Stimuli from:

Skuk, V. G., Schweinberger, S.R. (2014) . Influences of Fundamental Frequency, Formant Frequencies, Aperiodicity and Spectrum Level on the Perception of Voice Gender. *Journal of speech, Language, and Hearing Research, 57(1),* 285-296, http://dx.doi.org/10.1044/1092-4388(2013/12-0314)

Content: 880 audio-files in wav-format

Parameter specific morphs across vocal gender for two male-female speaker pairs of young (mean age = 22.8 years, *SD* = 3) adults uttering /aba/, /aga/, /ibi/, and /igi/.

Initial voice recordings were time-standardized to a mean plosive release at 292 ms (*SD* = 2) and a duration of 620ms (*SD* = 0) by carefully adding and deleting friction, silence, single vowel period, or glottal closure phases.

Speakers mUA and fDK formed speaker pair P1, and speakers mJN and fEM formed speaker pair P2. To prepare voices for morphing, we set corresponding time anchors at key features (i.e., onset and offset; initial burst of consonant; beginning, middle, and end of formant transitions; stable phase of the vowel) for each of eight (4 Syllables × 2 Speaker Pairs) male–female voice pairs. After morphing, we equated the mean intensity of all stimuli of Experiment 1 and 2 to 70 dB SPL using root-mean-square normalization.

The auditory voice morphing software TANDEM STRAIGHT (Kawahara et al., 2008) analyses and decomposes speech samples into source and filter information that can be modified based on five independent parameters, before resynthesis is done. Both the fundamental frequency track F0(t) and the aperiodicity spectrogram AP(f,t) relate to the voice-source. The time-varying filter configuration is represented by a smoothed spectrogram Ps(t,f) that does not have interference due to periodicity. Each of five parameters (i.e., F0, AP, Ps, t, and f) can be varied either individually or in combination.

Experiment 1 was designed to measure and compare the relative impact of four individual parameters. Table 1 shows the resulting morph types and highlights which of the TANDEM STRAIGHT parameters were morphed (in steps of 10 %) along the male-female morph continua. The residual parameters were set constant to the empirically determined perceived-as-androgynous morph level (also called point of subjective equality, PSE). The PSEs for each speaker pair and utterance are depicted in Table 2. All individual morph types, with a notable exception of AP, had a measurable impact on voice gender classification (see Table I). In Experiment 2, we then varied all combinations of parameters that contributed to voice gender categorization in Experiment 1. The Full morph condition served as reference and was the same in both experiments.

Morph type	TANDEM STRAIGHT parameter					Exp.	Impact on voice gender classification
	F0	AP	Ps	f	t		
F0	morphed	const	const	const	const	1	Yes
AP	const	morphed	const	const	const	1	No
SL	const	const	morphed	const	const	1	Yes
FF	const	const	const	const	const	1	Yes
Full	morphed	morphed	morphed	morphed	const	1, 2	Yes
F0SL	morphed	const	morphed	const	const	2	Yes
F0FF	morphed	const	Const	morphed	const	2	Yes
FFSL	const	const	morphed	morphed	const	2	Yes
F0FFSL	morphed	const	morphed	morphed	const	2	Yes

Table 1. Overview of morphed TANDEM STRAIGHT parameters for each experimental condition (morph type) of experiments 1 and 2.

Speaker Pair	Utterance			
	/aba/	/aga/	/ibi/	/igi/
P1	42 (58)	36 (64)	35 (65)	38 (62)
P2	33 (67)	34 (66)	33 (67)	27 (73)

Table 2. Morph level at the point of subjective equality (with the proportion of "femaleness" in parentheses), separately for both speaker pairs and each of four utterances.

Resulting stimuli and filename convention

- Experiment 1: 440 stimuli, i.e. 2 male-female speaker pairs x 4 utterances x 5 morph types (Full, F0, FF, SL, AP) x 11 morph level
- Experiment 2: 440 stimuli, i.e. 2 male-female speaker pairs x 4 utterances x 5 morph types (Full, F0FF, F0SL, FFSL, F0FFSL) x 11 morph level

File name conventions (e.g. P1-ABA-AP-SL058-FF058-AP000-F0058-T058.wav) code for the following information in order: Speaker pair (P1 or P2), utterance (ABA, AGA, IBI, or IGI), morph type (F0, AP, SL, FF, Full, F0SL, F0FF, FFSL, or F0FFSL) and the five TANDEM STRAIGHT parameters SL, FF, AP, F0, and T, each with the respective proportion of femaleness (000 to 100; 000 corresponds to 0% of the female voice and 100 corresponds to 100% of the female voice) of the current stimulus.

Note that these stimuli have been provided to you as an academic researcher on the agreement that you:

- will make sure that the stimuli are not passed on to other researchers without previous request to, and agreement from, the authors of the original puplication (present mail address of corresponding author: verena.skuk@uni-jena.de"> verena.skuk@uni-jena.de)
- will cite the following articles in any published report of your research

  - Skuk, V. G., & Schweinberger, S. R. (2014). Influences of Fundamental Frequency, Formant Frequencies, Aperiodicity and Spectrum Level on the Perception of Voice Gender. *Journal of Speech, Language, and Hearing Research, 57(1),* 285-296. doi: http://dx.doi.org/10.1044/1092-4388(2013/12-0314)

  - Kawahara, H., Morise, M., Takahashi, T., Nisimura, R., Irino, T., & Banno, H. (2008, March 31 2008-April 4 2008). Tandem-STRAIGHT: A temporally stable power spectral representation for periodic signals and applications to interference-free spectrum, F0, and aperiodicity estimation. Paper presented at the *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*. http://dx.doi.org/10.1109/ICASSP.2008.4518514

Attachments:

- stimuli-Skuk-2014-JSLHR-Exp1.zip (440 .wav files, 22.882 kB)
- stimuli-Skuk-2014-JSLHR-Exp2.zip (440 .wav files, 22.979 kB)