

Rhythm, tempo and f0 in language discrimination

Tara Rodriquez & Amalia Arvaniti
UC San Diego



*162nd ASA meeting
Special Session on "Speech rhythm in production, perception and acquisition"
4th November 2011*

Acknowledgements

Tara Rodriquez



- Spyros Raptis (ILSP, Athens) and his team
- Carolyn Quam (UCSD/U. Arizona)
- Cody Brimhall (formerly UCSD; now Apple)
- UCSD Committee on Research for funding


Rhythm classes and empirical evidence

- No strong evidence for rhythm classes has been found in production (Bertinetto 1989; Kohler 2009; Arvaniti 2009; Arvaniti to appear/2012; Horton & Arvaniti, forthc.)

But....

- There is evidence from perception that supports the idea of rhythm classes
- Processing experiments (Cutler and colleagues)
- Discrimination experiments (e.g. Ramus, Dupoux, Mehler 2003)

Some caveats

- Discrimination experiments involve English, Dutch, Spanish, Japanese
- These languages are characterized by substantial differences in tempo 
- Discrimination could be due to a number of factors
- Could tempo be the reason behind the successful results of these experiments?



Some additional evidence

- Ramus et al (2003): AAX (oddball) experiments
- (Stress-timed) Polish was discriminated from both English and Spanish/Catalan
- Polish stimuli were longer in duration and shortened proportionally
 - increased tempo
(by approximately 1 syllable/second)

Hypotheses

- Discrimination is due to differences in tempo
- Fundamental frequency (f_0) should aid discrimination
 - and could lead to different responses (Arvaniti, in press)

Experiments

- Five AAX experiments
- Context (AA): English
- Test (X): Danish, Spanish, Greek, Korean, Polish
- 2 m 2 f monolingual speakers of English (2 for context, 2 for controls)
- 1 m, 1 f monolingual speaker each of Danish, Spanish, Greek, Korean and Polish

Stimuli

- sentences converted to **sasasa**
 - consonantal intervals convert to [s]
 - vocalic intervals convert to [a]
- diphone synthesis using STRAIGHT in Matlab
(thanks to Spyros Raptis)
 - voices of one male and one female Greek speaker
 - diphone taken from the stressed syllable of the name ['sasa]



Conditions: tempo & f0

- **Original tempo:** that of the original utterances, measured in syllables/second (after *sasasa* conversion)
- **Original f0:** modulation of the original utterances
- **Same tempo:** all stimuli have average tempo of the two languages in each experiment
- **Same f0:** slightly declining f0



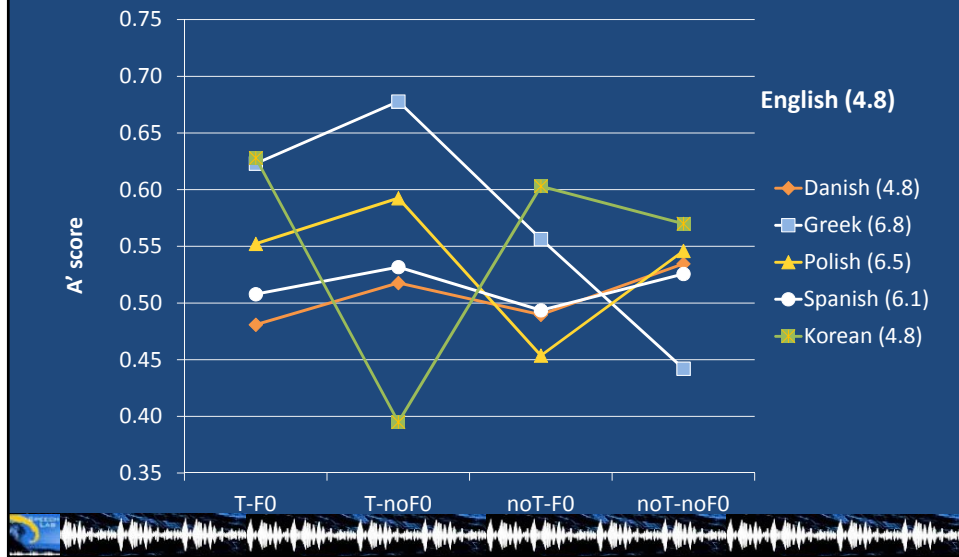
Tempo × F0

- Original tempo, original F0: T-F0
 Original tempo, flat F0: T-noF0
 Same tempo, original F0: noT-F0
 Same tempo, flat F0: noT-noF0

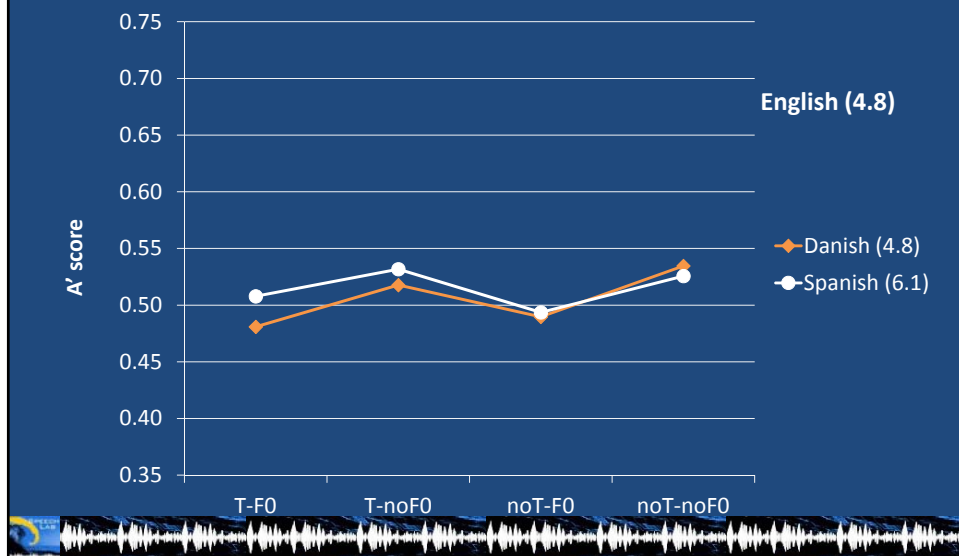
Language choice & predictions

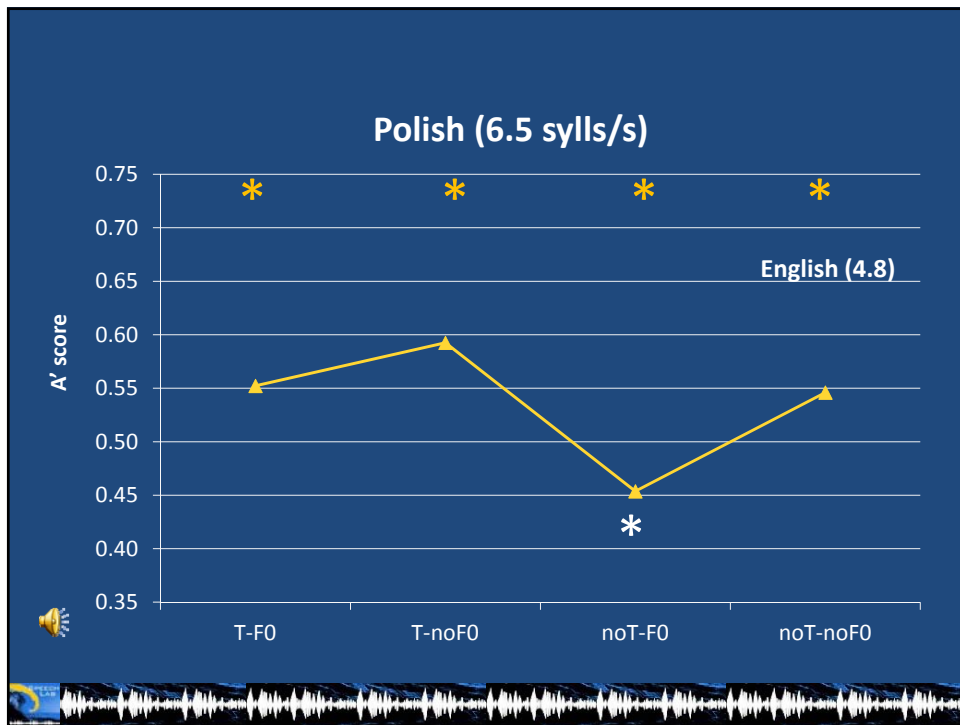
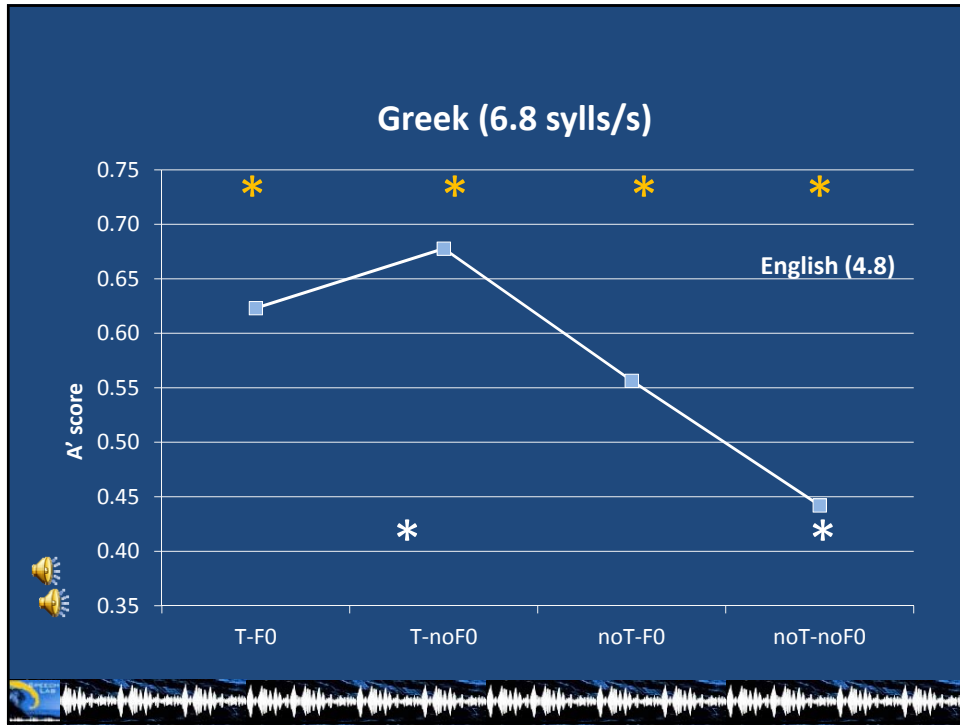
language	rhythm class	tempo wrt English	predictions
Danish	stress-timed	D ~ E	ND - ND
Polish	stress-timed	P > E	ND - D
Spanish	syllable-timed	S > E	D - D
Greek	((syllable-timed))	G > E	D - D
Korean	((syllable-timed))	K ~ E	D - ND

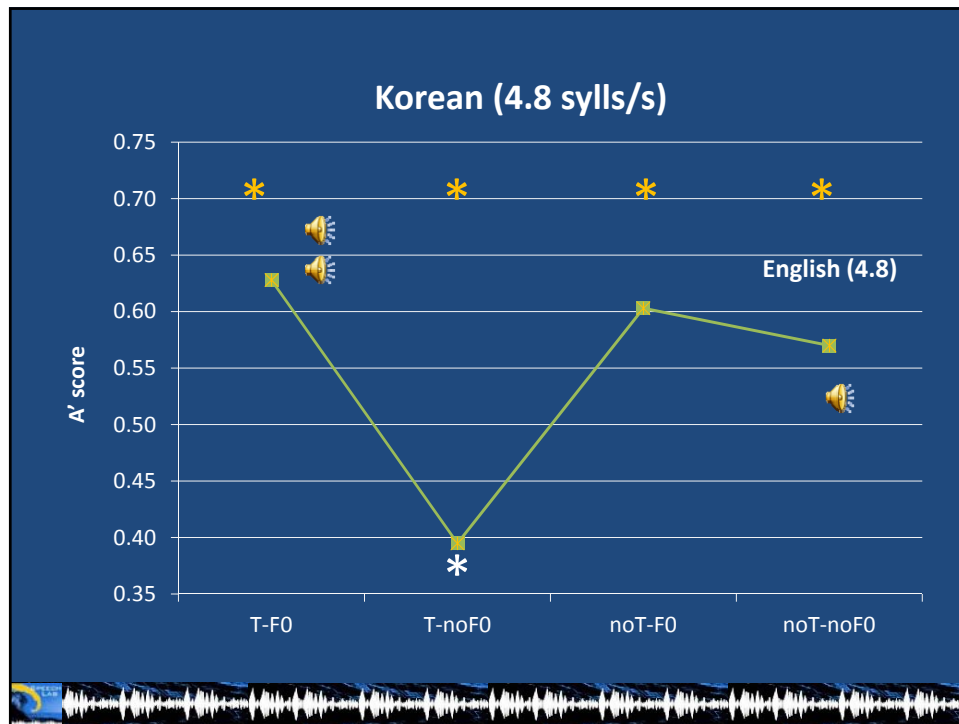
Tempo × f0 × language interaction



Spanish and Danish





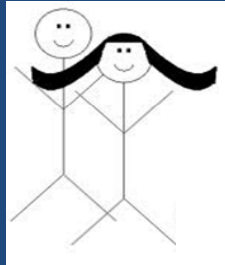


Discussion

- Language discrimination using impoverished stimuli is not an easy task
- Listeners focus on whatever aspects of the signal will help them with the task
- One such salient feature is tempo
- Failing tempo, F0 information is used if sufficiently different between languages
- Otherwise F0 can confuse listeners
- Timing plays a part when all else fails

Arvaniti & Rodriguez forthc.

What “rhythm discrimination” is like



Based on Hawkins, 2003

Conclusion

- The oddball paradigm does not really support the idea of rhythm classes
- Previously results are probably due largely to a confound between rhythm and tempo
- Timing is probably not a dimension of the speech signal that is independently processed
 - *sasasa* is not an ecologically valid method of signal transformation
- Rhythm is a multi-dimensional phenomenon
- More scrutiny of results said to support rhythm classes is needed

Arvaniti & Rodriguez forthc.