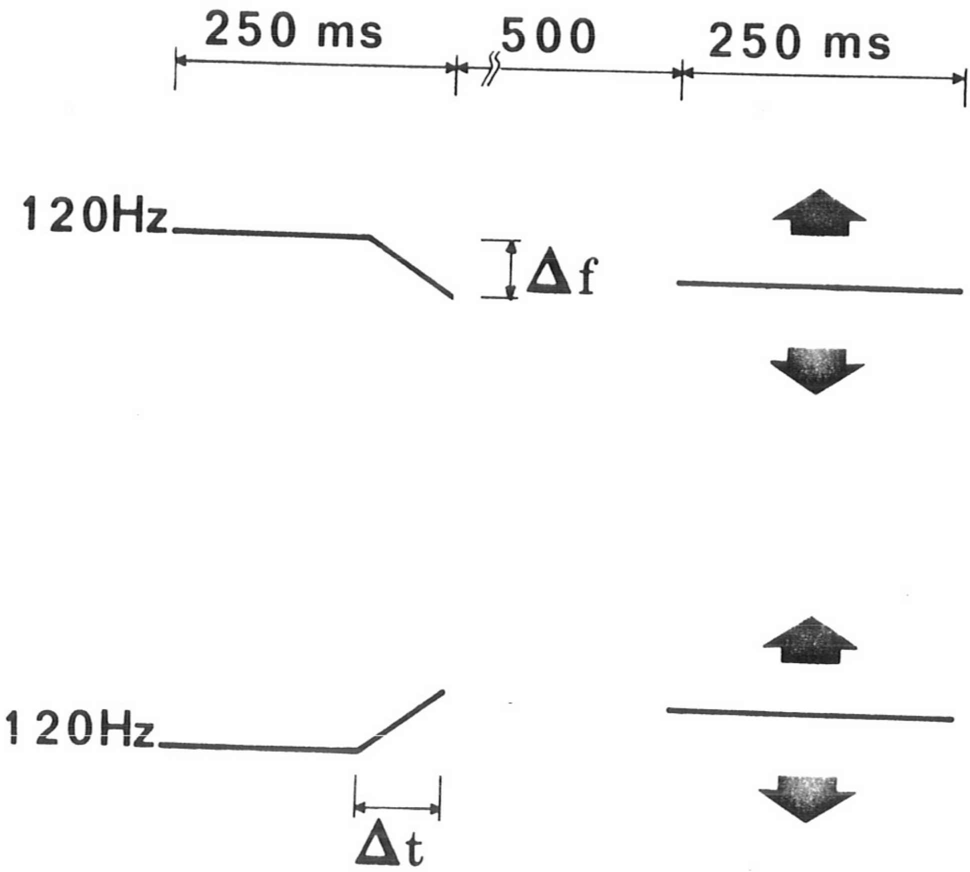

Figure 2

PERCEPTION EXPERIMENT

Stimulus presentation format

(Δt = 40, 60, 100, 150, 250 ms
Δf = -50, -20, -10, +10, +20, +70 Hz)

In summary, I have shown that:

1) [h] and [ʔ] produce significant perturbations on the Fo of the preceding vowel.

2) These perturbations can be perceived by the auditory system.

The overlap between these two sets of data validate and explain the development of rising and falling tones from the loss of [h] and [ʔ] postvocalically. The following historical stages can be proposed:[7]

Stage 1	Stage 2	Stage 3
aʔ	ǎʔ	ǎ
ah	âh	â

Stage 1: The language has a contrast between final [h] and [ʔ]; the pitch information is redundant.

Stage 2: The pitch perturbations are exaggerated and start to play a linguistic role but final consonants are still present.

Stage 3: The final consonants are lost. The pitch distinction becomes the only distinctive feature.

ACKNOWLEDGEMENTS

I am grateful to the members of the Phonology Laboratory at Berkeley and the Phonetics Laboratory at UCLA for their help and suggestions. Special thanks to V. Fromkin, S. Greenberg, R. Harshman, L. Hyman, H. Javkin, R. Krones, P. Ladefoged, M. Mazaudon, B. Michailovsky, J. Ohala and L. Rice for their helpful comments on an earlier draft of this paper. This work was partially supported by NSF Grants made to the Phonology Laboratory, Berkeley and to the Phonetics Laboratory, UCLA.

---

7. ˇ = rising tone and ˆ = falling tone.

References

Haudricourt, A. G. (1954). De l'origine des tons en Vietnamien. Journal Asiatique, 242:69-82.

Haudricourt, A. G. (1961). Bipartition et tripartition des systèmes de tons dans quelques langages d'extrême-orient. Bulletin de la Societe Linguistique de Paris, 56:163-180.

Hombert, J. M. (1974). Production and perception: their respective roles in one case of tonal development. Paper presented at the annual meeting of the Linguistic Society of America, New York, December 1974.

Hombert, J. M. (1975a). Perception of contour tones: an experimental investigation. In Proceedings of the First Annual Meeting of the Berkeley Linguistics Society, pp. 221-232.

Hombert, J. M. (1975b). Towards a theory of tonogenesis: an empirical, physiologically and perceptually-based account of the development of tonal contrasts in language. Ph.D. dissertation, U.C. Berkeley.

Maran, L. R. (1971). Burmese and Jinghpaw: a study of tonal linguistic processes. Occasional Papers of the Wolfenden Society on Tibeto-Burman Linguistics, Vol. IV.

Matisoff, J. (1970). Glottal dissimilation and the Lahu high-rising tone: a tonogenetic case-study. Journal of the American Oriental Society, 90-1:71-96.

Matisoff, J. (1973). Tonogenesis in Southeast Asia. In:Consonant Types and Tone (ed. L. Hyman), pp. 71-96.

Mazaudon, M. (1975). Tibeto-Burman tonogenetics. To appear in Current Trends in Sinotibetan Linguistics, L.R. Maran, ed.

Mei, T. L. (1970). Tones and prosody in Middle Chinese and the origin of the rising tone. Harvard Journal of Asiatic Studies 30:86-110.

Møller, A. R. (1973). Coding of amplitude and frequency modulated sounds in the cochlear nucleus. Paper presented at the Symposium on Auditory Analysis and Perception of Speech, Leningrad, August 21-24, 1973.

Ohala, J. J. (1974). Experimental historical phonology. In Historical Linguistics II. Theory and Description in Phonology, Anderson J. M. and C. Jones (eds.) (Proceedings of the First International Conference on Historical Linguistics, Edinburgh). Amsterdam:North Holland, pp. 353-389.

Ohala, J. J. (1975). Phonetic explanations for nasal sound patterns. In: Nasálfest; Papers from a Symposium on Nasals and Nasalization, C. A. Ferguson, L. M. Hyman, J. J. Ohala, eds., Language Universals Project, Stanford, pp. 289-316.

Pike, E. V. and Wistrand, K. (1974). Step-up terrace tone in Acatlan Mixtec (Mexico). In: R. M. Brend (ed.) Advances in Tagmemics. Amsterdam: North Holland Pub. Co., Pp. 81-104.

Pulleyblank, E. G. (1962). The Consonantal System of Old Chinese, Part II. A. M. 9, pp. 206-265 and 58-144.

Purcell, E.T. (1974). From dynamic-stress to word-tone: the consonantal origins of Balto-Slavic tone. In:Papers from the Tenth Regional Meeting of the Chicago Linguistics Society, La Galy et al eds., University of Chicago, Chicago, Ill.

Purcell, E., S. Young, G. Villegas, J. Cromshaw and R. Smith (1975). The Acoustics of Tonogenesis. In:Proceedings of the First Annual Meeting of the Berkeley Linguistics Society, Berkeley, Calif., pp. 366-378.

Whitfield, I. G. and Evans, E. F. (1965). Responses of auditory cortical neurons to stimuli of changing frequency. J. Neurophysiology, 28:655-72.