Automaticity vs. feature-enhancement in the control of segmental F0

Phil Hoole[1] & Kiyoshi Honda[2]

[1]Institute of Phonetics and Speech Processing, Munich University
[2]ATR Human Information Science Labs, Kyoto

hoole@phonetik.uni-muenchen.de, honda@atr.jp

This study looks at two areas where electromyographic data from the cricothyroid muscle can be used to improve our understanding of the role of F0 in the implementation of key segmental features of both consonants and vowels.

(1) Consonant voicing.
There is some evidence (e.g Löfqvist et al., 1989) that there is increased cricothyroid activity for voiceless consonants. This may help to suppress voicing by tensing the vocal folds. Results in the literature are, however, not very extensive and not completely consistent. Thus data for an additional language (German) should extend our understanding of the realization of the voicing contrast by the motor system. However, the implications of this activity are potentially much wider. It is very well known that F0 is generally higher following voiceless than voiced consonants. Can the higher CT activity for consonant devoicing be responsible for the higher F0 on the following vowel? Radically different points of view can be found in the literature. Some (e.g Löfqvist et al.) believe that F0 differences are contingent on the articulatory adjustments for the consonant. Others (e.g Kingston & Diehl, 1994) believe that the F0 differences on the vowel are independent of the specific laryngeal adjustments for the consonant, and have to be related to a more abstractly conceived [voice] feature, in which the speaker aims to enhance an auditorily defined property such as low-frequency emphasis in the [+voice] cognates.
Results: In /C_1 V C_2 ə/ utterances with stress on the first syllable we found clear evidence of higher CT activity on voiceless C1. Whether CT was higher on the vowel itself in voiceless contexts was more variable over speakers; nevertheless, the F0 differences on the vowel were, as expected, completely clear-cut.
We then developed the following scenario to account for F0 on the vowel in terms of CT differences on the consonant: Firstly, we contend that it is probably not the main purpose of the CT activity to suppress voicing. If this were the case, then CT activity in voiceless consonants would be expected to be most in evidence at consonant ***onset***. But then, as Kingston has pointed out, it would most likely be too far away to affect the following vowel. In fact, however, we estimate that the CT activity proceeds more or less in parallel with the glottal abduction/adduction gesture, and speculate that its purpose is not to suppress vocal fold vibration as such, but rather to increase the mechanical efficiency of the abductory movement of the arytenoids.
A further consideration that moves the locus of CT effects from consonant onset towards the following vowel was introduced by Hanson & Stevens (2002). They estimate that increased vocal fold tension at the transition from voicelessness to voicing (at least in aspirated plosives) may help prevent the vocal folds from vibrating while the arytenoids are still abducted - and thus help to achieve a well-marked transition from aspiration noise to modal voicing.
A final consideration that can help to explain why CT differences on the consonant can particularly affect the ***following*** vowel is that evidence was found in our data (with some support from the literature) that the effect of an increase in CT on F0 may be more rapid than the effect of relaxation. Thus our basic scenario suggests that the driving force for voicing contrast effects on F0 does come from the contingencies of consonantal articulation, and not from active enhancement of an auditory property. However, overall we in fact favour a hybrid scenario: Some speakers do show pronounced

CT differences right through the vowel. Thus, once such an articulatorily-based pattern is up and running speakers have the possibility of actively enhancing it as part of their linguistic behaviour.

(2) Vowel intrinsic pitch in vowel-height and tense-lax contrasts

The basic questions are very similar: Is intrinsic pitch explainable as a mechanical effect of vowel articulation (see e.g. Whalen & Levitt, 1995), or does it represent an active enhancement of an auditory property such as F1-F0 distance to support vowel-height contrasts (see Kingston, 1992)? Once again, we believe that a hybrid scenario turns out to be the most persuasive.

This area will in turn be divided into two sub-areas: Firstly, the specific case of intrinsic pitch in the tense-lax opposition in German vowels, secondly the more general intrinsic pitch case of high vs. low vowels.

Intrinsic pitch in German has often been quoted as a potential problem for mechanical theories, since tense-lax vowel pairs tend to have very similar F0 despite substantial differences in tongue position. A mechanical explanation would remain viable, however, if it could be shown that speakers actively raise F0 on the lax vowel series. Analysis of our CT data indicated that this could indeed be the case. In particular, regression analysis of the relationship between CT and F0 indicated that tense and lax vowels are characterized by different regression lines, such that at a given level of CT activity F0 is lower for lax vowels. This can be interpreted as revealing the presence of a mechanical effect on F0 that is overlaid in this specific case by active articulatory adjustments. Thus, on the one hand, the presence of biomechanical effects of the kind often assumed is confirmed; on the other hand, active manipulation of F0 as part of what would traditionally be regarded as segmental distinctions is clearly a control strategy that is available to speakers.

The second part of the intrinsic pitch analysis points in a similar direction. Here we examined the basic intrinsic pitch finding of differences between high and low vowels. The key analysis technique was again regression analysis of F0 vs. EMG. This revealed even more clearly than the tense-lax data the presence of a mechanical effect: At a given level of CT activity, F0 is lower for low vowels than high vowels. This was completely consistent over speakers, as indeed ought to be the case if it emerges as an automatic consequence of vowel articulation. In addition, however, it was observed that some speakers typically showed higher EMG activity for the higher vowels. Thus, once again a hybrid scenario appears the most balanced one: The driving force behind intrinsic pitch differences on vowels emerges from the articulatory contingencies of vowel production; this physically given bias may then open up an articulatory strategy that speakers can, but do not necessarily have to adopt for enhancing linguistic distinctions (cf. Honda & Fujimura, 1991).

References

Hanson, H.M. & Stevens, K.N. (2002). "A quasiarticulatory approach to controlling acoustic source parameters in a Klatt-type formant synthesizer using HLsyn". J. Acoustical Soc. America, 112(I), 1158-1182.

Honda, K. & Fujimura, O. (1991). "Intrinsic vowel F0 and Phrase-final F0 lowering: Phonological vs. Biological Explanations". In: Gauffin, J. & Hammerberg, B. (eds) Vocal Fold Physiology, pp. 149-157.

Kingston, J. (1992). "The phonetics and phonology of perceptually motivated articulatory covariation". Language and Speech, 35, 99-113.

Kingston, J. & Diehl, R. (1994). "Phonetic knowledge". Language, 70, 419-454.

Löfqvist, A., Baer, T., McGarr, N. S. & Story, R. S. (1989). "The cricothyroid muscle in voicing control". J. Acoust. Soc. Am., 85, 1314-1321.

Whalen, D.H. & Levitt, A. G. (1995). "The universality of intrinsic F0 of vowels". J. Phonetics, 23(3), 349-366.