EMG evidence for the automaticity of intrinsic F0 of vowels

D. H. Whalen*, Bryan Gick†, Masanobu Kumada*, and Kiyoshi Honda‡

*Haskins Laboratories, 270 Crown Street, New Haven, CT 06511-6695 USA
†Dept. of Linguistics, Yale University, New Haven, CT 06520 USA
‡ATR Human Information Processing Research Laboratories, 2-2 Hikaridai Seika-cho Soraku-gun, Kyoto, Japan

Abstract: Although vowels can be produced with many F0 values, FOs of high vowels tend to be higher than of low vowels, a tendency found for every language examined. Nonetheless, this "intrinsic FO" (or F0) has been called a deliberate "enhancement" of the speech signal, aiding vowel height perception. Since F0 remains constant with the number of vowels in a language, this enhancement seems unlikely. The only positive evidence for it has been that activity of the cricothyroid (CT) muscle, which mostly raises FO, is greater for high vowels than for low. The present experiment explores this by having subjects produce vowels at slightly different FOs, with the differences being similar to the F0 for each subject. Our subjects showed no vowel height difference in CT; further, the slopes of FO by CT plots differ for high vs. low vowels, indicating that different levels are needed to effect an F0 change. Even the EMG evidence, then, makes F0 an unlikely speech enhancement. Rather, F0 is likely to be an automatic effect of vowel production.

INTRODUCTION

Although the vowels of a language can be produced with a variety of F0 values, high vowels tend to have a higher F0 on average (by about 15 Hz) than low vowels. This tendency has been found for every language that has been examined for it, and the size of the effect does not seem to differ across languages [1]. Nonetheless, it has been proposed that this "intrinsic F0" (or F0) of vowels is a deliberate "enhancement" of the speech signal, one that must be specified in the phonology of the language [2, 3]. On this account, the perception of vowel height should be based on the difference between F1 and F0. Thus having a high F0 with the high vowels will make this difference even smaller, and the difference for low vowels correspondingly larger, making vowel height more easily perceived. Since F0 does not differ in magnitude based on the number of vowels a language has, this enhancement seems unlikely. The only positive evidence in its favor has been the finding that the activity of the cricothyroid (CT) muscle, which primarily raises F0, is larger for high vowels than for low vowels, just as we would expect if F0 were being deliberately controlled[4-7]. The present experiment explores this muscle activity by having subjects produce vowels at slightly different F0s, with the differences in F0 being equivalent to the F0 differences for that subject.

METHOD

We investigated the activity of the crico-thyroid (CT) muscle in four subjects, two female and two male. Hookwire electrodes were inserted into the CT and the placement checkd by comparing the effects of pitch changes, swallowing, opening the jaw, and pushing the head against resistance. Bilateral insertions were attempted for all subjects, but only M1 had successful signals on both sides.

Changes in F0 were obtained by having the subjects match the pitch of a tone presented aurally to them. These tones were generated for each subject based on his or her natural range of F0 values for the vowels /a/ and /u/ produced in isolation. The means for /a/ served as one value, and the average of the means of /i/ and /u/ served as the other. The difference between these two was the step size for going three steps up and one step down for each vowel. Thus each vowel to be produced had five target F0 values, with the ranges overlapping except for the low value for /a/ and the high values for /i/ and /u/ (see Table 1). (All F0 values are in Hz; all CT values are in microvolts.)

| TABLE 1. F0 values in production and test targets. |
|-------------------|---------------|---------------|
| Subject | Production | Target Values |
| | | /a/ | /i/ | /a/ | /i/ | /a/ | /i/ | /a/ | /i/ | /a/ | /i/ |
|-------------------|---------------|---------------|
| M1 | 117 | 123 | 111 | 117 | 123 | 129 | 135 | 141 |
| M2 | 100 | 104 | 96 | 100 | 104 | 108 | 112 | 116 |
| F1 | 155 | 160 | 150 | 155 | 160 | 165 | 170 | 175 |
| F2 | 192 | 204 | 180 | 192 | 204 | 216 | 228 | 240 |

Subjects were instructed to match the tone at the beginning of their utterance (again an isolated vowel), and then to drop their F0 as if they were saying a declarative sentence. These instructions were intended to keep the subjects in a speech mode, rather than falling into a singing mode. Singing appears to recruit muscles in ways that are unusual for speech, and we are concerned with how the CT behaves both in the speech range and mode. Fifteen repetitions of each vowel spoken in response to each target tone were collected in pseudorandom order.

All subjects except M1 also produced fifteen repetitions of the vowels at their normal speaking level.
Analysis was performed both on the F0 of the vowel produced and the CT activity leading up to that F0. F0 was measured at the beginning of the syllable by an autocorrelation function. CT activity was measured as the RMS amplitude over the 250 ms preceding vowel onset. This window is larger than has been used in other studies [6], but was necessary to include the significant portion of the EMG burst, especially for the female talkers.

RESULTS

In the isolated condition, with no target F0, all three subjects measured again showed IF0, but the difference in CT activity did not replicate. In only one subject did a high vowel show greater activity (F2's /i/) but that same subject showed lower activity for the other high vowel. (Two subjects also produced higher F0s (about 14%) than they did in the pretest, for unknown reasons.)

| TABLE 2. F0 values and CT activity for isolated vowels (no F0 targets). |
|---------------------------|----------------|----------------|----------------|----------------|
| Subject: | a | i | u |
| F0 | CT | F0 | CT | F0 | CT |
| M2 | 112 | 26.4 | 120 | 26.1 | 124 | 25.9 |
| F1 | 176 | 20.8 | 180 | 20.7 | 180 | 20.4 |
| F2 | 192 | 13.2 | 203 | 15.1 | 198 | 10.9 |

All four subjects were significantly accurate at reproducing the target F0s (correlations of .63, .69, .38 and .88 for M1, M2, F1 and F2 respectively). Analyses comparing F0 and CT activity were made on the actual F0, not the target F0.

Three of the four subjects increased CT activity with higher F0, as expected. For these subjects, there was a higher slope for function relating CT and F0 for the high vowels than there was for the low vowels, indicating that more CT activity was required to achieve the same change in F0 (see Table 3). Speaker F1 had a negative correlation; it may be that she habitually uses the lower portion of her F0 range and so recruits her muscles in an unusual way.

| TABLE 3. Slope CT activity vs. F0. |
|---------------------------|----------------|----------------|----------------|
| Subject: | a | i | u |
| Left | Right | Left | Right | Left | Right |
| M1 | 1.44 | 1.47 | 1.94 | 2.22 | 2.56 | 3.43 |
| M2 | - .57 | 2.83 | 2.49 |
| F1 | - .99 | - .68 | - 1.19 |
| F2 | 2.82 | 4.27 | 4.09 |

DISCUSSION

IF0 does not seem to be a deliberate enhancement of the speech signal. CT activity is not universally higher for high vowels, as seen in the present subjects. Further, the amount of CT activity needed to effect a change in F0 differs for the different vowels, making comparisons of absolute CT activity levels meaningless. Whatever may ultimately be found to account for IF0, IF0 appears to be an automatic consequence of successful vowel articulation.

ACKNOWLEDGMENTS

The authors wish to thank Anders Löfqvist and Donald S. Hailey for their assistance. This work was supported by NIH grant DC-02717.

REFERENCES