Influence of Environmental Intensity on Intrinsic f_0

S. Ng and T. Wada

1 February, 2018

Abstract

In this study we examine the relationship between fundamental frequency, intensity, and vowel quality (height). We use background/environmental noise as a constraint on speaker intensity, prompting speakers to adjust their intensity to the environment. Based on our collected data, we cannot affirm that the intensity of the speech environment (and the way that speakers accommodate this noise) impacts the If_0 correlation.

1   Purpose

Intrinsic f_0 (If_0) is considered a physiologic property of general speech (Whalen and Levitt 1995). However, not all general properties of speech hold in all speech environments. For example, we posit voiced and unvoiced obstruents in English, but then in whispered speech there is no distinction.

We know that besides practical purposes of intensity adjustment (i.e. being able to be heard in a loud environment), intensity also serves as a marker for linguistic behavior (Fry 1954).

According to Fry, intensity has proven to be problematic in linguistic research in the past. Linguistic stress has been correlated with intensity in the past, however this has been disputed by later experimental testing (Mol and Uhlenbeck 1955). The impact of intensity on continuous speech production is not something that has a great deal of prior research. Intensity is normally the means by which we make ourselves more audible in noisy environments. Previous observations on speech intensity have primarily been on stressed syllables rather than on continuous speech at a higher intensity (Sluijter and Heuven 1996).

The goal of this experiment is to identify any interactions of If_0 and F1 when speakers are in environments of varying background noise. By replicating environments where a speaker will alter their speech in order to compensate for noisy environments we will gather data that may demonstrate the impact that higher intensity may have on speech overall.

2   Hypothesis

H_0: There is no significant effect between environmental intensity and the relationship between vowel height and fundamental frequency.

H_1: In environments where speakers experience background noise, the correlation of vowel height and fundamental frequency will have a higher coefficient.

H_2: In environments where speakers experience background noise, the correlation of vowel height and fundamental frequency will have a lower coefficient.

3   Procedure

3.1   Experimental Design

In order to elicit speech samples at different frequencies without altering any performative value of the speech act, we chose to introduce environmental noise in the form of background audio clips. Sample audio clips were public domain and procured from www.soundbible.com.

We determined 3 levels of intensity for testing: silence (i.e. only existing ambient noise), moderate ambient noise (avg. 60dB) and loud ambient noise (avg 80dB). These intensity levels are based on perfunctory environmental barriers in speech processing (Linsday and Norman 2013) For both the moderate and loud noise categories, two sound clips were obtained: one containing non-linguistic noise (e.g. rainfall) and one containing linguistic noise

Table 1: Recorded Vowels

Vowel	Word	Elicitation
i	heat	
u	hoot	
ɪ	hit	
ʊ	hʊt	
o	oat	
ɛ	et	"Say the word _____ again."
ə	hut	
ɚ	hurt	
ɔ	aught	
æ	hat	
ɑ	hot	

(e.g. chatter in a restaurant). These audio clips were looped to obtain an appropriate length (> 1.5 min), and then scaled for intensity in Praat.

Stimulus sentences were designed to elicit a full range of vowels in minimally contrastive environments. The words chosen begin with voiceless glottal fricatives and end with voiceless alveolar stops where possible. One nonce word was included in the set ([hʊt]), however it is a plausible English construction. The set is given in Table 1. The order of items on the reading list was randomized for each background noise environment. The environmental noise was simulated through playback on Sony MDR-7506 headphones and audio impedance was manually adjusted. and played throughout the reading task. The goal of the simulated noise was to influence the speaker into producing speech in the same manner as real-world noisy environments.

The speaker was a male, native English speaker in his 20s. Speech samples were recorded on a MicroMic C520 Vocal Condenser Microphone (head-mounted style) with a Zoom H46 Recording device. Audio files were sampled at 44.1 kHz and encoded in 16 bit. The recording space (i.e. the silence environment) was a quiet room with minimal fluctuation in ambient noise.

All files (sound files, elicitations, etc.) can be found at `github.com/SaraBlalockNg/553-labs/1.`

3.2   Analysis

Speech samples were analyzed using Praat. Researchers manually extracted the vowel space from the sample. The beginning of the vowel space was defined as the zero crossing nearest to the start of periodicity of the vowel, erring on the conservative side. The completion of the vowel was defined as the point in the spectrogram when the higher formants (especially F2 and F3) experienced sharp loss in energy.

Using these heuristics, the files were anotated via TextGrid. The f_0 and F1 were calculated at the midpoint of each vowel noted in the TextGrid using a Praat script developed by Mietta Lennes and modified by Dan McCloy (see Lennes and McCloy (2003) for the original

scripting). The modified version can be found at `https://depts.washington.edu/phonlab/resources/getDurationPitchFormants.praat`.

The metric by which we measured If_0 was the arithmetic difference of normalized F1 values and normalized fundamental frequency.

4 Results

Z-scores were calculated based on this metric, given in Table 2. These scores represent deviation from the mean for a given vowel.

Table 2: Z-scores for Vowel If_0

		Environmental Noise				
		None Added	Rain (60dB)	Chatter (60dB)	Siren (80dB)	Stadium (80dB)
Vowel	i	1.1617064	0.49615936	0.7293767	-1.30041654	-1.08682592
	u	-0.71446495	1.06776599	-0.92180991	1.36547618	-0.79696731
	ɪ	1.11560008	0.85681526	0.40384381	-1.32024782	-1.05601133
	ʊ	0.09453912	-0.34041282	0.49406052	1.39252904	-1.64071586
	o	0.8549252	-1.58346887	1.24101464	-0.4687192	-0.04375177
	ε	1.55843828	0.32527953	0.1767649	-0.86442303	-1.40972668
	ə	1.18038547	0.91699935	0.6869689	-1.90505712	-0.1122433
	ɚ	0.69106531	0.6392662	-0.12527302	-0.20702227	-1.55142251
	ɔ	1.09745698	-0.82235959	0.82123286	-1.50835247	0.41202222
	æ	1.24310294	-0.10052823	-1.55621331	0.89137786	-0.47773925
	ɑ	1.85143356	-0.08641264	-0.16354847	-1.15824253	-0.44322992

The vowels are arranged by approximate height for convenience. Across vowels, the control environment performed on average 0.92128985 standard deviations above the mean. This tells us that the correlation between f_0 and F1 is to some degree more controlled/consistent when there is not environmental noise. However, the data cannot support any difference between high and low vowels in this relationship, nor can it support any gradient effect on F1.

5 Conclusion

We set out to determine whether environmental intensity would effect the attested If_0 of vowels in English. Unfortunately, our does not imply any sort of pattern in this respect, and we cannot reject the null hypothesis. If we were to repeat an experiment of this type in the future, we believe it would be worthwhile to consider whether instructions should be explicitly provided to the speaker to try and talk over the environmental stimuli, or whether stimuli are necessary/useful in iliciting different intensities of speech.

References

Fry, D. B. (1954). Duration and Intensity as Physical Correlates of Linguistic Stress. The Journal of the Acoustical Society of America, 26(1), 138-138. doi:10.1121/1.1917773

Lennes, M., & McCloy, D. (2003). Collect_formant_data_from_files. praat.

Lindsay, P. H., & Norman, D. A. (2013). Human information processing: An introduction to psychology. Academic press.

Mol, H., & Uhlenbeck, E. (1955). The linguistic relevance of intensity in stress. Lingua, 5, 205-213. doi:10.1016/0024-3841(55)90010-3

Sluijter, A. M., & Heuven, V. J. (1996). Spectral balance as an acoustic correlate of linguistic stress. The Journal of the Acoustical Society of America, 100(4), 2471-2485. doi:10.1121/1.417955

Whalen, D. H., & Levitt, A. G. (1995). The universality of intrinsic F0 of vowels. Journal of phonetics, 23(3), 349-366. Chicago