The relationship between fundamental frequency and vowel quality

Santiago Barreda & Terrance M. Nearey

Social Sciences and Humanities Research Council of Canada | Conseil de recherches en sciences humaines du Canada | Canada

Introduction

The relationship between fundamental frequency (f0) and vowel quality may be:

- Direct, in the same way the formants affect vowel quality. [1,2]
- Indirect, in that it might affect vowel quality only by changing the listener's impression of the speaker. [3,4]
- Nonexistent, the two may be uncorrelated. [5,6]

Objective: to investigate the relationship between intrinsic f0, perceived vowel quality and speaker judgments. Does f0 affect vowel quality directly or does it do so by affecting the assumed speaker.

What's new here: Simultaneous judgements on a stimulus-by-stimulus basis of vowel quality and assumed speaker traits.

Methods

Participants: 19 native English speakers.

Stimuli: A seven-step F1/F2 /ʌ/-/æ/ continuum fully crossed with 3 f0 and 3 F3 steps. 63 stimuli in total. The continuum went from least to most open. F1 and F2 increased at equal rates and were perfectly correlated.

	f0 Levels		
	Low	Mid.	High
Mean	108	153	215

F3 Levels (log Hz)		
Low	Mid.	High
2475	2755	3068

Step #	1	2	3	4	5	6	7
F1	684	735	789	848	911	978	1051
F2	1354	1455	1563	1679	1803	1937	2081

Table 1 *– Stimuli formant and f0 frequencies.*

Procedure: Participants were presented with a vowel and simultaneously reported vowel quality by clicking on the horizontal vowel response bar, and speaker size and gender by using one of the two vertical size bars. Size and vowel quality responses were continuous within category. Participants responded to each stimuli 6 times.

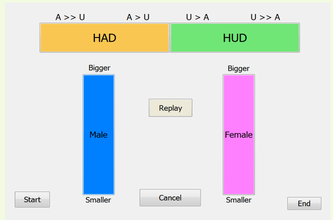

Figure 1 *– Screenshot of experiment layout.*

Results

Indirect normalization and frame of reference: Open vowels occur with F1 frequencies near a speaker's maximum F1. A speaker with larger vocal tract (VT) has lower maximum F1 than a speaker with a shorter VT. If interpreted as coming from a speaker with a larger VT, a stimuli with an intermediate F1 will be nearer to the that speaker's max F1 and hence sound more open. A summary of related predictions follows:

- **Higher F1/F2 (= higher step number) => More open**
- **Higher F3 => Shorter VT => Less open**
- **Higher f0 => Shorter VT => Less open**
- **Higher formants/f0 => Less Likely to be Male**
- **Higher formants/f0 => Smaller Size**
- **Bigger Size => Longer VT => More open**
- **Male => Longer VT => More open**

6,921 responses were collected in total. Listeners reported hearing a male speaker 65.7% of the time overall.

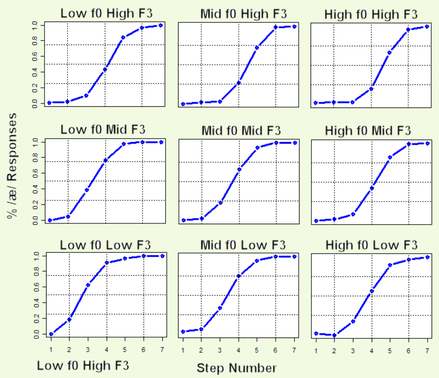

Figure 2 *– Classification functions divided by f0 and F3 level, pooled across all listeners. Step numbers have identical F1/F2 in all conditions*

- Vowels in the same row in Figure 2 differ only in terms of f0, those in the same column differ only in terms of F3.
- It's clear that both f0 and F3 are having an effect on the classification functions.

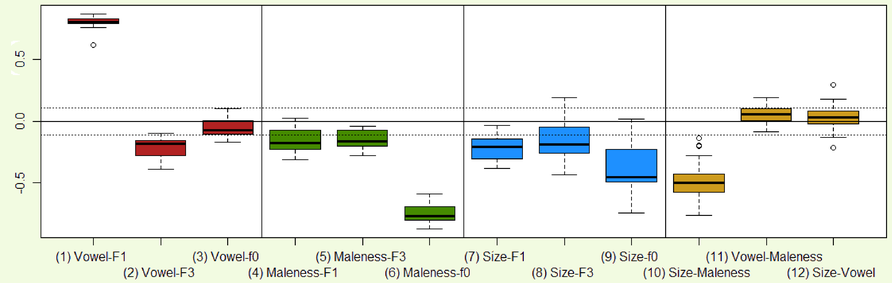

Figure 3 *– Boxplot of partial correlation coefficients across all participants for each pair of variables considered. Vowel openness is encoded as a positive value. Dotted lines represent bounds at which an individual participant's coefficient reaches significance ($p < .05$)*

	1	2	3	4	5	6	7	8	9	10	11	12
Mean	0.802	-0.216	-0.053	-0.152	-0.148	-0.744	-0.212	-0.153	-0.374	-0.476	0.049	0.027
t value	64.34	-10.94	-3.03	-6.49	-9.04	-41.00	-9.18	-4.05	-8.59	-12.18	2.98	1.04
p value	0	0	0.007	0	0	0	0	0.001	0	0	0.008	0.314

Table 2 *– Results of t-tests for the correlation coefficients plotted in Figure 3.*

Partial Correlations: We looked for correlations between pairs of variables after controlling for all of the remaining variables. We did this for: F1/F2, F3, f0, Size, Maleness and Openness for each listener individually.

We performed between-participants t-tests on the resulting partial correlation coefficients to see which pairs of relationships were significantly different from zero, on average, across all listeners.

- F1/F2 and F3 show the expected relationship with openness.
- f0 has a weak but persistent effect on openness, independently of any other variable.

- F1/F2, F3 and f0 show the expected relationship with Maleness. Lower frequencies are associated with a male speaker and, presumably, a longer VTL.

- F1/F2, F3 and f0 also show the expected relationship with speaker size. Higher frequencies are associated with a smaller speaker and, presumably, a shorter VTL.

- Males overall are associated with smaller size. This is because where listeners might have heard adult females, they more often than not heard young males.
- Despite the negative association between size and maleness, maleness consistently affects vowel quality in the expected direction after controlling for size.
- Size is positively correlated with openness.
- The size-openness relationship is the only one to not reach significance. However, 14 of 19 participants showed a correlation between these two variables in the expected direction.

Figure 4 presents the average correlations across all listeners. The direction of the arrows indicate the presumed direction of the effect. The broken arrow indicates the only relationship to not reach significance.

All of the stimulus variables have both direct and indirect effects on vowel quality.

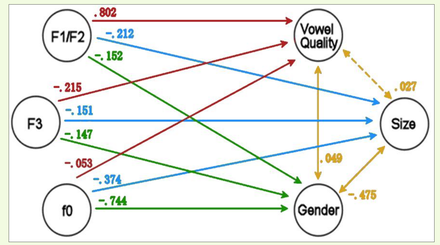

Figure 4 *– Average partial correlations for pairs of variables across all participants. Effects are direct (red), indirect (green/blue) and those resulting from speaker changes (brown).*

Conclusion

Although f0 is clearly having an effect on vowel quality (there is a 17% increase in the number of /æ/ identifications from the low to the high f0 condition), its average direct effect (Table 2, #3), after controlling for all other variables is not particularly strong.

However, it does strongly affect assumed speaker size and gender (Figure 4, in green/blue), which themselves also exert influence on vowel quality (Figure 4, in brown).

The formants, on the other hand, have a relatively stronger effect on vowel quality directly, although they also affect assumed speaker traits.

In Summary:

- f0 has a strong cumulative effect on vowel quality.
- This effect is both direct AND indirect.
- The direct effect, after controlling for assumed speaker traits, is relatively weak.
- The indirect effect, via the strong influence it exerts on assumed speaker traits, is the more significant effect.
- Results confirm Johnson's (1990a, 1990b) hypothesis that f0 affects vowel quality mainly by affecting assumed speaker traits.

References

[1] Syrdal, A. K. & Gopal, H. S.. (1986) . A perceptual model of vowel recognition based on the auditory representation of American English vowels. *Journal of the Acoustical Society of America* 79: 1086-1100.
[2] Miller, J. D.. (1989). Auditory-perceptual interpretation of the vowel. *Journal of the Acoustical Society of America* 85: 2114-2134.
[3] Johnson, K. (1990a) The role of perceived speaker identity in F0 normalization of vowels. *Journal of the Acoustical Society of America*. 88: 642-654.
[4] Johnson, K. (1990b) Contrast and normalization in vowel perception. *Journal of Phonetics*. 18: 229-254.
[5] Smith, David R. R. & Roy D. Patterson. (2005). The interaction of glottal-pulse rate and vocal-tract length in judgements of speaker size, sex, and age. Journal of the Acoustical Society of America 118: 3177-3186.
[6] Smith, D. R. R., Patterson, R. D., Turner, R., Kawahara, H. & Toshio Irino. (2005). The processing and perception of size information in speech sounds. *Journal of the Acoustical Society of America* 117: 305-318.

Contact Information

Santiago Barreda (sbarreda@ualberta.ca)
Terrance M. Nearey (t.nearey@ualberta.ca)
Department of Linguistics
4-32 Assiniboia Hall, University of Alberta,
Edmonton, Alberta, Canada T6G 2E7