hdk

Zürcher Hochschule der Künste
Zurich University of the Arts

Universität Zürich[UZH]

"Flat" vowel spectra revisited in vowel synthesis

Dieter Maurer, Heidy Suter

Institute for the Performing Arts and Film, Zurich University of the Arts, Switzerland
dieter.maurer@zhdk.ch, heidy.suter@zhdk.ch **www.phones-and-phonemes.org/asa**

Acoustics '17, Boston
173rd Meeting of the Acoustical Society of America and the 8th Forum Acusticum

Journal of the Acoustical Society of America
141(5): 3469–3469

Introduction

Vowel qualities of natural and of synthesised sounds which exhibit "flat" spectral envelopes or envelope parts in terms of vowel-related frequency ranges with harmonics equal in amplitude can be recognised [1, pp. 147–157; 2]. The present investigation addresses this question in details in a vowel synthesis experiment in which sounds related to various series of harmonics, multiples of 200Hz and equal in amplitude, were created. The frequency range of investigation was 200–4000Hz.

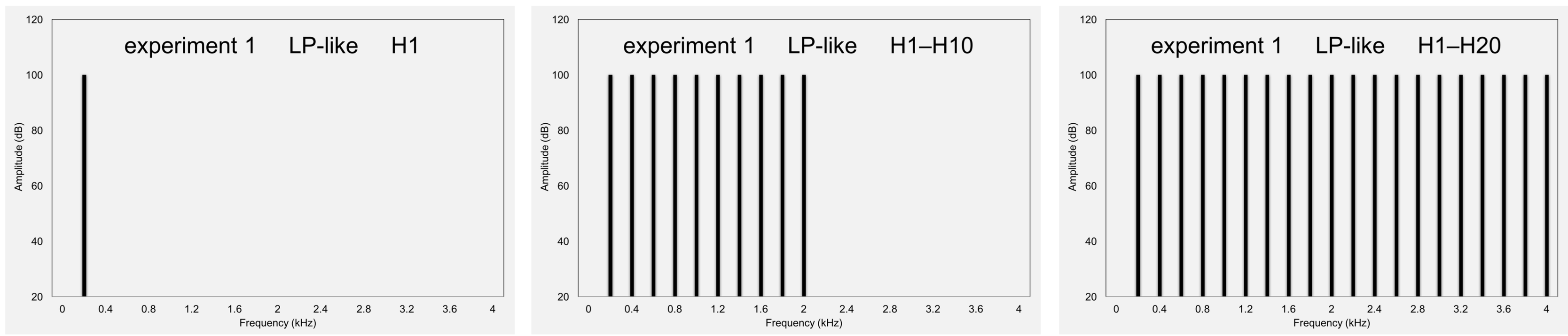
Harmonic synthesiser, experiments 1 to 3

Synthesis: Monotonous sounds of 1.2 sec. (incl. 0.1 sec. fade in/out) were produced with a harmonic synthesiser, the harmonic frequencies set as multiples of a reference frequency H1 = 200Hz, all their amplitudes set as equal, and all phases set as 0.
Three different types of sounds according to three different experiments were created (see Figure 1).
LP filter-like – increasing the number of higher harmonics stepwise from H1 to H1–H20 (experiment 1): A series of sounds was produced beginning with only one harmonic H1 (first sound) and then stepwise increasing the number of harmonics.
HP filter-like – decreasing the number of lower harmonics stepwise from H1–H20 to H20 only (experiment 2): A series of sounds was produced beginning with H1–H20 (first sound) and then stepwise decreasing the lower number of harmonic until H20 only.
BP filter-like – increasing the number of harmonics stepwise from a middle harmonic to a band of higher harmonics (experiment 3). Several series of sounds were produced, beginning with only one harmonic H2 or H3 or H4 and so on up to H15 as the first sound, and then increasing the number of harmonics stepwise up to a band of eleven consecutive higher harmonics in maximum (frequency band = 2000Hz) or of consecutive higher harmonics up to H20.
Note. Harmonics are given always in reference to H1 = 200Hz, even if a sound spectrum shows only one harmonic on a higher frequency.
Listening test. Vowel recognition was investigated by means of a listening test: five phonetic expert listeners assigned the synthesised sounds to Standard German vowel qualities or to "no vowel" (for details, see the link below).

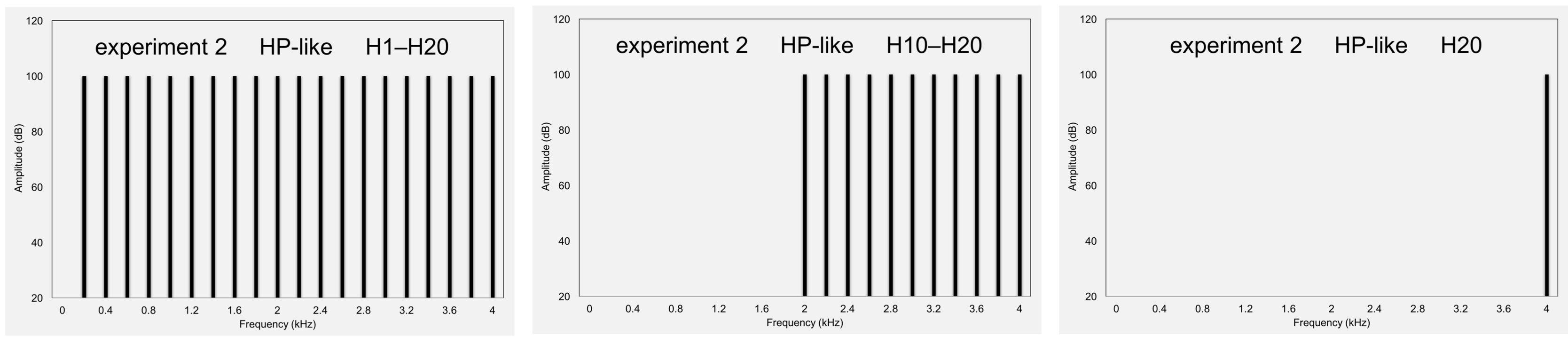
Results

Sounds with "flat" harmonic spectra, for which a majority of listeners perceived the same vowel quality, were found for the following qualities:
Experiment 1: /u, a, ɛ/ (5/5), /ɔ/ (4/5), /o/ (3/5)   (note: 5/5, 4/5 and 3/5 = results of the listening test, i.e., 5 or 4 or 3 of 5 listeners perceived the same vowel quality)
Experiment 2: /i/ (5/5), /a/ (4/5), /ɛ, e/ (3/5)
Experiment 3: /u, o, ɔ, a, ɛ, e, y, i/ (5/5), /ø/ (3/5)
Examples of "best cases for vowel recognition" are shown in Figure 2a to 2c.
For details of the listening test and for the entire table of the corresponding results (confusion matrices), see
https://www.phones-and-phonemes.org/asa/Poster-ASA-2017b-Add-170618.pdf

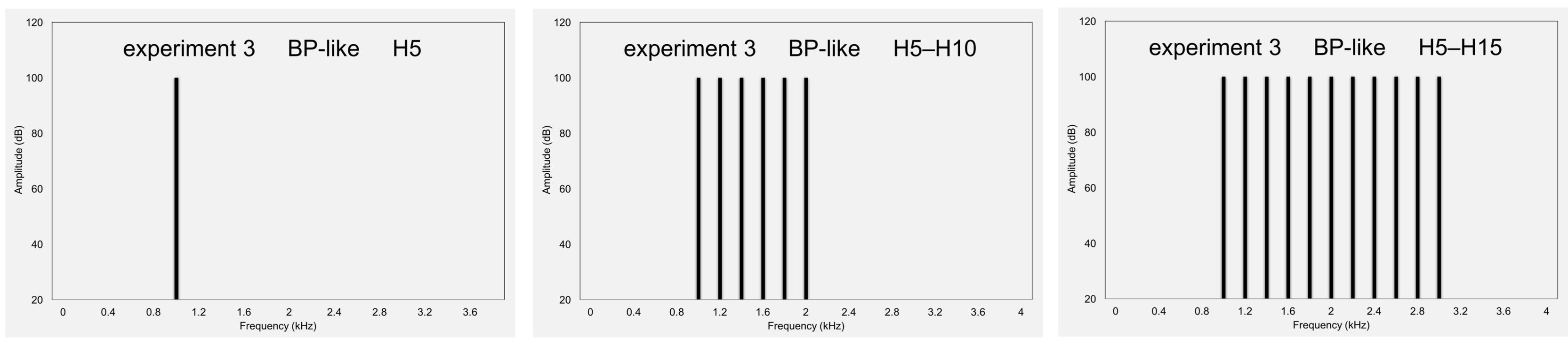
Figure 1, illustration of the experiments 1 to 3



Experiment 1: Increasing the number of harmonics stepwise from H1 to H1–H20.



Experiment 2: Decreasing the number of harmonics stepwise from H1–H20 to H20 only.



Experiment 3: Increasing the number of harmonics stepwise from a middle harmonic to a band of harmonics.

Figure 1: Illustration of the experiments 1 to 3. For each experiment, the harmonic spectra of the first, the middle and the last sound of a series is shown.

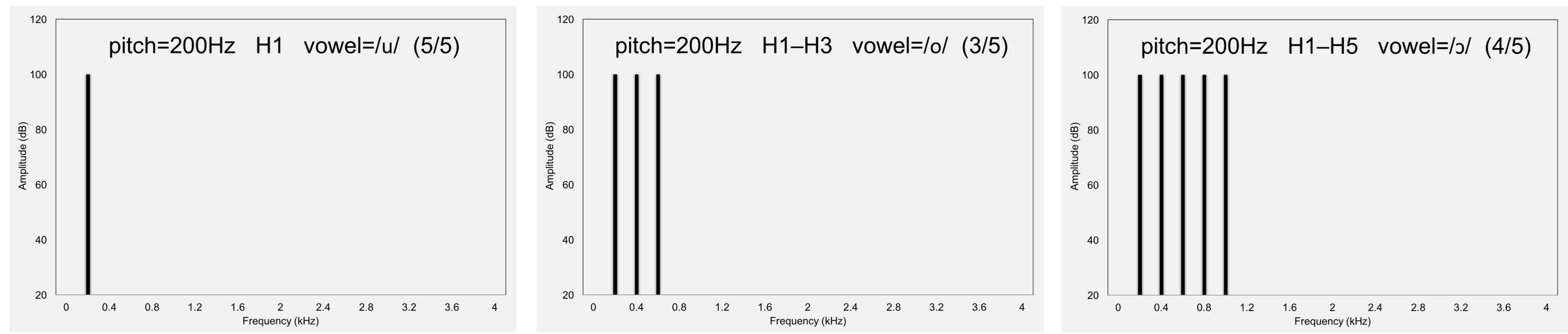
Discussion

(1) For all investigated vowel qualities except /u/ and /ø/, examples of unambiguously recognised vowel sounds related to "flat" harmonic spectra were found.
(2) Some sounds of /u/ related to a single harmonic ≤ 1kH were also unambiguously recognised.
(3) For sounds of front vowels, vowel recognition was maintained in cases of harmonic spectra lacking spectral energy < 1kHz (usually attributed to F1).

These synthesis experiments confirm the above mentioned findings from earlier studies: **vowel recognition does not in principle base on spectral peaks.**
Further research on the matter should address the question of the effect of the reference frequency H1 on vowel recognition: for a given frequency range, decreasing the reference H1 frequency increases the resolution of harmonics and may therefore increase the number of clearly recognisable sounds, and vice versa.

References
[1] Maurer, D. (2016): Acoustics of the Vowel – Preliminaries. Bern/Frankfurt: Peter Lang.
[2] Carpenter, A., and Morton., J. (1962): The perception of vowel colour in formantless complex sounds. Language and Speech, 5(4), 205–214.

Figure 2, recognisable vowel sounds (examples)



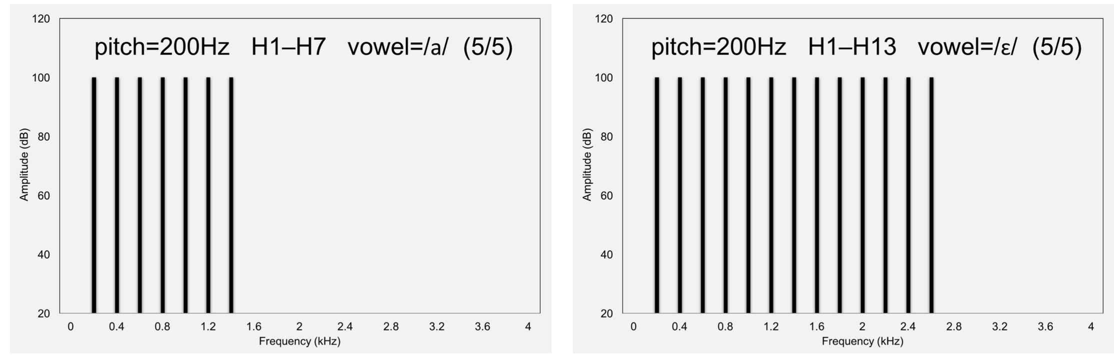
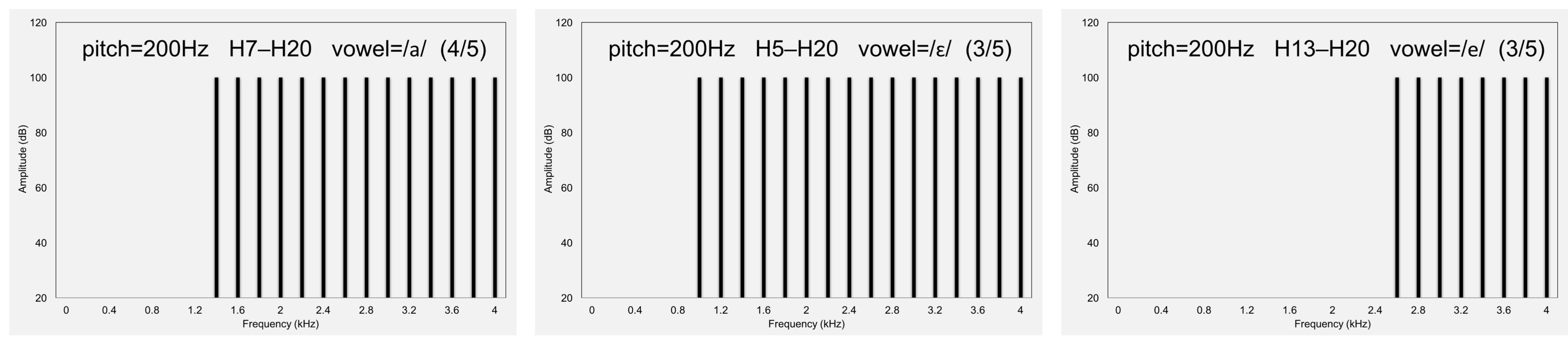


Figure 2a: Examples of harmonic spectra of recognised vowel sounds for experiment 1 (ratio given in a graphic on the right = result of the identification test; e.g. 5/5 = 5 of 5 listeners).



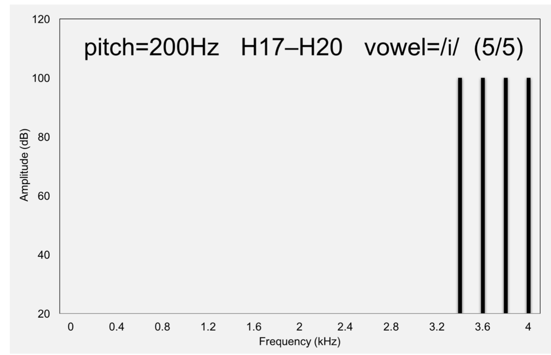
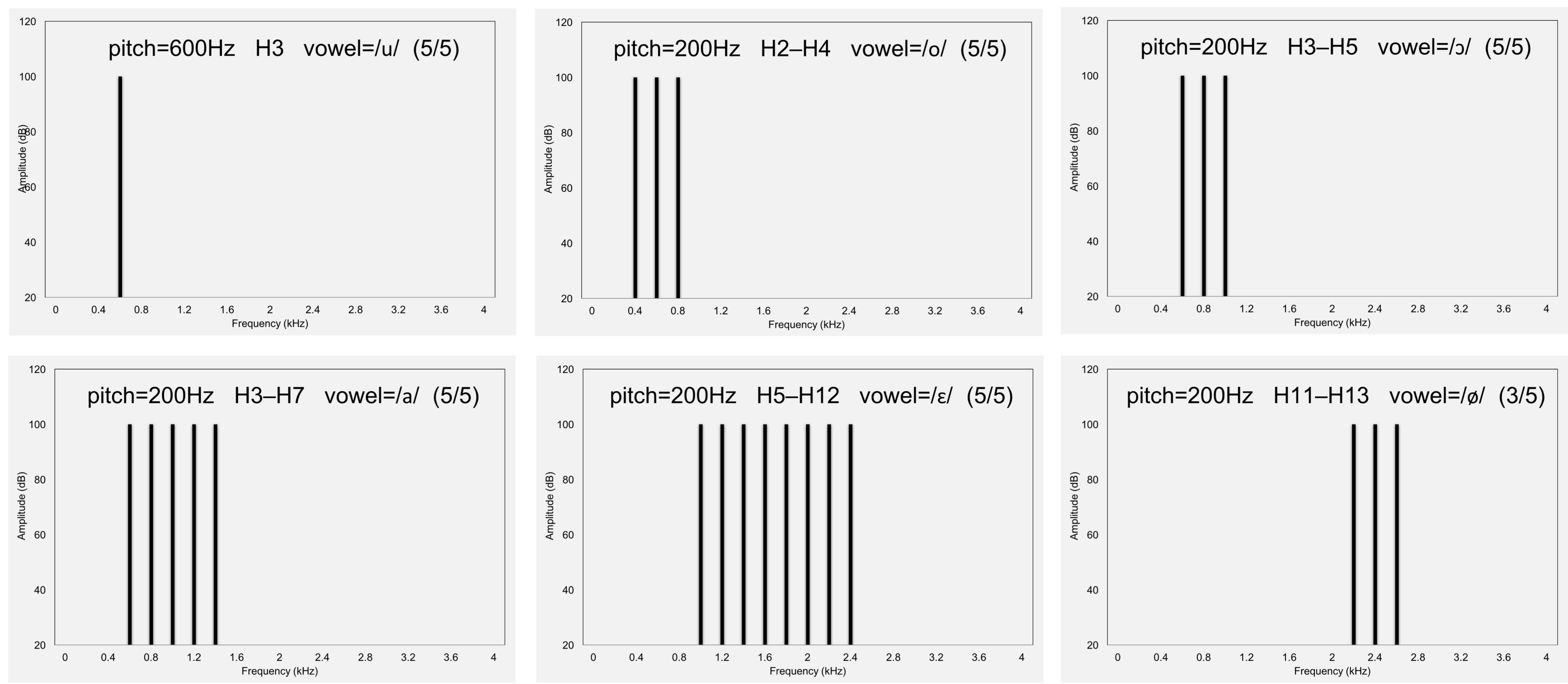


Figure 2b: Examples of harmonic spectra of recognised vowel sounds for experiment 2.



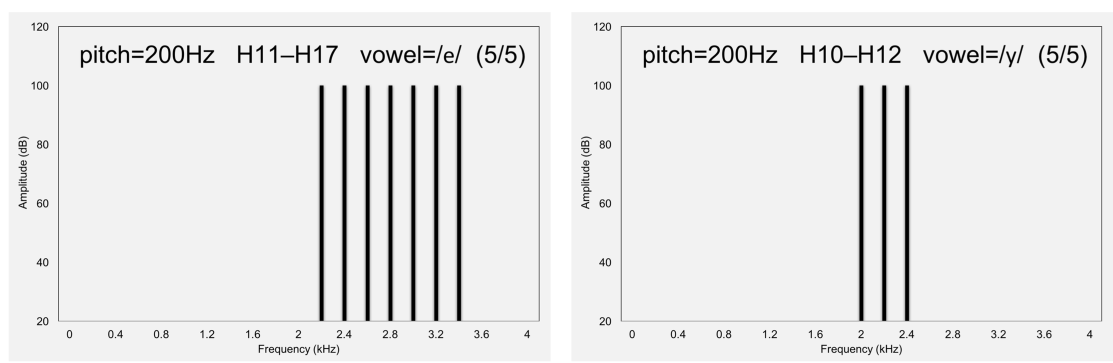


Figure 2c: Examples of harmonic spectra of recognised vowel sounds for experiment 3. For the sound example of /i/, see Figure 2b.