

Average Speaking Pitch vs. Average Speaker Fundamental Frequency – Reliability, Homogeneity, And Self Report Of Listener Groups

Sven Grawunder¹, Ines Bose²



¹Max Planck Institute for Evolutionary Anthropology, Leipzig,
Germany

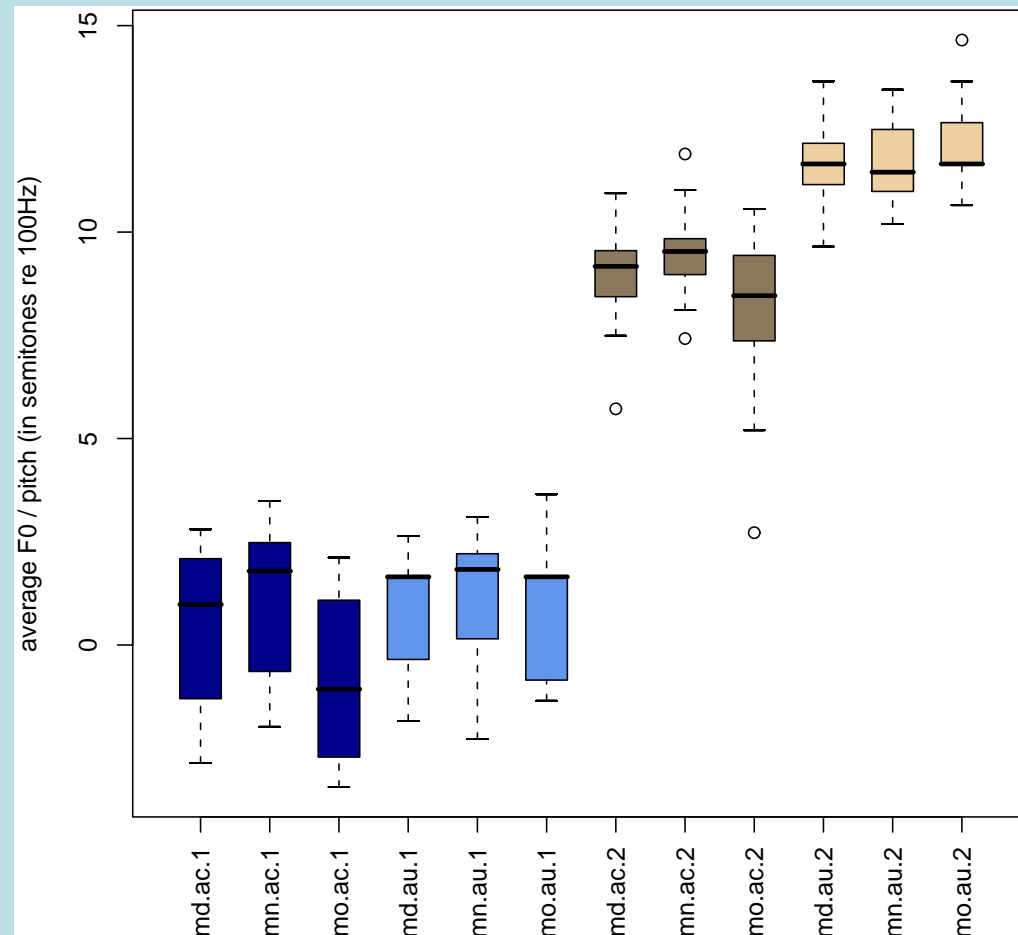
²Martin Luther University, Halle, Germany

introduction

- ASP is considered as an indicator of subjective (tonal) perception, as well as an overall impression of a speaking voice
- ASP as established measure [1,3] needs to be evaluated regarding objective parameters
- preliminary work has shown average pitch (ASP) and speaking fundamental frequency (SFF) to be different [5] and not identical as it has been assumed before [3, 15, 13]
- the specific process of perception, i.e. how the listener decides what to focus on over time in the voice of a speaker, is still opaque [2]
- model: listeners actively follow the pitch movement during the utterance (silently or half-loud); only then do they decide on the virtual center of the perceived tones [7]

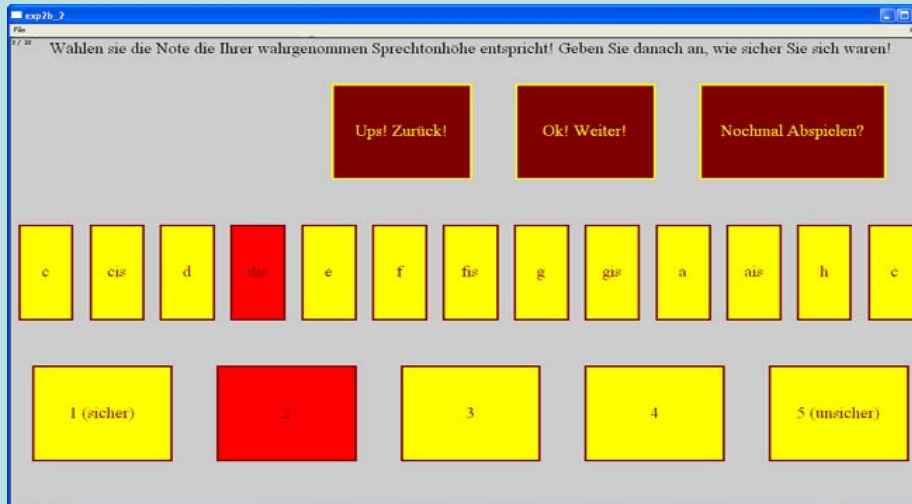
identification experiment

- Average Pitch vs. Speaking Fundamental Frequency
- female speakers are perceived 2.5semitones higher than their 'average' F0-values
- 19 expert listeners provided pitch estimation as musical note



Acoustic values of SFF vs. auditory values of ASP values by 19 subjects on 15 female and 15 male voices; 1=male,2=female, ac=acoustic, au=auditory, mn=mean, md=median, mo=modal.

self reports of listeners

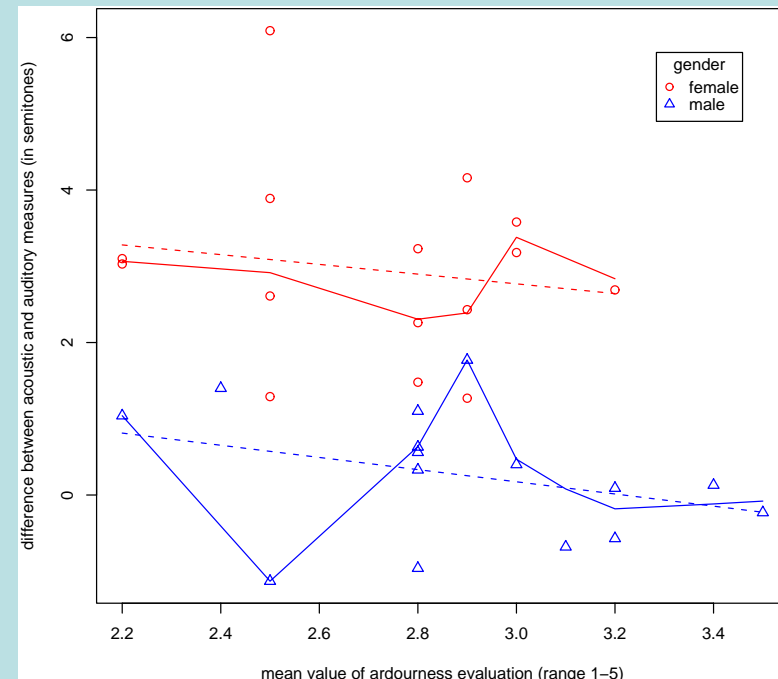


pitch estimation certainty appears as to be not related to the deviation of ASP from SFF

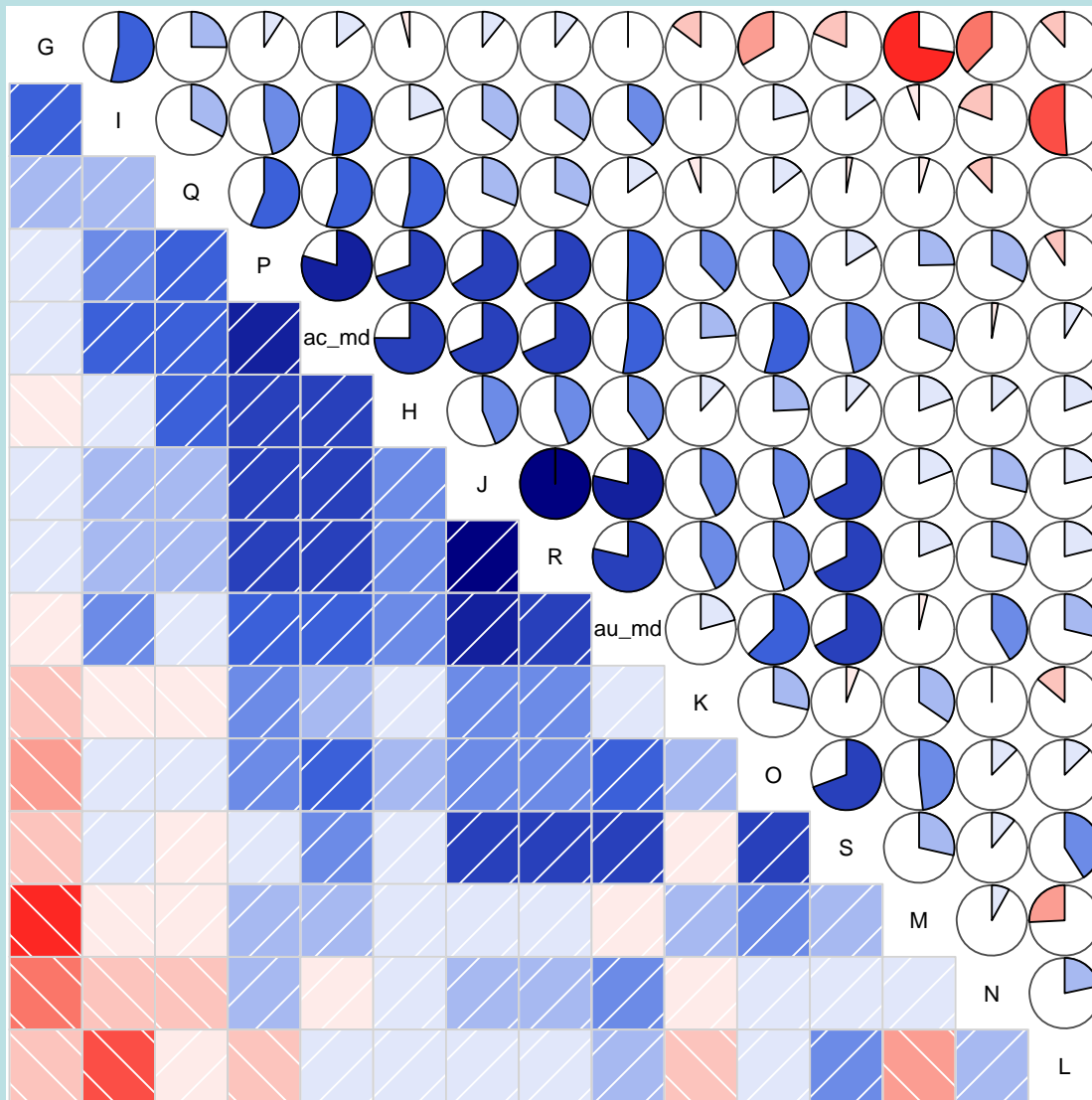
scatterplot of self report ratings about certainty vs. difference between median values of ASP and SFF for 13 raters / 30 samples. Rating 1(= "very easy") to 5 (= "very difficult").

mask in PRAAT (4.5.01)

the listeners were provided with a stimuli (ca. 25sec German radio news) and were asked to estimate ASP on the basis of the musical scale and to give a self report their certainty



listener consistency

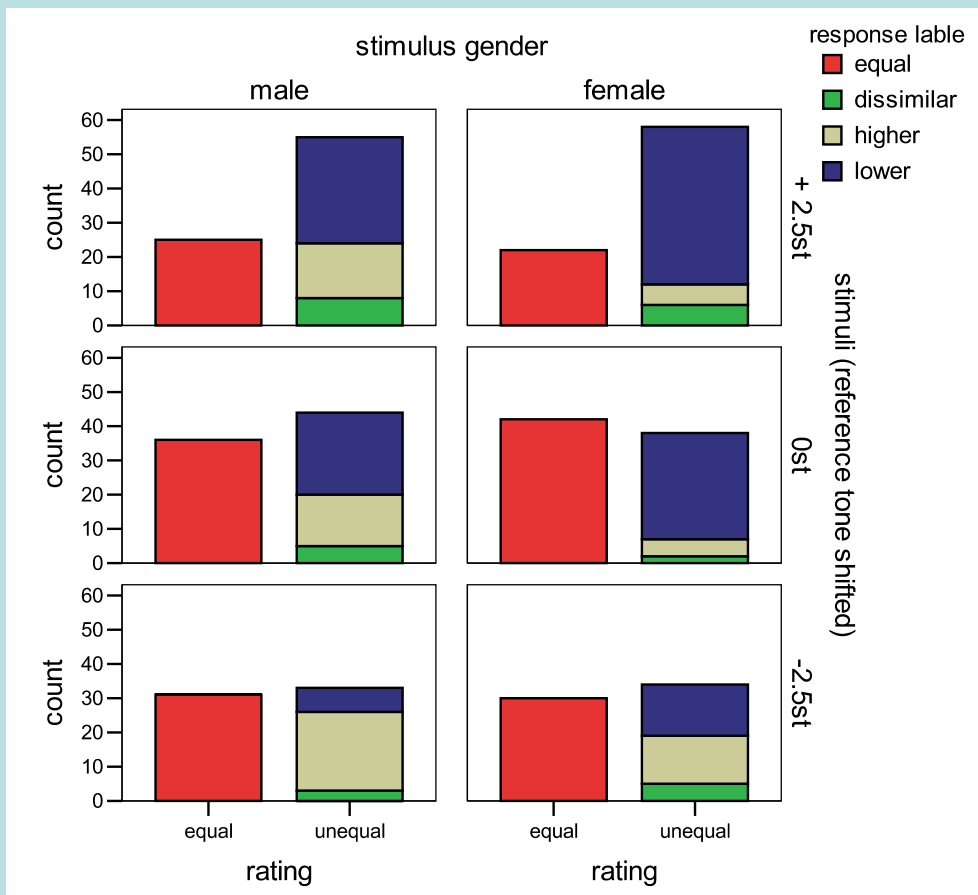


Correlogram [4] of acoustic measures and auditory average measures (median values) as well as individual auditory measures of 13 listeners for 15 female voices. The degrees of correlation (negative in red, positive in blue) are represented either as degrees of shading (lower panel) or completeness of the pie (upper panel). The letters (G to S) represent the estimations per listener (ac md=acoustic median; au md=auditory median).

the majority of pitch estimations clusters positively either around the acoustic median or the (auditory) median

discrimination experiment

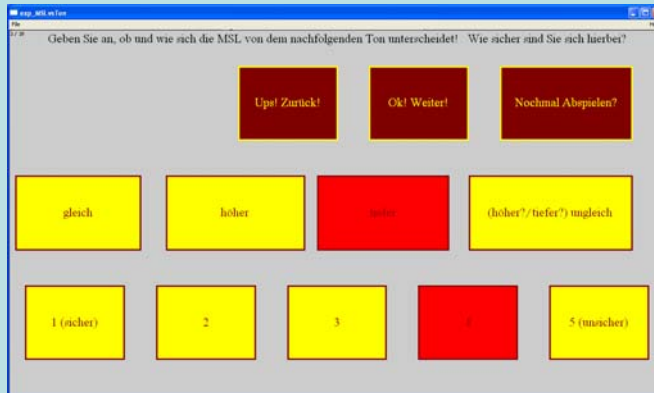
28 samples (14 male / 14 female speakers) were presented as stimuli. The samples were opposed to reference tones which had been partially shifted (-2.5st ; $\pm 0\text{st}$; $+2.5\text{st}$).



Ratings of 16 subjects' discrimination regarding perceived the ASP and a given reference tone (RT); 80 cases of 2.5st upward shifted RT (top), 80 cases of 0st shifted RT (mid), and 64 cases of 2.5st downward shifted RT (bottom) with median F0 as a basis

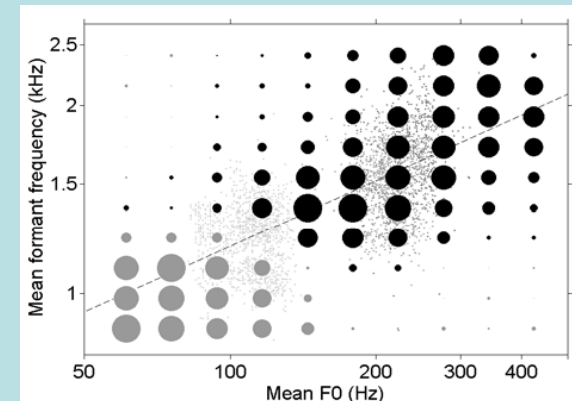
discussion

listener task is extremely difficult



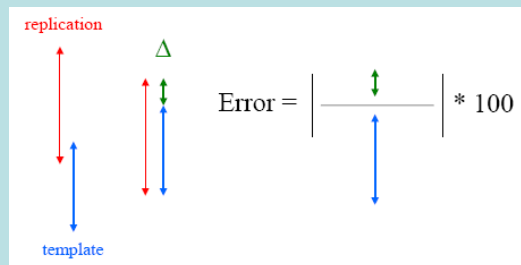
gender perception [16]

effected
by
F0 & FF



intonation perception [8] regarding speakers' intuitions about equivalence of intonational span

across
speakers



•JND around 2.5st

•re-conceptualizing ASP as range instead of tone

•F0-range and micro-variation as factor of influence on ASP

•gender effects in formant structure and ASP seem to be reversible

acknowledgements

This project has been mainly supported by the MPI EVA Leipzig. We are grateful to Benard Comrie (EVA MPI L), our subjects and Roger Mundry (EVA MPI L)

references

- [1] Baken, R. J., Orlikoff, R. F., 2000. Clinical measurement of speech and voice. San Diego: Singular.
- [2] Carlson, R., Elenius, K., & Swerts, M., 2004. Perceptual Judgments of Pitch Range. Speech Prosody 2004, International Conference ISCA. Nara, Japan.
- [3] Braun, A., 1994. Sprechstimmlage und Muttersprache. Göschel (ed.): Z. f. Dial. & Ling., 1, LXI: 170–178.
- [4] Friendly, M., 2002. Corrgrams: Exploratory Displays for Correlation Matrices. The Am. Statistic., 56(4): 316–325.
- [5] Grawunder, S., Bose, I., Hertha, B., Trauselt, F., Anders, L. C., 2006. Perceptive and acoustic measurement of average speaking pitch of female and male speakers in German radio news. Proc. ICSLP, Pittsburgh: 885–888.
- [6] Boersma, P., Weenink, D., 2007. Praat: doing phonetics by computer (Version 4.5.16) [Computer program].
- [7] Kurka, E., Fredrich, R.-B., 1968. Zur Bestimmung des physiologischen Hauptsprechtons. Wiss. Z. Univ. Halle, XVII'68 G, Vol. 5, 45–52.
- [8] Nolan, F., 2003. Intonational equivalence: an experimental evaluation of pitch scales. Proc. 15th ICPhS, 771–774.
- [9] Schultz-Coulon, HJ., 1975. Bestimmung u. Beurteilung d. individ. mittl. Sprechstimmlage. Fol. Phon., 27: 375–386.
- [10] Terhardt, E., 1998. Akustische Kommunikation. Berlin [u.a.]: Springer.
- [11] Traunmüller, H., Eriksson, A., 1995a. The frequency range of the voice fundamental in the speech of male and female adults. 12-Dec-05 retrieved from <<http://www.ling.su.se/staff/hartmut/aktupub.htm>>
- [12] Traunmüller, H., Eriksson, A., 1995b. The perceptual evaluation of F0 excursions in speech as evidenced in liveliness estimations. J Acoust Soc Am, 97(3): 1905–1915.
- [13] Vaissière, J., 2005. The Perception of Intonation. In D. Pisoni and R. Remez, editors, The handbook of speech perception, 236–263. Oxford: Blackwell Publishing.
- [14] Wirtz, M.A., Caspar, F., 2002. Beurteilerübereinstimmung und Beurteilerreliabilität. Göttingen [u.a.]: Hogrefe.
- [15] Wittlinger, I., & Sendlmeier, W.F., 2005. Stimme und Sprechweise erfolgreicher Frauen. In W. Sendlmeier, editor, Sprechwirkung - Sprechstile in Funk und Fernsehen, 71–119. Berlin: Logos.
- [16] Assmann, P. F., Dembling, S., & Nearey, T. M. (2006). Effects of Frequency Shifts on Perceived Naturalness and Gender Information in Speech.