QUANTIFYING VOWEL CHARACTERISTICS IN HEBREW AND ARABIC

Noam Amir¹, Osnat Tzenker¹, Ofer Amir¹, Judith Rosenhouse²

¹Dept. of Communication Disorders, Sackler Faculty of Medicine, Tel Aviv University
²Swantech Ltd.
What’s in a formant (or two)?

- Formants are the resonant frequencies of the vocal tract.
- Formants, principally the first two, are the basic characterizations of vowels.
- If you’ve never seen this diagram, you’ve been missing out on life:

The axes are phonological – representing tongue position:

- **High**:
  - **Close**
  - **Open**
- **Mid**:
  - **Close**
  - **Open**
- **Low**:
  - **Close**
  - **Open**

The diagram illustrates the positions of different vowels, with blue indicating unrounded and red indicating rounded.
If you’d rather see numbers…

- Interestingly, the phonological properties correspond roughly to formant values, when the axes are arranged like this:
So who cares?

- Linguists?
- Phoneticians?
- Speech technologists? (TTS? ASR?)

**Most likely**: phoneticians

- But possibly technologists too. So-
  - We’ll tell you about what we found
  - You guys tell us if it’s interesting
Formants of Hebrew

• Formants of prototypical Hebrew were reported by Most et al (2000)
  • Utterances were “Beeg” “Beg” “Bag” “Bog” “boog” in a carrier sentence- “Isolated phonation”
  • Participants: 30 subjects in each group
Question: what happens in everyday speech?

- As opposed to isolated phonation – everyday speech has higher speech rates
- Speakers don’t articulate as well as in isolated phonation -
  - Is the result just some kind of articulatory undershoot?
  - Is it affected by neighboring phonemes?
  - Is it affected by lexical stress?
  - (Do speech technologists care about this?)
  - (Can speech technologists help with this kind of research?)
Formants in spontaneous Hebrew speech

- Participants: 5 men and 5 women (7 of them taken from the CoSIH corpus)
- Participants wore microphones and a recording device
- Out of hours upon hours of speech a research assistant extracted 5,582 vowel tokens:

<table>
<thead>
<tr>
<th></th>
<th>/i/</th>
<th>/e/</th>
<th>/a/</th>
<th>/o/</th>
<th>/u/</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>stressed</td>
<td>288</td>
<td>353</td>
<td>539</td>
<td>248</td>
</tr>
<tr>
<td></td>
<td>unstressed</td>
<td>260</td>
<td>244</td>
<td>509</td>
<td>152</td>
</tr>
<tr>
<td>Women</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>stressed</td>
<td>325</td>
<td>330</td>
<td>547</td>
<td>408</td>
</tr>
<tr>
<td></td>
<td>unstressed</td>
<td>216</td>
<td>217</td>
<td>477</td>
<td>154</td>
</tr>
</tbody>
</table>

- This involved a LOT of manual labor. 
*Could technology help???
Extracting formant values

- Standard software (using LPC) performed pathetically on this task
  - Maybe our mobile phones do LPC 50 times a second…
    - For voice coding this works well: extracting the accurate formant values is not so important
  - However, for formant analysis – sampling frequency and model order must be tweaked constantly to get reasonable results
    - In spontaneous speech this is much more of a challenge than isolated phonation.
- Once again we employed our super-computer: A research assistant
  - Using a custom written Matlab GUI
  - She went through all tokens manually

- This involved a LOT of manual labor. *Could technology help???
Results

- I could talk an hour about these graphs
  ...but would you guys find it Interesting?

- Highlights:
  - Centralization (different in men and women)
  - Stress affects low vowels more than high ones
Centralization was not “pure”

- Spontaneous /a/ moved from stressed to unstressed values in different directions, depending on tongue placement of preceding consonant
- This is called “assimilation to context”
- *Maybe TTS people should know this?*
Formants of Arabic

• Arabic is a tough language to study:
  • **Literary/written/received** Arabic is quite uniform, but not spoken much in daily life
  • **Spoken/Colloquial** Arabic (CA) has many many dialects, but no agreed writing
  • Vowel duration is a distinctive property of a vowel
  • There are very few acoustic studies of Arabic!

• Studying CA –
  • How do you decide how many vowels there are?
  • How do you write down the sentences you want the subjects to utter?
  • Which dialects do you study?
  • How do you make sure your subjects speak these dialects (to quote a student: “my mother is from Jasr, but my father is from Maqr, and I speak something in between…”)

Synopsis of methodology

• Two dialects were studied:
  1. Galilean (GD)
  2. Muthallath (MD)
• 20 men and 20 women from each region participated
• Each subject uttered 30 words in carrier sentences
  • 3 words for each of the 10 Arabic vowels (5 short, 5 long)
  • Emphatic vowels were left out
  • Words were written in a hybridized non-standard form
• Analysis:
  • In this study also – all 1600 tokens were analyzed manually
  • Once more: This involved a LOT of manual labor.

*Could technology help???
Results (formants)

- Scatter plots: (interesting, but hard to read)
Results (formants)

- Polygons:
  - Outer – long vowels
  - Inner – short vowels
  - Clearly – the difference between long and short vowels is not only in duration
  - Differences between dialects are clear, mainly in short vowels
Results (duration)

- No differences between dialects or genders were found
- Looking at absolute and relative durations:
  - Long vowels are about twice as long as short ones
  - **Absolute** Duration is longer for low vowels
  - Duration **ratio** is smaller for low vowels
- *Is any of this important for TTS or ASR people?*
Conclusions:

• Interesting differences were found between isolated phonation and spontaneous speech in Hebrew
  • Also between stressed and unstressed vowels
  • Also between vowel /a/ in different contexts

• Interesting differences were found between two CA dialects
  • Also between short and long vowels
  • Differences involved duration AND formants

• A never ending story: there are hundreds of additional CA dialects
  • Even in such a small country as Israel