THE RELATIONSHIP BETWEEN GENDER IDENTITY AND SIX F0 MEASURES IN POLISH

Kamil Kaźmierski

Faculty of English, Adam Mickiewicz University in Poznań
kamil.kazmierski@wa.amu.edu.pl

ABSTRACT

This paper reports on a study of the relationship between gender identity, as measured by the Bem Sex-Role Inventory (BSRI) and six characteristics of F0 (mean, median, minimum, maximum, range and standard deviation) in semi-spontaneous speech of 24 Polish female speakers.

Despite ample evidence suggesting that gender identity exerts an influence on voice quality, notably on pitch, whose main acoustic correlate is F0, the statistical analysis of the data gathered for the present study does not attest to a relationship between the BSRI score and F0 measures. An elaboration of culture-specific tool for measuring gender identity, as well as supplementing acoustic measurements with physiological and articulatory data are suggested as methodological developments needed to verify whether there are links between gender identity and F0 in Polish.

Keywords: F0, gender identity, Polish

1. INTRODUCTION

Despite clearly being influenced by physiological factors [14], individual characteristics of a person's voice have been shown to be co-determined by learning. Developmental studies suggest that gender-specific formant frequencies start to transpire before the relevant sex-typed physiological differences set in [16]. Cross-linguistic differences in F0 measures [9, 10] including inter-linguistic differences in bilinguals [15] suggest that there is an acquired component to F0 characteristics. Previous research has also found links between gender identity and voice quality. Sexual orientation has been found to influence vowel production [11] and F0 [3]. Further, it has been found [2] that speakers hailing from societies with different cultural norms about masculinity and femininity exhibit different F0 values, and that some self-reported gender-typed traits and behaviors correlate with F0 measures [4]. To verify whether gender-identity influences voice quality of Polish speakers, the present study investigated the relationship between cultural gender as measured by the Bem Sex-Role Inventory and six F0 measures of female Polish speakers during semi-spontaneous speech.

2. METHODOLOGY

2.1. Recordings

2.1.1.Procedure

The recordings took place in the anechoic chamber of the Centre for Speech and Language Processing at Adam Mickiewicz University in Poznań. This ensured high sound quality needed for acoustic analysis and invariability of situational factors. The speakers were 25 female undergraduate students between the ages of 19 and 22, all native speakers of Polish.

With every participant, a conversation was conducted according to a script inspired by the classic sociolinguistic interview method [8, 13] to elicit speech as close to the vernacular as possible, given the constraints of the setting. The same set of questions, relating to high school life and entering college was used with every participant, though not necessarily in the same order. All conversations lasted at least 10 minutes.

2.1.2. Analysis: F0 measurements

The acoustic analysis was conducted in *PRAAT* [6]. The first minute of every recording was discarded as it tended to include a good deal of preparatory talk. Starting from a point 60 seconds into each recording, all turns of the interviewee that lasted between 10 and 30 minutes were selected. Out of this pool, intervals were selected at random until their total length has reached or exceeded 120 seconds. If a randomly selected interval contained a non-speech sound such as laughter, loud in- or exhalation, or if it contained overlap, it was discarded, unless the non-speech sound was at the very beginning or at the tail-end of the interval allowing the shortening of the interval and the exclusion of the non-speech sound. Hesitation noises were not treated as non-speech sounds. The resulting combined intervals for each speaker ranged from 120 to 140 seconds.

The combined intervals were then analyzed with respect to F0 by means of the autocorrelation method [5]. The pitch range was set to 100Hz – 500 Hz, and the unit selected was semitones with the reference point of 1Hz. For every speaker, six F0 measures were taken: average F0, median F0, minimum F0, maximum F0, F0 range and standard deviation. Average F0 is given in *PRAAT* automatically when pitch analysis is active. For the remaining measurements, full pitch listing was retrieved. For minimum F0 and maximum F0, the bottom 3% and top 3% of the F0 listings respectively were excluded to factor out measurement errors such as octave jumps. The range was calculated by subtracting the minimum from the maximum. During the analysis, the recording of one speaker had to be discarded as the autocorrelation method yielded highly irregular results, indicating constant radical changes in F0, not confirmed by auditory inspection. Consequently, recordings of 24 speakers were analysed.

2.2. Measuring gender identity: the Bem Sex-Role Inventory

Following the recorded conversation, each participant filled out a questionnaire designed to measure their gender identity.

For measuring gender identity, the shortened version [7] of BSRI [1] was selected. Despite it being developed already in the 1970's [1], and despite some criticism of its assumptions, it is still, together with the Personality Attributes Questionnaire one of the two most widely used measures of (cultural) gender. One of its merits is its reliability. Another is expedience, as it has been shown [7] that the shortened version is at least as reliable as the original one. The original version comprised 60 traits, 20 masculine, 20 feminine and 20 neutral ones. The shortened version consists of 6 feminine traits and 6 masculine traits. Each trait is rated by the participant on a 7-point Likert scale with regard to how often this trait applies to them. The English terms were translated into Polish by the author.

For each participant, a mean score on the feminine and masculine traits was calculated. Additionally, participants were divided into four categories based on their BSRI score. The four categories – feminine, masculine, androgynous and undifferentiated – are established with reference to median scores of the whole group. If a given participant's score on femininity traits is above the median, and their masculinity score is below the median, they are classified as feminine. If the opposite is the case, they are classified as masculine. If they score above the median for both sets of traits, they are classified as androgynous, and if they score below the median for both sets of traits they are classified as undifferentiated.

3. RESULTS

First, it was tested whether there are correlations between the femininity (F) and masculinity (M) scores and the F0 measures. All statistical analyses were conducted in R [12]. Table 1 below shows correlation coefficients produced by Pearson's Product-moment correlation tests. None of those values has reached statistical significance.

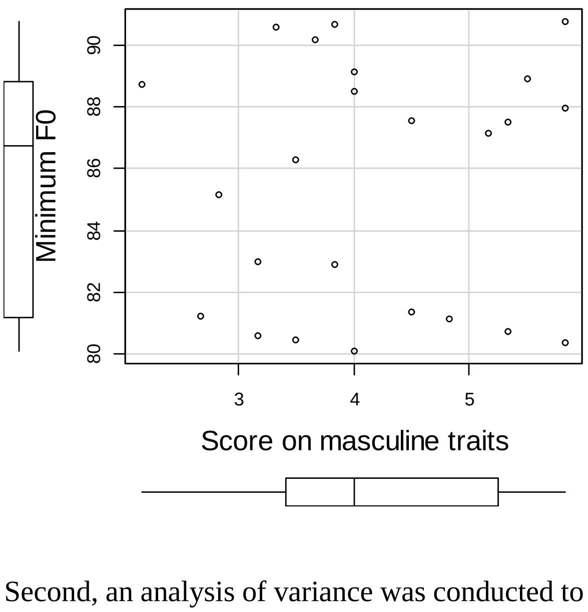
Table 1: Correlation coefficients for femininity (F) and masculinity (M) scores and six F0 measures: mean, median, minimum, maximum, range and standard deviation.

	mean	median	min	max	range	st.dev.
F	-0,11	-0.03	-0,16	0,02	0,09	0,11
M	-0,01	-0,25	0,07	0,09	-0,01	-0,05

Person's product-moment correlation, df = 22

A typical scatter plot of the data fed into the correlation tests is presented in Figure 1 below.

Figure 1: Scatter plot of the relationship between the score on masculine traits in the BSRI (mean for six answers on a seven-point scale) and minimum F0 (in semitones re 1Hz, with the bottom 3% measurements excluded).



Second, an analysis of variance was conducted to see whether there are differences with respect to the six F0 measures depending on which of the four BSRI categories the participants were found to belong. The results of the four-way classification are shown in Table 2.

Table 2: The classification of participants based on their BSRI scores.

	Feminine	Masculine	Androgynous	Undifferentiated
No.	8	6	4	6

One-way ANOVA tests for the six measures have not reached significance in any case. The p-values were as shown in Table 3 below.

Table 3: Results of one-way ANOVA tests with four categories (feminine, masculine, androgynous, undifferentiated) for the six F0 measures.

Measure	*p*-value
Average F0	0,96
Median F0	0,66
Minimum F0	0,92
Maximum F0	0,95
F0 range	0,95
F0 standard deviation	0,97

Neither analysis has corroborated the existence of a relationship between gender identity and F0 measures.

4. DISCUSSION

The analysis of the data gathered for the present study has not shown a relationship between gender identity, as measured by the BSRI, and any of the six F0 measures used. Drawing any strong conclusions about a lack of such a relationship in Polish would be premature, though. Given the extensive number of studies supporting the hypothesis that F0 is influenced by learning, and that sexual and gender identity plays an important role in acquiring voice characteristics, further research is needed.

One potential reason for the lack of a positive result of the present study pertaining to its design is the use of translated terms of the BSRI. First, the translations are not full equivalents, as no translation can ever be; second, for a really reliable culture-specific measure, one would have to develop a questionnaire based on a comprehensive study of gender roles in Poland, much along the lines of the comprehensive research that went into developing the BSRI [1].

Further, a shortcoming common to all purely acoustic studies of variation in F0 is that, despite the use of between-speaker normalization in the form of the use of semitones of ERB's, physiological differences cannot be ruled out as a factor contributing to inter-speaker F0 variation.

5. FURTHER RESEARCH

The development of a gender identity measurement tool for Polish would provide, among other benefits of having such an instrument, a means of a more reliable verification of the relationship between gender identity and the voice of Polish speakers.

Further, to overcome the second shortcoming of the study, a shortcoming inherent in all solely acoustic investigations of inter-speaker variation in F0, a study supplementing acoustics with physiological and articulatory data should be conducted. The measurement of vocal fold length of speakers, together with electromyographic measurements of activation levels in muscles responsible for controlling the tension of the vocal folds would provide the data needed to answer the question of the relative role of physiological versus learned factors influencing F0 variation. It would allow pinning down the physiological, fixed factors underpinning F0, as well those that are under the control of speakers.

6. REFERENCES

[1] Bem, S. L. 1974. The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology* 42(2), 155–162.

[2] van Bezooijen, R. 1995. Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech* 38(3), 253–265.

[3] van Borsel, J., Vandaele, J., Corthals, P. 2013. Pitch and pitch variation in lesbian women. *Journal of Voice* 27(5), 656.e13–656.e16.

[4] Biemans, M. 1998. The effect of biological gender (sex) and social gender (gender identity) on three pitch measures. *Linguistics in the Netherlands* 15, 41–52.

[5] Boersma, P. 1993. Accurate short-tem analysis of the fundamental frequency and the armonics-to-noise ratio of a sampled sound. *Proceedings of the Institute of Phonetic Sciences* Amsterdam 17, 97–110.

[6] Boersma, P., Weenink, D. 2013. Praat: doing phonetics by computer [Computer program]. Version 5.3.59, retrieved 11 December 2014 from http://www.praat.org/

[7] Fernandez, J., Coelleo M. T. 2010. Do the BSRI and PAQ really measure masculinity and femininity? *The Spanish Journal of Psychology* 13(2), 1000–1009.

[8] Labov, W. 1984. Field methods of the project on Linguistic Change and Variation. In: Baugh, J., Sherzer, J. (eds), *Language in use*. Englewood Cliffs: Prentice Hall, 28–53.

[9] Mennen, I., Schaeffler, F., Docherty, G. 2007. Pitching it differently: A comparison of the pitch ranges of German and English speakers. *Proc. ICPhS Saarbrücken*.

[10] Mennen, I., Schaeffler, F., Docherty, G. 2012. Cross-language differences in fundamental frequency range:

A comparison of English and German. *Journal of the Acoustical Society of America* 131(3), 2249–2260.
[11] Pierrehumbert, J. B., Bent, T., Munson, B., Bradlow, A. R., Bailey, J. M. 2004. The influence of sexual orientation on vowel production (L). *Journal of the Acoustical Society of America* 116(4), 1905–1908.
[12] R Core Team. 2012. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, http://www.R-project.org/
[13] Tagliamonte, S. A. 2006. *Analysing sociolinguistic variation.* Cambridge: Cambridge University Press.
[14] Titze, I. R. 1989. Physiologic and acoustic differences between male and female voices. *Journal of the Acoustical Society of America* 85(4), 1699–1707.
[15] Ullakonoja, R. 2007. Comparison of pitch range in Finnish (L1) and Russian (L2). *Proc. 16th ICPhS* Saarbrücken, 1701–1704.
[16] Vorperian, K. H., Kent, R. D. 2007. Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. *Journal of Speech, Language and Hearing Research* 50(6), 1510–1545.