

Development of tones from vowel height?

Jean-Marie Hombert

Department of Linguistics, University of California, Los Angeles, U.S.A.

Received 9th February 1976

Abstract:

The development of contrastive tones on vowels due to the loss of a voicing distinction on obstruents in prevocalic position is widely attested and rather well understood. On the other hand intrinsic fundamental frequency variations caused by vowel height rarely, if at all, give rise to the development of phonological tones. Ten subjects were asked to compare the pitch of synthesized vowels different in quality but equal in fundamental frequency. The data indicate that the vowel [a] has a tendency to be judged higher in pitch than the high vowel [i] or [u]. Two possible origins of this effect are discussed. These results are taken as a partial explanation for the lack of tonal development from vowel height.

Introduction

In this paper I will attempt to provide one of the major reasons why the intrinsic fundamental frequency variations caused by vowel height, rarely if at all, give rise to the development of phonological tones. Explanations of sound changes based on physiological constraints of our articulatory and/or auditory mechanisms have been proposed for a long time (Durand, 1956; Grammont, 1933; Haden, 1938; Passy, 1891; Paul, 1909; Rousselot, 1891; Sweet, 1888). The interest in finding such explanations for widely (i.e. non-language specific) attested sound changes has been revived recently (Hombert, 1974, 1975*a, b, c*; Hombert, Ohala & Ewan, 1975, in preparation; Ohala, 1974*a, b*, 1975). These explanations imply that the pronunciation intended by the speaker may get distorted by the time it is perceived by the listener. This may be due to the action of articulatory constraints which affect the way the sounds were uttered or by the action of auditory constraints which affect the way the sounds were analyzed by the listener's ear. In order to demonstrate that a sound change is phonetically motivated, one has to demonstrate first that these intrinsic perturbations are present in the speech signal (at least in the case of articulatory motivated sound changes) and second, that their magnitudes are sufficient to be perceived.

The role played by experimental phonetics for explaining sound changes is clear from this scheme. But the interaction between phonetics and historical linguistics is not unidirectional. In this paper, I want to illustrate a case in which historical data can lead us to new insights concerning perception and speech processing. I mentioned earlier that a phonetically motivated sound change will be explained if experimental data show that intrinsic perturbations (non-intended by the speaker) are present in the speech signal and of sufficient magnitude to be heard by the listener. Now let us suppose we can demonstrate these two points despite the fact that historical data do not show that this historical change is widely attested (i.e. probably phonetically motivated). At least two

reasons can be proposed to account for this discrepancy: (1) linguists have failed to report this type of sound change perhaps because of a lack of focus in the linguistic literature on such a possibility; (2) this sound change simply did not occur. If this second case is valid, this may imply that our previous investigations of the existence and the perception of the intrinsic perturbations may have been too hasty or too superficial and consequently may have to be reconsidered.

Phonetic data

It has been shown (Hombert, 1974) that the most well documented type of tonogenesis, namely the development of contrastive tones on vowels due to the loss of a voicing distinction on obstruents in prevocalic position is phonetically motivated. In this paper I want to address myself to the following question: Why is it that such a development does not occur from the intrinsic fundamental frequencies caused by vowel height? First, let us examine data concerning the intrinsic fundamental frequencies of vowels in American English. Several studies have shown that vowels have an intrinsic fundamental frequency related to their height: high vowels (low F_1) have a higher fundamental frequency than low vowels (House & Fairbanks, 1953; Lehiste & Peterson, 1961; Peterson & Barney, 1952) as shown in table I.

Table I Intrinsic fundamental frequency of vowels (in Hz)

		i	a	u
House & Fairbanks (1953)	(10 subjects)	127.9	118.0	129.8
Lehiste & Peterson (1961)	(5 subjects)	129	120	134
Peterson & Barney (1952)	(33 subjects)	136	124	141

The data presented in Table I indicate that the intrinsic variations of fundamental frequency due to vowel height are of the same order of magnitude as the intrinsic differences caused by prevocalic consonants (Hombert, 1974; House & Fairbanks, 1953; Lea, 1972, 1973; Lehiste & Peterson, 1961; Löfqvist, 1975; Mohr, 1971). In fact, Lehiste (1970, p. 71) mentions "The influence of an initial consonant could counterbalance the influence of intrinsic pitch: the average for /kæ/ sequences was 171 Hz, while that of /gi/ sequences amounted to 170 Hz". Since vowel height and prevocalic obstruents seem to cause comparable intrinsic perturbations, one would expect tonal development resulting in vowel merging to be as frequent as tonal development resulting in the loss of some voicing contrast in prevocalic position and one would expect the development to show that high vowels give rise to high tones and low vowels to low tones. However, historical data illustrating such developments are not straightforward.

Historical data

Pilczkowa-Chodak (1972) suggests that tone assignment of verbs and noun plurals in Hausa is largely predictable from the height of the final vowel: a high (*vs* low) final vowel

predicting a high (*vs* low) tone. This analysis, however, has been criticized by Hausa scholars (Newman, 1975; Leben & Schuh, personal communications). It seems that Middle Chinese words with checked tones (i.e. p, t, k endings) and voiceless initial consonants developed a relatively lower tone when the vowel nucleus was [a] than when it was [ə] (Baron, in preparation; Pulleyblank, 1970–1). In some Cantonese dialects, this tone development has sometimes been analyzed as originating from a length contrast. In the Omei dialect of Mandarin, two tones rearranged themselves depending on vowel height; the “new” high tone regrouping high vowels (Baron, in preparation; Cheung, 1973). In Ngizim (Schuh, 1971) and in Bade, the tone patterns of verbs are partially predictable from the vowel of the first syllable; if the vowel is [a], the verb will have a high tone. The literature offers a few other examples which seem to indicate an interaction between vowel height and tone. In Foochow (Yuan *et al.* 1960; Chen & Norman, 1965; Maddieson, 1975) the vowel is raised if the tone of some lexical items is replaced by a higher tone. A similar phenomenon occurs in Lahu (Matisoff, 1973) where the rising tone can raise a vowel. These last two cases (Foochow and Lahu) can be interpreted as an effect of tones on vowels. It would seem that the interaction between tones and vowel height works only in one direction; tone can affect vowel height but not vice-versa. As I mentioned earlier, historical data do not suggest that the development of contrastive tones from vowel height is a widely attested process; furthermore, the reverse direction of interaction (i.e. low vowels giving rise to high tones) as observed in Ngizim and Bade seems inexplicable phonetically. It would seem reasonable then to find an explanation for the infrequency of this type of effect.

Several reasons can be proposed. One would expect that such a tonal development would start where the intrinsic differences are maximum, that is from the merging of low and high vowels. This seems unlikely since even very restricted vowel systems contain i, a and u. One can also argue that the intrinsic perturbations caused by prevocalic consonants are of a different nature than those caused by vowel height. A voiced (*vs* voiceless) consonant causes a rising (*vs* falling) fundamental frequency pattern at the onset of the following vowel. On the other hand, the intrinsic fundamental frequency associated with different vowel qualities is manifested by differences in steady-state fundamental frequency levels. Our auditory system is more “efficient” at detecting changes in varying fundamental frequency signals rather than differences (of the same magnitude) between two steady-state fundamental frequency signals.

Another possibility is that our perception of pitch of vowels is affected by vowel quality. Since intrinsic pitch and vowel quality are always associated for a given vowel (as opposed to the case of consonantal perturbation where the consonant can be removed) it is possible that either the pitch of vowels is affected by the spectrum envelopes (effect of loudness at certain frequencies) or that the pitch of high (*vs* low) is perceived lower (*vs* higher) as a result of some sort of normalization. The following experiment was designed to test this hypothesis.

Experimental paradigm

Ten subjects (five females and five males) all native speakers of American English participated in this experiment. They were asked to compare the pitch of two synthesized vowels (generated by a software synthesizer) of different quality. Only those who were able to achieve a score of 95% or better in a control experiment involving the comparison of the pitch of two pure tones were selected as subjects. The formant values of the three vowels used in this experiment are given in Table II.

Table II Formant values of synthesized vowels i, a, u

	i	a	u
F_1 (Hz)	270	730	300
F_2 (Hz)	2600	1090	870
F_3 (Hz)	2800	2440	2240

These values are taken from Peterson & Barney (1952) except for the second and third formants of [i] which have been raised in order to sound more speech-like (Delattre, Liberman, Cooper & Gerstman, 1952). Three fundamental frequencies, 115, 120 and 125 Hz (with an accuracy of $\pm 0.5\%$) were superimposed on each of these three vowels which had a duration of 250 ms (with a rising and decay time of 20ms). The interval separating the first (V_1) and second vowel (V_2) was 500 ms. The fundamental frequency of the second vowel was either 5 Hz below, equal to, or 5 Hz above the fundamental frequency of the first vowel (in other words, the fundamental frequency range of V_2 was from 110 to 130 Hz). V_1 and V_2 always differed in quality (i.e. there were no i-i, a-a, or u-u sequences). Six repetitions of all possible V_1 and V_2 comparisons were presented (excluding cases where V_1 and V_2 had the same quality) making a total of 324 judgements ($3V_1$ qualities \times $3V_1$ $F_0 \times 2V_2$ qualities \times $3V_2$ $F_0 \times 6$ repetitions). Overall amplitude levels were equalized for the three vowels. Subjects were asked to judge whether the first or the second vowel was higher in pitch (i.e. a two-way forced choice). They were instructed to mark the corresponding vowel on their answer sheet. They had three seconds in which to make a response. The experiment was divided into two parts preceded by a short training session in which six pairs were presented. The same stimulus presentation format was adopted in the control experiment in which synthesized vowels were replaced by pure tones. The role of this control experiment was twofold; it was used to determine the accuracy of the subjects' pitch perception and also to investigate the effect of stimulus ordering on the perception of pitch especially when the two tones had the same frequency.

Results and discussion

Since the criterion used to select subjects was quite strict (minimum of 95% correct in the control experiment), they made very few mistakes when the vowels were 5 Hz apart (90% correct or better). As a result, in the following analysis I will only consider the pairs which had identical F_0 on both vowels. The subjects' responses for these pairs are presented in Table III.

Table III Number of times each vowel was judged to be the higher in pitch when comparing the pitch of vowels of different quality but equal fundamental frequency (10 subjects)

i: a comparison (i-a and a-i pairs)	u: a comparison (u-a and a-u pairs)	u: i comparison (u-i and i-u pairs)
a = 257 i = 103	a = 261 u = 99	i = 185 u = 175

From this table, it is clear that the low vowel [a] has a tendency to be judged higher in pitch than the high vowels [i] or [u] (although their fundamental frequencies were in fact equal).

Let us now look in more detail at the pairs involving comparisons of low (a) *vs* high (i or u) vowels. In table IV, data showing the effect of fundamental frequency level and ordering for these pairs are presented.

Table IV Number of times each vowel was judged to be higher in pitch when comparing the pitch of vowels of different quality (but equal F_0) as a function of the fundamental frequency level and the order of presentation of these two vowels (10 subjects)

	i:a comparison						u:a comparison					
	i-a			a-i			u-a			a-u		
	F_0			F_0			F_0			F_0		
	115	120	125	115	120	125	115	120	125	115	120	125
Low vowel a judged higher in pitch	44	39	48	46	36	44	48	36	42	57	47	34
High vowel i or u judged higher in pitch	16	21	12	14	24	16	12	24	18	3	16	26

These data were subjected to a three-way analysis of variance. The hypothesis being tested (i.e. low vowel being judged higher in pitch than high vowel of equal fundamental frequency) was found to be statistically significant ($P < 0.0005$). The ordering effect was found to be non-significant. It can be noticed in Table IV that the differences in the number of "high" judgements for the low vowel [a] *vs* the high vowel [i] or [u] is smaller at 120 Hz and moreover the responses for the a-u pair at 115 Hz deviates considerably from the ratio found under other conditions. These two points suggest that the location of the relevant harmonics with respect to the formants may play a role in pitch perception. In fact, the statistical analysis revealed that the effect of fundamental frequency level was significant at the 0.02 level. Although this last point requires further investigation, the basic point has been demonstrated. Since the low vowel [a] is perceived higher in pitch than the high vowels [i] and [u] when compared at equal fundamental frequency, this implies that vowel quality affects pitch perception of vowels. The pitch difference between high and low vowels is smaller than their fundamental frequency differences. This is an important point for explaining why intrinsic fundamental frequency differences caused by vowel height did not give rise to tonal development as frequently as we would have expected considering the magnitude of these intrinsic differences.

These data demonstrate that the effect exists, that the perceived pitch difference between high (low F_1) and low (high F_1) vowels is smaller than their fundamental frequency difference would indicate. However, the origin of this effect has not been demonstrated. Two possibilities should be investigated:

- (1) It is a low level phenomenon which can be explained by vowel spectra characteristics
- (2) It is a higher level phenomenon involving some normalization of fundamental frequency depending on vowel height.

Although it is clear that a controlled experiment has to be done in order to accept or reject one of these hypotheses, available relevant experimental data do not seem to favor the

first possibility. It has been known for quite a long time that intensity affects our pitch perception. Stevens (1935) shows that the pitch of a tone whose frequency is 150 Hz is lowered by as much as 11 % when its intensity is raised by 50 dB. More recent data (Cohen, 1961) suggest a smaller effect (2–4 %). In order to be able to apply these data to vowels, we have to know which frequency region is the most relevant for pitch perception. Ritsma (1967) indicates that for fundamental frequencies in the range of 100–400 Hz (i.e. the relevant range for speech), the frequency region consisting of the third, fourth and fifth harmonics plays a dominant role in the perception of pitch. Since the vowels used in this experiment had a F_0 of 115, 120 or 125 Hz, the dominant frequency region is located between 345 (3×115) and 600 Hz (5×125). Since we know (Cohen, 1961; Snow 1936; Stevens, 1935; Zürmühl, 1930) that by increasing loudness we lower the pitch (assuming that we are dealing with a tone below 1000 Hz), our results would be explained if we could show that the amplitude of the spectrum of [i] or [u] is sufficiently greater than the amplitude of [a] in the relevant frequency region (i.e. approximately 345–600 Hz). In fact, as can be seen from Fig. 1, the intensity of [a] is higher than the intensity of [i] or [u] in a major part of this frequency region and consequently would not explain our results.

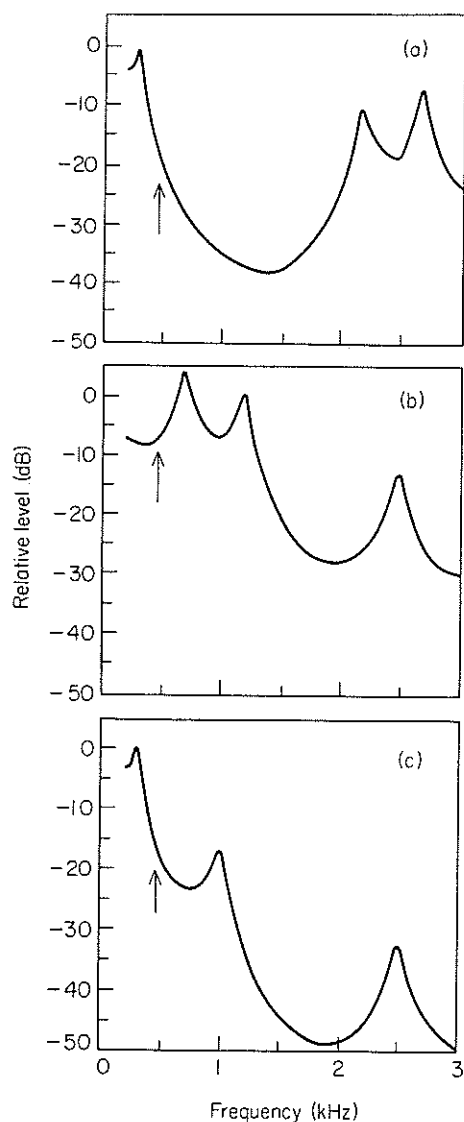


Figure 1.

Spectrum of vowels i, a, u from top to bottom (from Stevens & House, 1961). Arrows indicate dominant frequency region for pitch perception.

The second possibility would imply an interaction between F_0 and vowel quality at a higher neurological level. Since vowel quality and intrinsic F_0 are always produced simultaneously, it is possible that our auditory system subjects the speech signal to some form of normalization having the effect of raising the pitch of low vowels (or lowering the pitch of high vowels). It should be remembered that similar effects have been investigated and described; for instance, the size-weight effect by which a small size box is judged heavier than a bigger size box of equal physical weight (Birnbaum & Veit, 1974). It should be noted that this experiment indicates only the direction of the effect. The only quantitative information is that for these selected subjects, under these experimental conditions, the effect was smaller than 5 Hz. Before making definitive conclusions about the relevance of these data for explaining the non-development of tones from vowel height, two stages are required. First, a quantification of the effect under laboratory conditions (i.e. by presenting pairs of vowels differing both in quality and F_0) and second, a quantification of the relationship between laboratory conditions and "real world" conditions.

Conclusion

In summary I have shown that the study of historical developments, or more precisely in this case, the lack of historical developments, can lead to new insights concerning perception of speech. I presented data indicating that the pitch differences between high and low vowels was smaller than the corresponding intrinsic fundamental frequency differences. This is taken as a partial explanation for the lack of tonal development from vowel height.

The research reported in this paper was supported by a NSF Grant SOC75-07158, made to the UCLA Phonetics Laboratory. I wish to thank the members of the UCLA tone seminar for their valuable comments. Special thanks are owed to Peter Ladefoged, Victoria Fromkin, Hector Javkin, Steve Greenberg and Louis Goldstein, Larry Hyman, Richard Harshman and John Ohala for their criticism and discussion and also to W. Ballard, Steve Baron, Jim Matisoff, Martine Mazaudon and Russ Schuh for discussing the historical data with me.

References

- Baron, S. (in preparation). On tone and vowel height in Chinese.
- Birnbaum, M. H. & Veit, C. T. (1974). Scale-free tests of an additive model for the size-weight illusion. *Perception and Psychophysics* 16, 278-82.
- Chen, L. & Norman, J. (1965). *An introduction to the Foochow dialect*. San Francisco State College in cooperation with the U.S. Office of Education.
- Cheung, S.H. N. (1973). Tonal redistribution in the Omei dialect. *Paper presented at the 6th International Sino-Tibetan Conference, San Diego*.
- Cohen, A. (1961). Further investigation of the effects of intensity upon the pitch of pure tones. *Journal of the Acoustical Society of America* 33, 1363-76.
- Delattre, P., Liberman, A. M., Cooper, F. S. & Gerstman, L. J. (1952). An experimental study of the acoustic determinants of vowel color; observations on one- and two-formant vowels synthesized from spectrographic patterns. *Word* 8, 195-210.
- Durand, M. (1956). Du rôle de l'auditeur dans la formation des sons du langage. *J. Psychol. Norm. Pathology* 52, 347-55.
- Haden, E. F. (1938). The physiology of French consonant changes. *Language Dissertations* 26.
- 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. Also in *UCLA Working Papers in Phonetics* 33. (In press.)
- Hombert, J. M. (1975a). Towards a theory of tonogenesis: an empirical, physiologically and perceptually-based account of the development of tonal contrasts in language. Ph.D. dissertation, UC Berkeley. To be published in the *Working Papers from the Phonology Laboratory, UC Berkeley*.
- Hombert, J. M. (1975b). Perception of contour tones: an experimental investigation. In *Proceedings of the First Annual Meeting of the Berkeley Linguistics Society*. Pp. 221-32.

- Hombert, J. M. (1975c). Development of tone from segmentals: evidence from contour tone perception. In *Proceedings of the Eighth International Congress of Phonetic Sciences*. Leeds, August 1975. (In press.)
- Hombert, J. M. (in preparation). The mechanism of phonetically motivated sound changes.
- Hombert, J. M., Ohala, J. J. & Ewan, W. G. (1975). Tonogenesis: theories and queries. Paper presented at the Eighth Sino-Tibetan Conference Berkeley, October 1975.
- House, A. S. & Fairbanks, G. (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America* 25, 1. Reprinted in (Lehiste, I., Ed.) *Readings in Acoustic Phonetics*, M.I.T. Press.
- Grammont, M. (1933). *Traité de Phonétique*. Paris: Librairie Delagrave.
- Lea, W. (1972). Intonational cues to the constituent structure and phonemics of spoken English. Purdue Doctoral Dissertation.
- Lea, W. (1973). Segmental and Suprasegmental influences on fundamental frequency contours. In (Hyman, L., Ed.) *Consonant Types and Tone*. Pp. 15-70.
- Lehiste, I. (1970). *Suprasegmentals*. Cambridge Mass.: M.I.T. Press.
- Lehiste, I. & Peterson, G. (1961). Some basic considerations in the analysis of intonation. *Journal of the Acoustical Society of America* 33, 419-25.
- Löfqvist, A. (1975). Intrinsic and extrinsic F_0 variations in Swedish tonal accents. *Phonetica* 31, 228-47.
- Maddieson, I. (1975). The intrinsic pitch of vowels and tones in Foochow. Paper presented at the Fifth Californian Linguistic Association Conference, San Jose, May, 1975. Also in *UCLA Working Papers in Phonetics* 33. (In press.)
- Matisoff, J. A. (1973). The grammar of Lahus. *Publications in Linguistics* 75, Berkeley and Los Angeles: University of California Press.
- Mohr, B. (1971). Intrinsic variations in the speech signal. *Phonetica* 23, 65-93.
- Newman, P. (1975). The non-correlation of tone and vowel height in Hausa. *Studies in African Linguistics* 6, 207-13.
- Ohala, J. (1974a). Phonetic explanation in phonology. In (Bruck, A., Fox, R. A. & Lagaly, M. W., Eds) *Papers from the Parasession on Natural Phonology*. Pp. 251-74. Chicago, Ill.: Chicago Linguistic Society.
- Ohala, J. (1974b). Experimental historical phonology. In (Anderson, J. M. & Jones, C., Eds) *Historical Linguistics II. Theory and Description in Phonology*. (Proceedings of the First International Conference on Historical Linguistics, Edinburgh.) Pp. 353-89. Amsterdam.: North Holland Publishing Co.
- Ohala, J. (1975). Phonetic explanations for nasal sound patterns. Papers from the Nasalfest, Stanford.
- Passy, P. (1891). *Études sur les Changements Phonétiques*. Paris: Librairie Firmin-Didot.
- Paul, H. (1891). *Principles of the History of Language*. London.
- Paul, H. (1909). *Prinzipien der Sprachgeschichte*. Halle: Niemeyer.
- Peterson, G. E. & Barney, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America* 24, 175-84. Reprinted in (Lehiste, I., Ed.) (1967). *Readings in Acoustic Phonetics*. Cambridge, Mass.: M.I.T. Press.
- Pilszczikowa-Chodak, N. (1972). Tone-vowel height correlation and tone assignment in the patterns of verb and noun plurals in Hausa. *Studies in African Linguistics* 1, 399-422.
- Pulleyblank, E. G. (1970-1). Late middle Chinese. *Asia Major* 15, 197-239; 16, 121-68.
- Ritsma, R. (1967). Frequencies dominant in the perception of complex sounds. *Journal of the Acoustical Society of America* 42, 191-8.
- Rousselot, L'abbé (1891). *Les modifications phonétiques du langage*. Paris: H. Welter.
- Schuh, R. (1971). Toward a typology of Chadic vowel and tone system. Unpublished, UCLA.
- Snow, W. B. (1936). Change of pitch with loudness at low frequencies. *Journal of the Acoustical Society of America* 8, 14-19.
- Stevens, K. N. & House, A. S. (1961). An acoustical theory of vowel production and some of its implications. *Journal of Speech and Hearing Research* 4 303-320.
- Stevens, S. S. (1935). The relation of pitch to intensity. *Journal of the Acoustical Society of America* 6, 150-4.
- Sweet, H. (1888) *History of English Sounds*. Oxford: Clarendon Press.
- Yuan, Chia-hua et al. (1960). *Han-yu fang-yen kai-yao* (outline of Chinese Dialects). Peking.
- Zurmühl G. (1930). Abhängigkeit der Tonhöhenempfindung von der Lautstärke und ihre Beziehung zur Helmholtzschen Resonanztheorie des Hörens. *Z. Sinnesphysiol.* 61, 40-86.