A Test of Quantal Theory: the effect of the second subglottal resonance in speech perception

Asaf Bachrach

Steven Lulich

Nick Malyska

Theory

Stable regions separated by unstable regions in articulatory-acoustic space define regions that may act as anchors for phonological categories (e.g. distinctive features).

For the feature [±back], it is hypothesized that the region in which F2 becomes close to the 2nd subglottal resonance (Z2) acts as an unstable boundary region.

Speech Production

Chi and Sonderegger (2004) found evidence from a speech production experiment, and from a typological review of vowel formant structure, that supports the hypothesis that Z2 plays a role in defining a quantal boundary between front and back vowels.

However, production data alone cannot rule out the possibility that other mechanisms are responsible for shaping the vowel space – the production evidence for Z2 as a boundary could be merely coincidental.

Figures removed due to copyright restrictions.

Speech Perception

Quantal theory is normally discussed in the spectral domain. The notion of a landmark, however, is a temporal-domain extension of quantal theory.

In a transition from a front to a back vowel (or from a front vowel to a labial stop), F2 and Z2 will cross, forming a landmark – a discontinuity in time.

We hypothesized that – if Z2 is not a coincidental boundary between front and back vowels – manipulating the location of such a landmark (in time) will effect the perception of vowel frontness and consonant labiality in a gating experiment.

Experimental design (I)

- The role of the sub-glottal zero in the perception of V→C transitions
- Front vowel /æ/ → Labial consonant /p/
- Backward gated identification task
- Partially synthesized targets
- Free-choice design

Experimental design (II)

- Front vowel /æ/– Labial consonant /p/
- Three target forms:

1) /æptər/
2) /æ(p)ktər/ (k spliced)
3) /æ(p)t/ (t spliced)

- Two sub glottal zero conditions:

1400 Hz

1600 Hz

- 8 fillers (ukter, opter, etc.)

Experimental design (III)

- Backward gated identification task:
  - Targets:18 Gates; gate 1, the minimal bi- syllabic percept, Gate 18, the full form
  - Fillers: 6 gates
- Sample of the 18 gates :

Experimental design (IV)

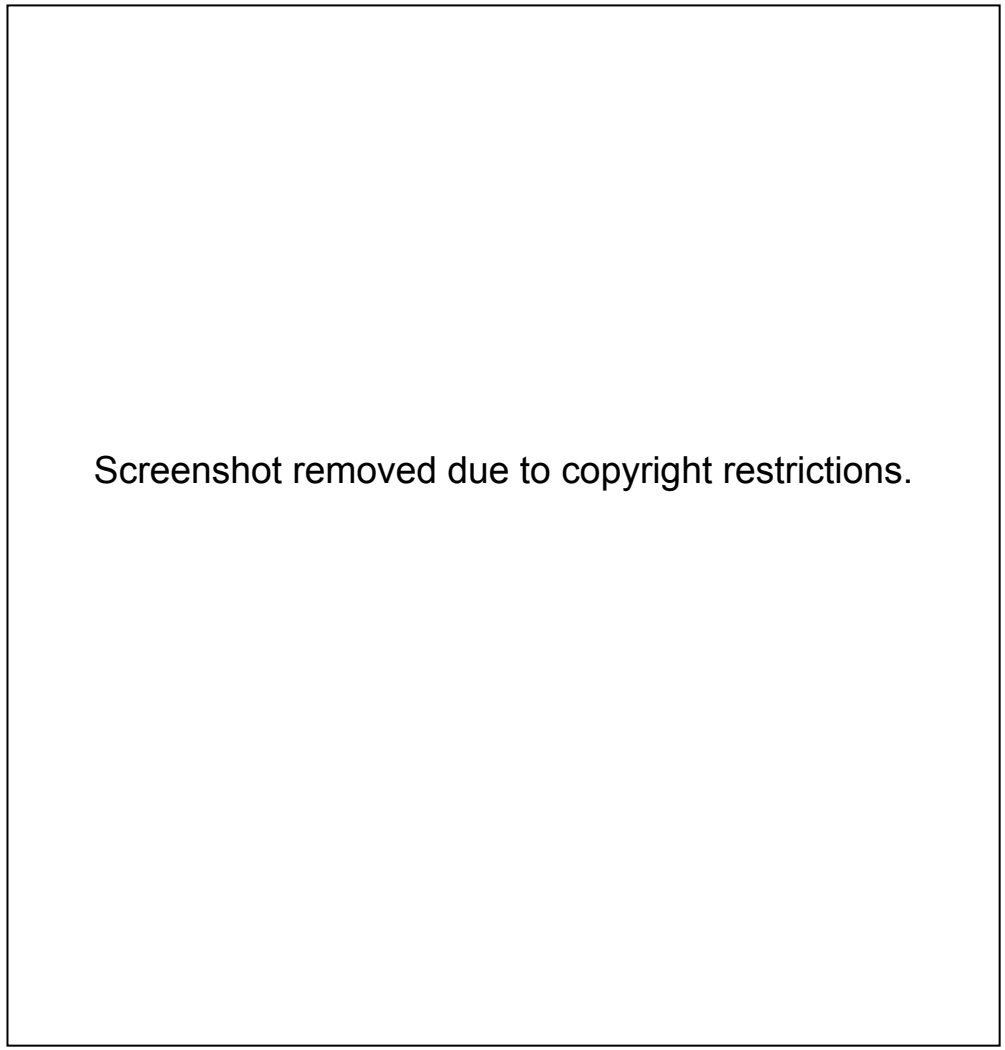
- Backward gated identification task:
- The 1400Hz sub glottal zero appears on an earlier gate than the 1600Hz one.
- Our hypothesis:
  - Subject preference for æ→p transition identification will be observed on shorter gates in the 1400Hz condition compared to the 1600Hz condition.

Experimental design (IV)

- Each gate presented twice for a total of 312 items
- Every item heard twice before identification
- Free-choice design:

Experimental design (V)

- Goodness ratings
- Self paced



Experimental design (VI)

Controlling for possible ‘top-down’ effects:

- Pseudo-randomized presentation:
- Block I: 32 items (1st-4th gate of each word)
- Block II: 280 items randomized
- Two groups:
  - Group A: First item on block II ‘apter’
  - Group B: First item on block II ‘a$_p$kter’
- Fillers contain similar V-C transitions

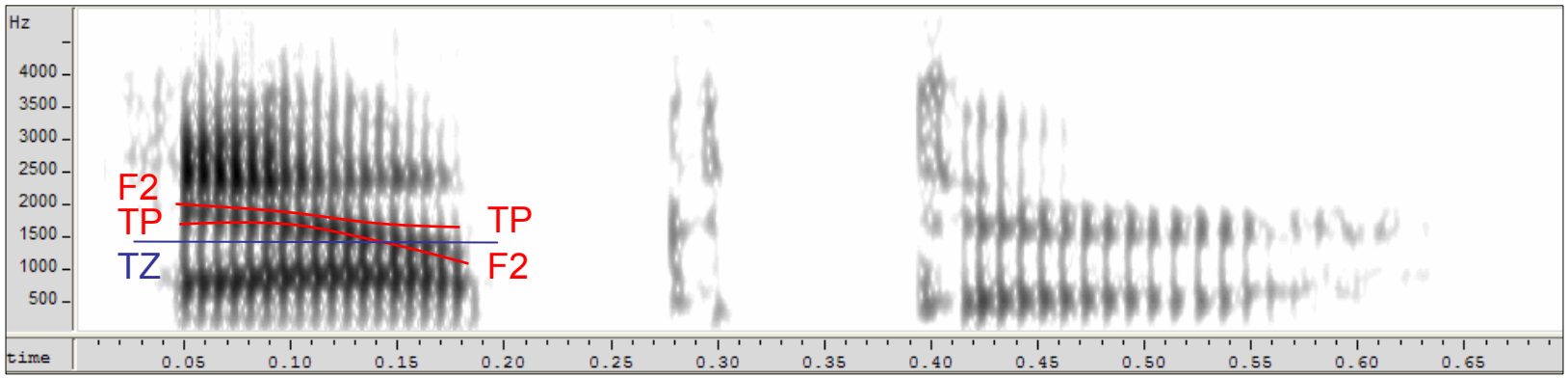
Experimental design (VI)

Data collection:

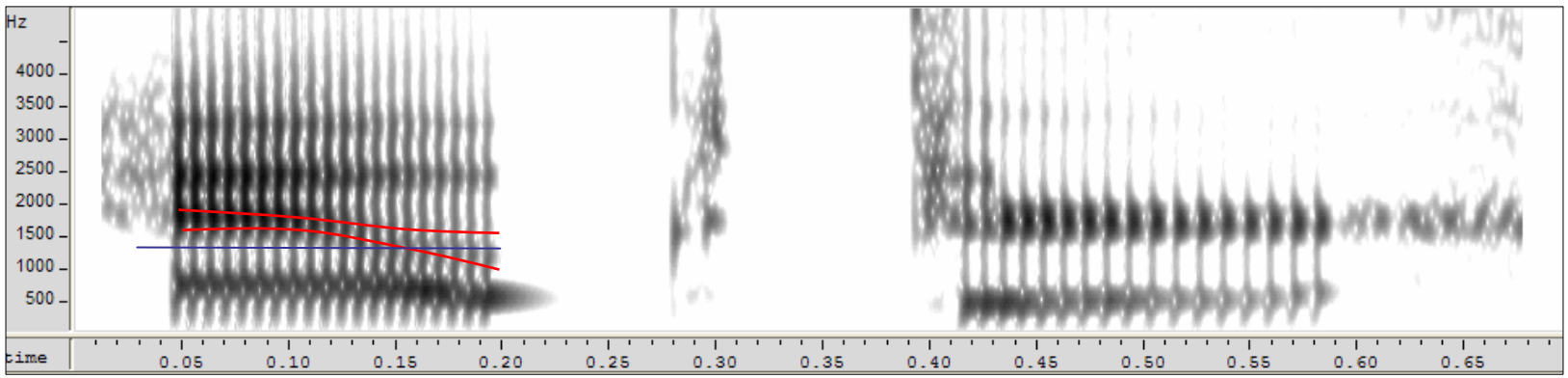
- 8 Subjects (native speakers of English)
- Responses were coded along two dimensions:
  1. +/- front vowel
  2. +/- labial percept (p/b/w)

Synthesis of Apter
Tracheal Resonance

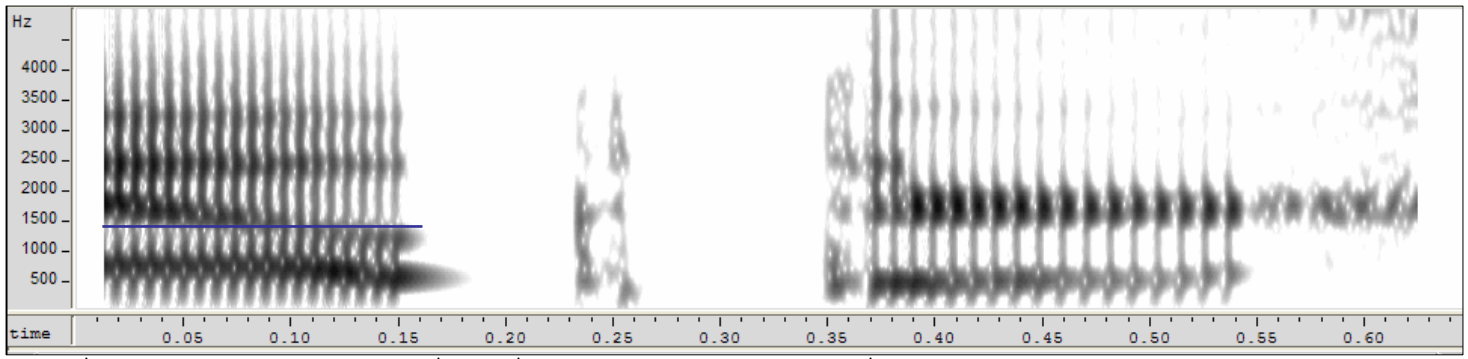
Original



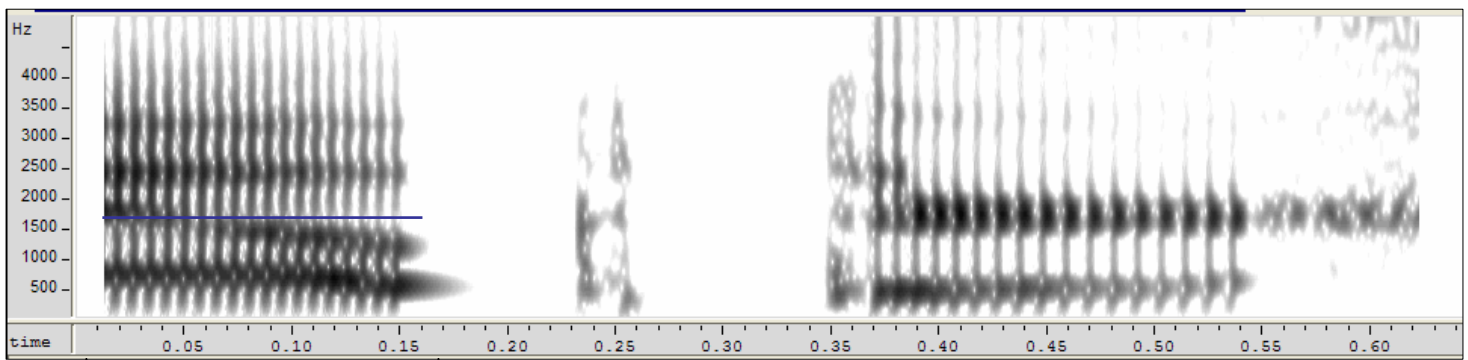
Copy Synthesis



Synthesis of Apter
Experimental Stimuli

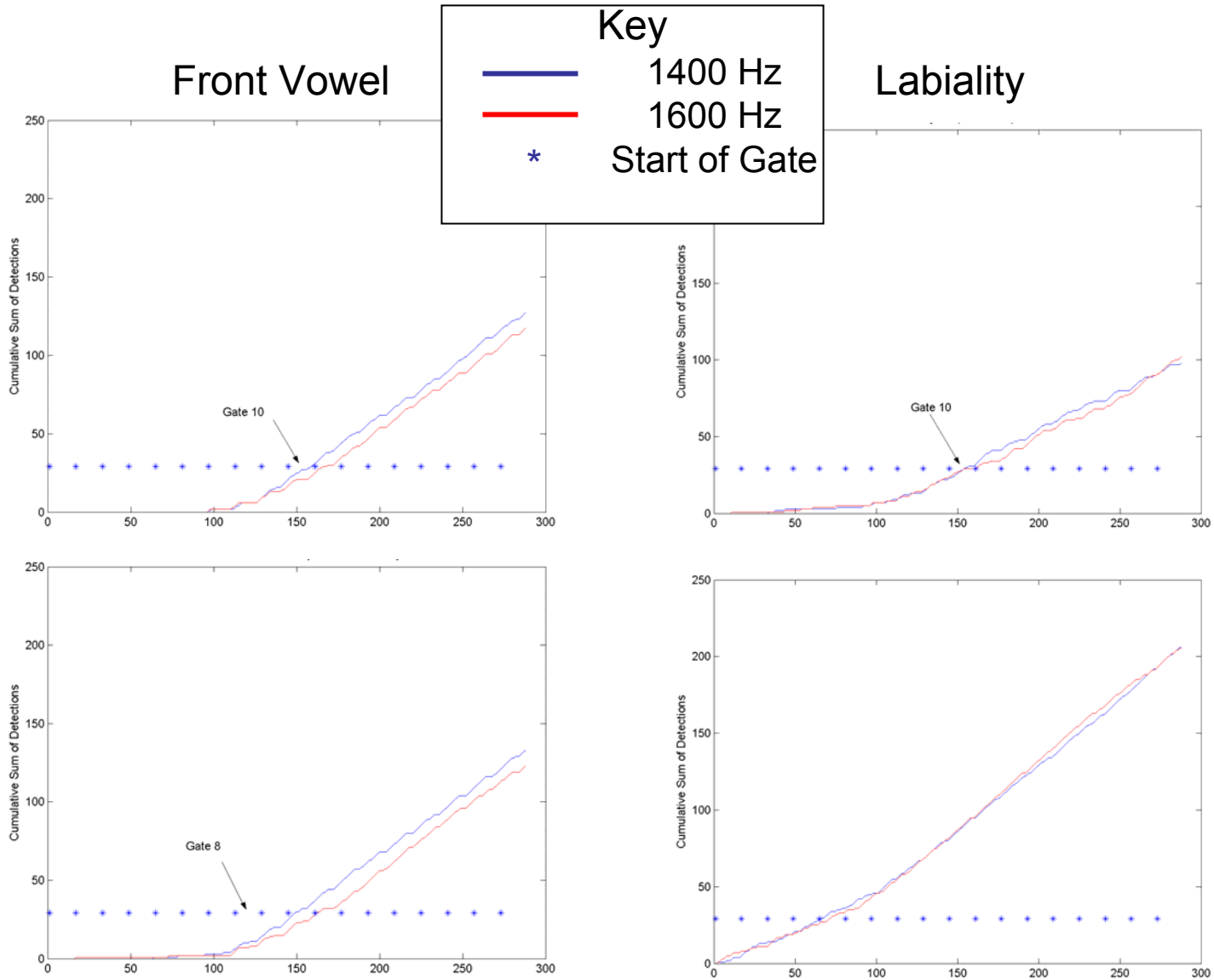


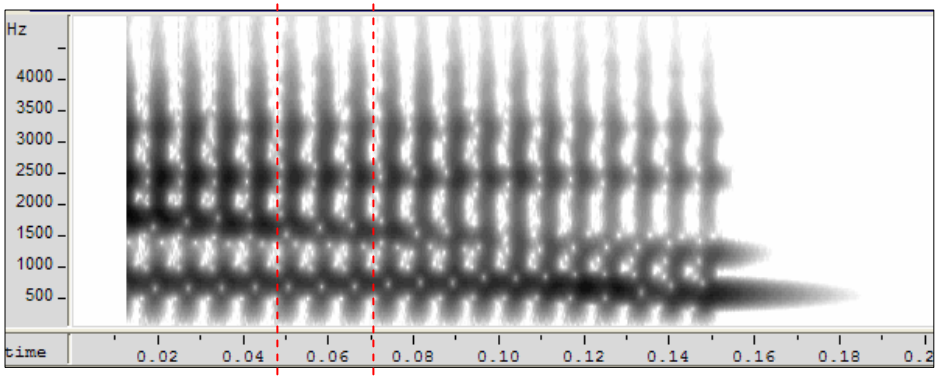
1400 Hz Zero Spliced-In Bursts



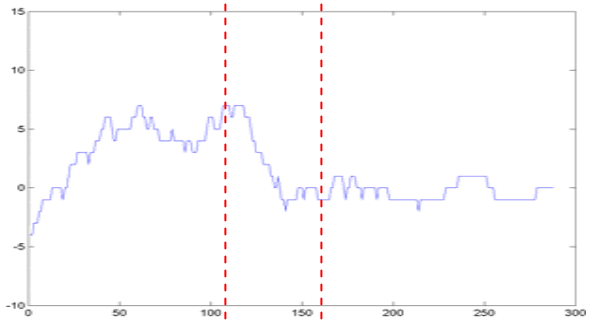
1600 Hz Zero Spliced-In Bursts

Results

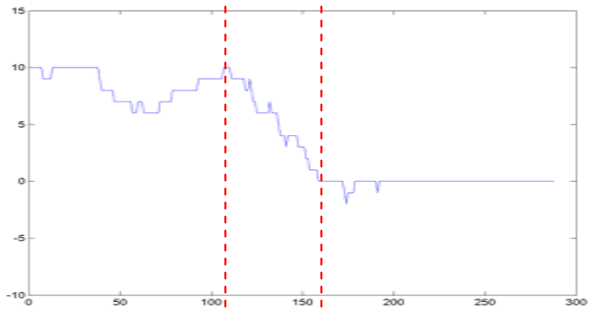

Key
1400 Hz
1600 Hz
*
Start of Gate
Front Vowel
Labiality
/k/
/p/
Cumulative Sum of Detections
Gate 10
Gate 10
Gate 8



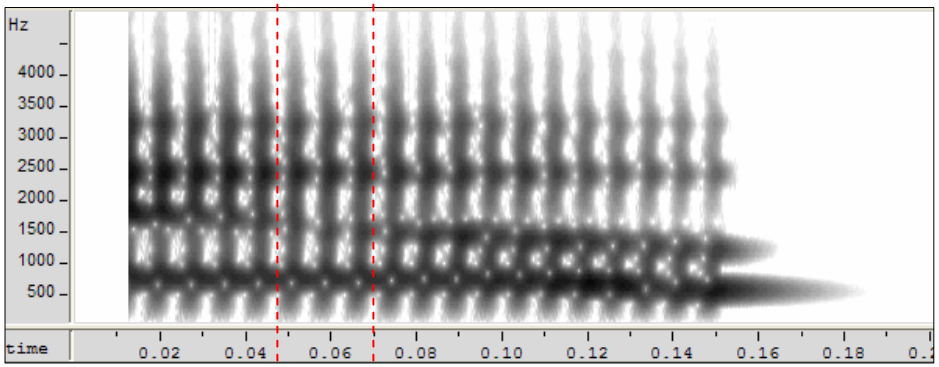
/k/ Splice, 1400 Hz Zero



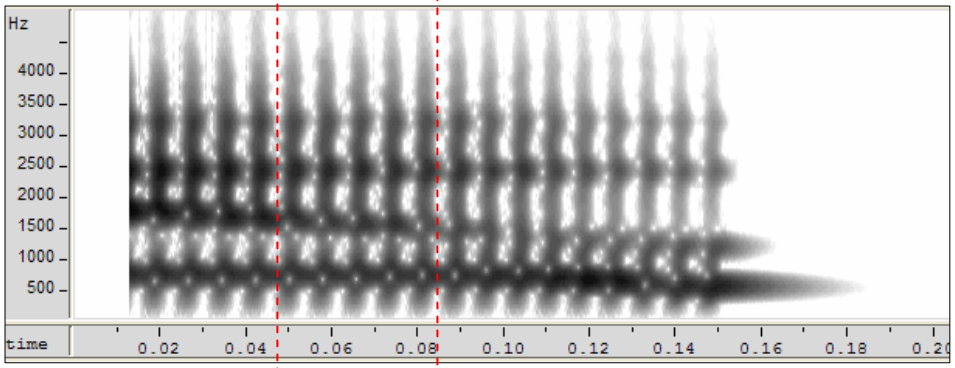
Difference in the number of labiality detections between the 1400 Hz and 1600 Hz stimuli



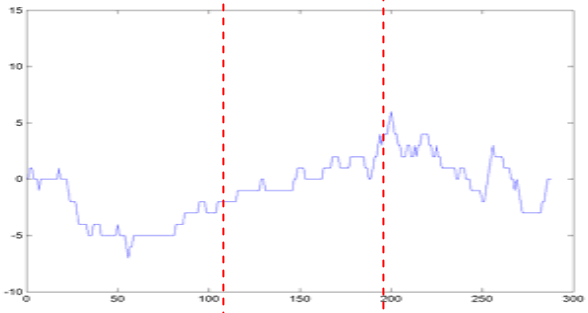
Difference in the number of front vowel detections between the 1400 Hz and 1600 Hz stimuli



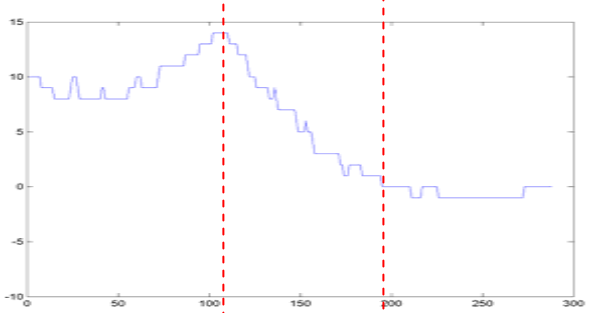
/k/ Splice, 1600 Hz Zero



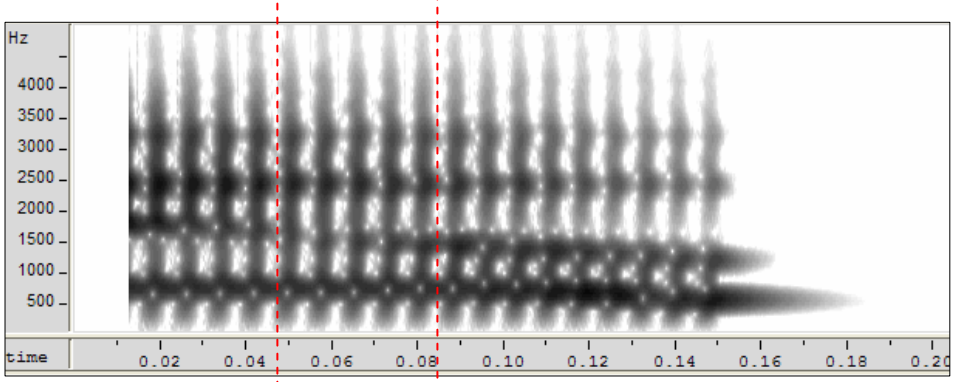
/p/ Splice, 1400 Hz Zero



Difference in the number of labiality detections between the 1400 Hz and 1600 Hz stimuli



Difference in the number of front vowel detections between the 1400 Hz and 1600 Hz stimuli



/p/ Splice, 1600 Hz Zero

Conclusions

- Our ***pilot study*** yielded positive results to the following:
  - The relevance of the interaction between F2 and Z2 for speech perception
  - The potential validity of extending quantal theory from the spectral domain into the temporal domain (landmarks)
  - The importance of dynamic temporal information in the speech signal for perception and identification