zhdk
Zürcher Hochschule der Künste
Zurich University of the Arts

Universität Zürich[UZH]

Formant pattern and spectral shape ambiguity of vowel sounds, and related phenomena of vowel acoustics – Exemplary evidence

Interspeech 2019

15–19 September 2019
Graz – Austria

Show and Tell Presentation

Dieter Maurer[1], Heidy Suter[1], Christian d'Heureuse[1], Volker Dellwo[2]
[1] Institute for the Performing Arts and Film, Zurich University of the Arts, Switzerland, [2] Department of Computational Linguistics, University of Zurich, Switzerland
dieter.maurer@zhdk.ch, heidy.suter@zhdk.ch, chdh@inventec.ch, volker.dellwo@uzh.ch http://is2019.phones-and-phonemes.org

Topic

Formant pattern and spectral shape ambiguity: Formant frequency patterns and spectral envelopes as such relate to different vowel qualities. First and foremost, this ambiguity phenomenon is a consequence of a non-systematic and complex interaction between pitch and spectral characteristics in perceived vowel quality.

Aim

In the form of an online documentation,

- exemplary evidence is given for the ambiguity phenomenon in terms of direct comparisons of natural vowel sounds linked to a Klatt synthesis tool for re-synthesis with varying fundamental frequency (f_o);
- the acoustic and perceptual context of the ambiguity phenomenon is extensively illustrated (see the Content list).

Sound playback and acoustic analysis (fundamental frequency, sound spectrum and spectrogram, formant patterns) is included.

For the various disciplines related to speech acoustics, the documentation addresses

- researchers to allow them to re-evaluate results of existing studies and to create experimental settings for future experiments, taking into account the actual variation and pitch-dependency of the vowel spectrum
- students to support the acquisition of state-of-the-art knowledge of vowel acoustics

Relevance

In the literature on vowel acoustics, there is an extensive and often controversial debate on whether the primary acoustic cues of vowel quality are contained in the formant patterns or, alternatively, in the spectral shape. Yet, recent studies have shown that neither formant patterns nor spectral shapes are vowel quality-specific but that they are ambiguous because of a complex interaction between pitch and vowel-related spectral characteristics.

Hitherto, the awareness and discussion of the ambiguity phenomenon

Making evident the ambiguity of formant patterns and spectral envelopes, with re-synthesis option

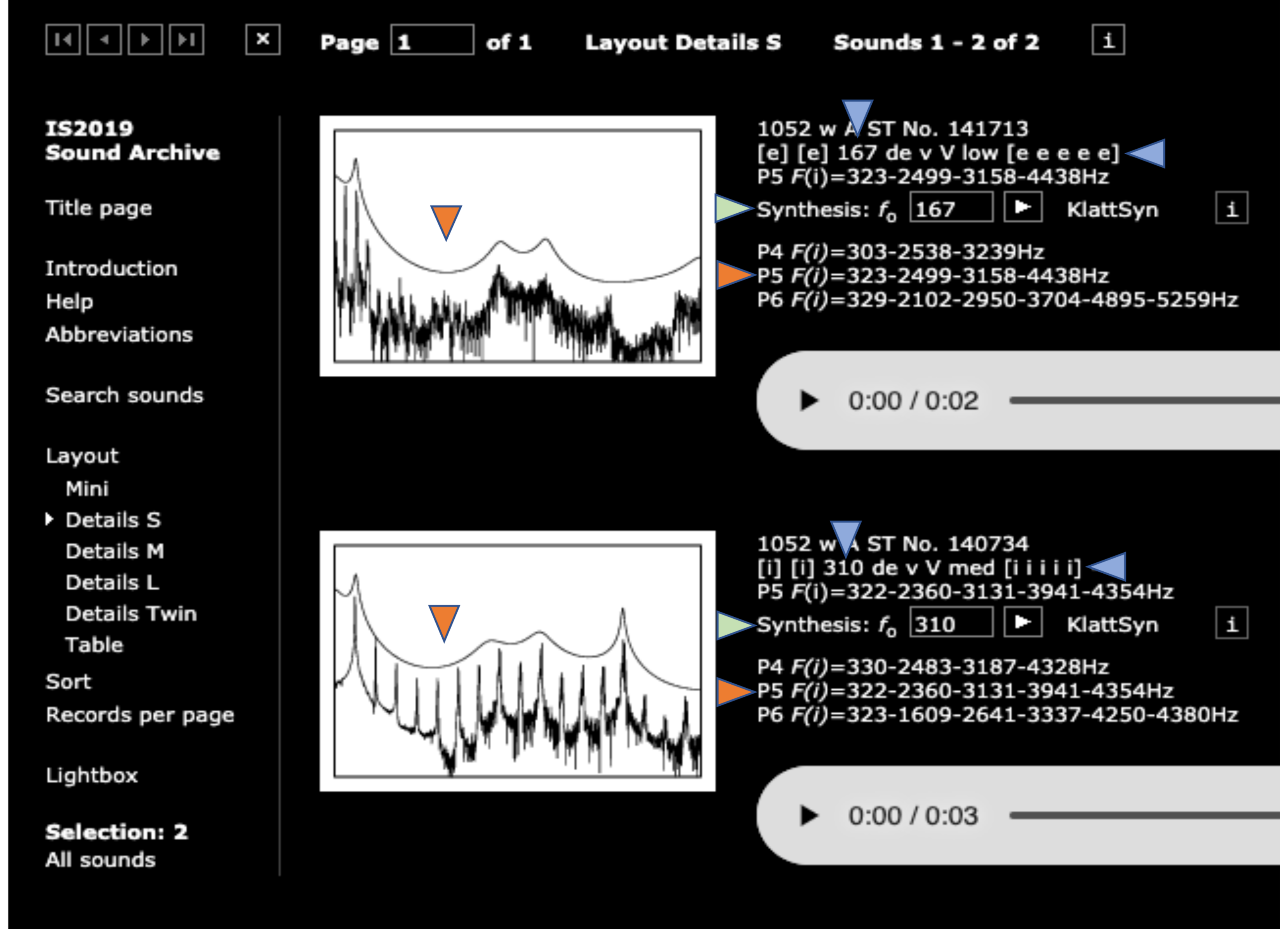

Illustration of an example: Two sounds of a single female speaker

- with similar formant frequencies F_1–F_2–F_3 (see ►)
- and similar spectral envelopes < 3kHz (see ▼)
- but different fundamental frequencies of 167Hz and 310Hz (see ▼)
- and different intended and recognised vowel qualities (see ◄ ; vowel recognition by five professionally trained speakers)

For verification, each natural sound can be re-synthesised with Klatt synthesis on both fundamental frequencies (see ▷).

Relevance (continuation)

was restricted to a limited number of specialists, and there is a lack of direct access to compilations of sound examples which allow for an understanding of the phenomenon. The present online documentation gives insight into the phenomenon for a wider audience of researchers of various fields dealing with speech acoustics: It allows to acquire basic knowledge, to extensively listen to exemplary sound series and to directly compare *F*-patterns and spectral shapes, and to crosscheck the perceptual role of f_o and *F*-patterns in vowel synthesis.

Content of the online documentation

Part I – Formant pattern and spectral shape ambiguity

1. **Natural sounds and their re-synthesis; sounds of different vowels at different f_o with similar formant patterns and/or spectral shapes**:
   Ambiguity for /ɛ–e–i/
   Ambiguity for /ɛ–ø–y/
   Ambiguity for /a–o–u/
2. **Synthesised sounds related to open-tube filter patterns, varying f_o:**
   Ambiguity for /ə–ø–y/

Part II – Context of the ambiguity phenomenon

3. **f_o contour and upper f_o ranges of speech:**
   Everyday speech samples and speech during artistic performance with upper f_o exceeding 350 Hz for men and 500 Hz for women
4. **Vowel recognition of natural isolated high-pitched sounds:**
   Sound of all long vowels at f_o = 700–800 Hz
   Sounds of the corner vowels /i–a–u/ at f_o = 1 kHz
5. **Pitch-dependency of the vowel spectrum in natural vocalises:**
   Sounds of all long vowels of a man, a woman and a child with f_o variation (C-major scale) of 22–34 semitones
6. **Other aspects of spectral variability for natural vowel sounds:**
   Sounds with different modes of phonation (voiced, breathy, whispered and creaky phonation)
   Sounds with different vocal effort
7. **Non-systematic relation between vowel-related spectral peaks or spectral envelopes for natural vowel sounds:**
   Sounds with different numbers of spectral peaks
   Sounds with "flat" or "sloping" spectral portions in their vowel-specific frequency range
   Differences related to vowel quality, range of f_o variation, formant levels, and harmonic configuration

Online documentation: **http://is2019.phones-and-phonemes.org**

Research homepage: **http://www.phones-and-phonemes.org**