

Constructing Physically Realistic VCV Stimuli for the Perception of Stop Voicing in European Portuguese



Daniel Pape
Luis M. T. Jesus
Pascal Perrier



universidade
de aveiro



PROPOR 2012, Coimbra

Outline

- Project description
- Biomechanical model (Perrier et al. 2003)
- Fitting EP data to the biomechanical model
- Recording of EP production data
 - Vowel and consonant durations
 - Stop voicing maintenance
- Synthesis procedure

Project description

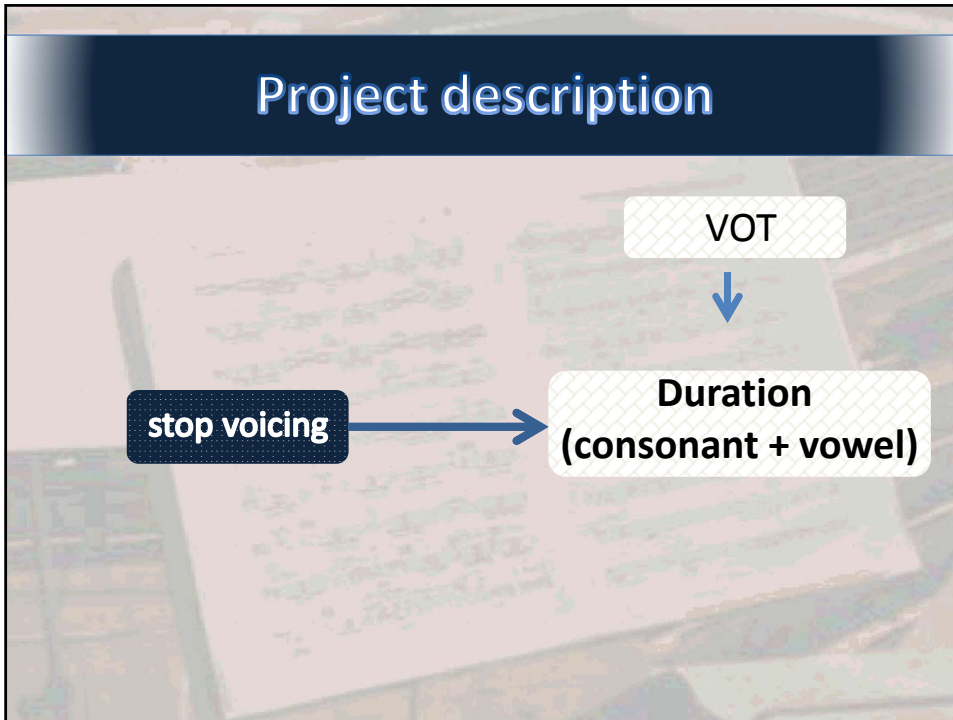
stop voicing

Project description

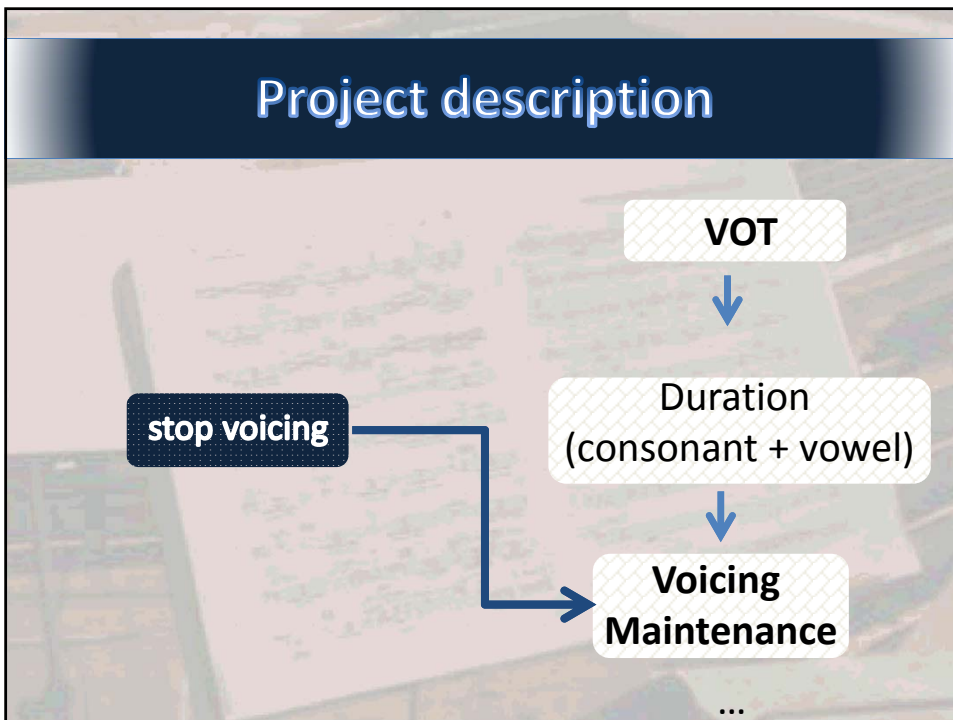
stop voicing

Voice
Onset
Time

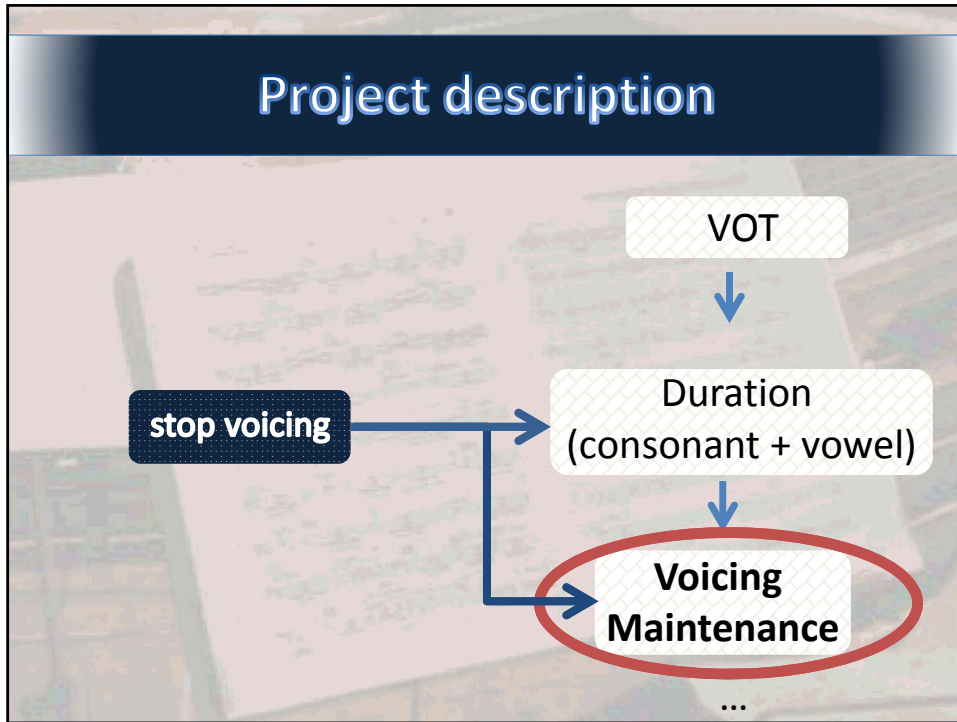
Project description



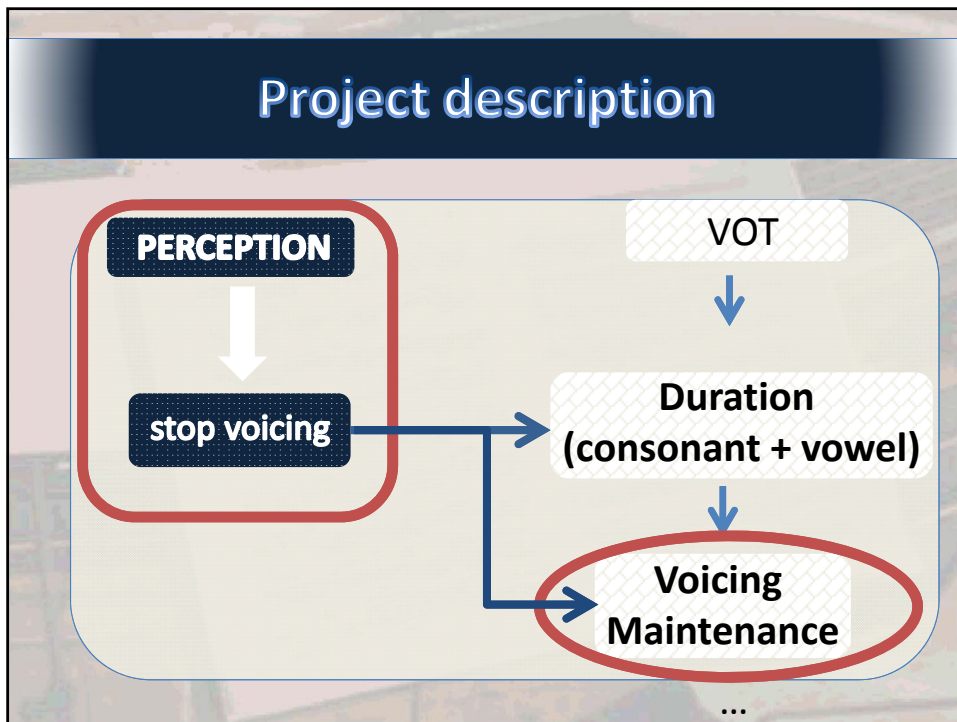
Project description



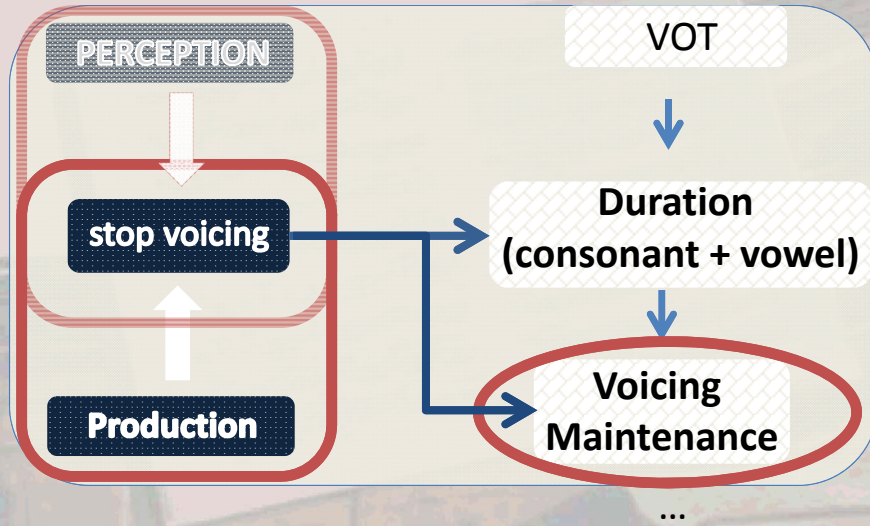
Project description



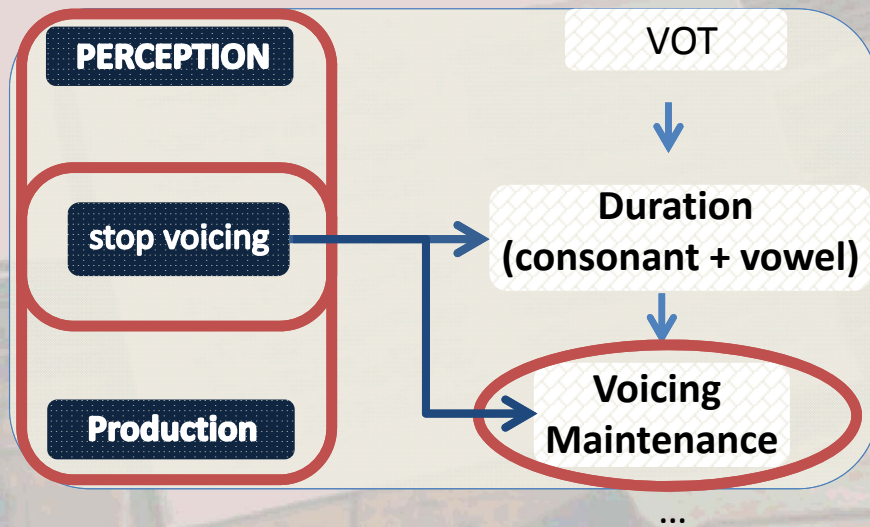
Project description



Project description



Project description



Voicing maintenance: Romance languages (selection)

- **Spanish:**
 - Solé (2011): high voicing and pre-voicing for stops
 - Shih & Möbius (1999): clear contrast to Germanic languages with higher voicing maintenance for Romance languages
- **European Portuguese:**
 - Jesus & Shadle (2003): high occurrences of devoicing for EP fricatives
 - Lousada, Jesus & Hall (2010): EP phonologically voiced stops often produced devoiced

Voicing maintenance: EP

- **Research questions for EP production:**
 - How and where does devoicing occur during the production of the stop closure?
 - Are the durational values in accordance to the literature (Veloso 1995, ...)?
- **Research questions for EP perception:**
 - Is (voicing) perception influenced by differences in devoicing patterns?
 - Is there a perceptual interaction with the durational cues (consonant and vowels)?
 - How robust is perception with a missing burst cue?

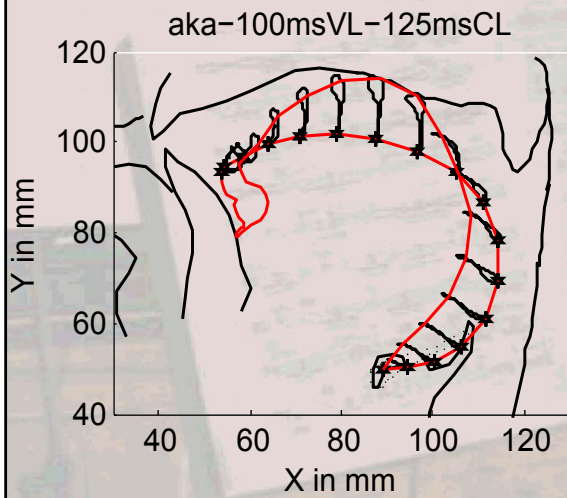
Synthesis for perceptual experiments

- Demand of highly natural synthesis
- No uncontrolled cues left (which rules out manipulating natural speech)
- Independent parameterization of:
 - Vowel and consonant durations
 - Transition lengths
 - Voicing maintenance
- “Physically correct” transitions between phonemes

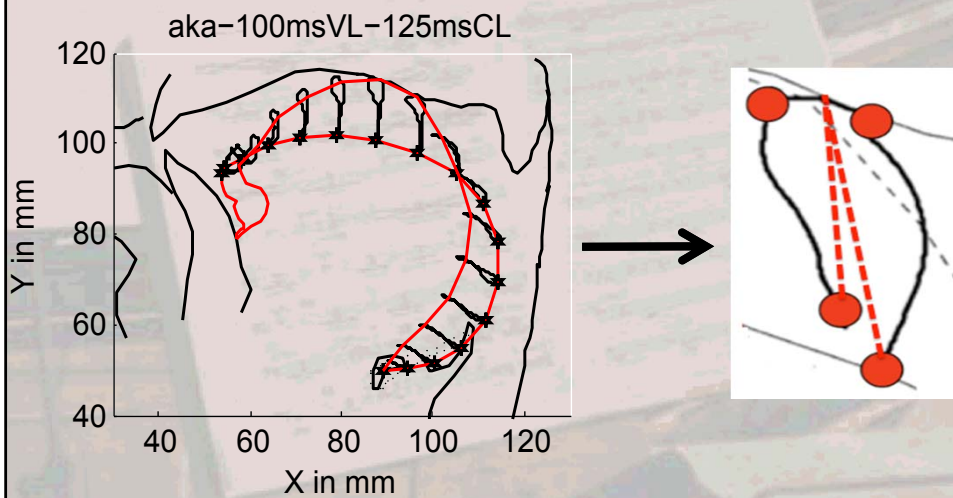
Synthesis for perceptual experiments

- Demand of highly natural synthesis
- No uncontrolled cues left (which rules out manipulating natural speech)
- Independent parameterization of:
 - Vowel and consonant durations
 - Transition lengths
 - Voicing maintenance
- “Physically correct” transitions between phonemes

“Articulatory Loops”



“Articulatory Loops”

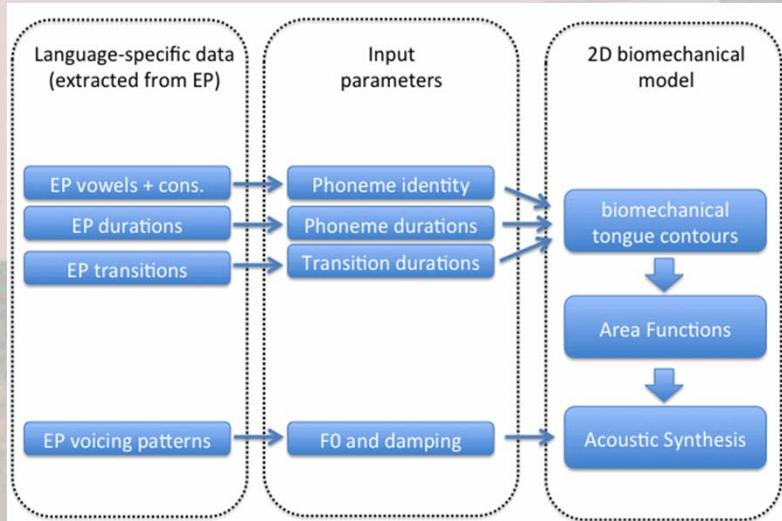


Solution: Biomechanical model

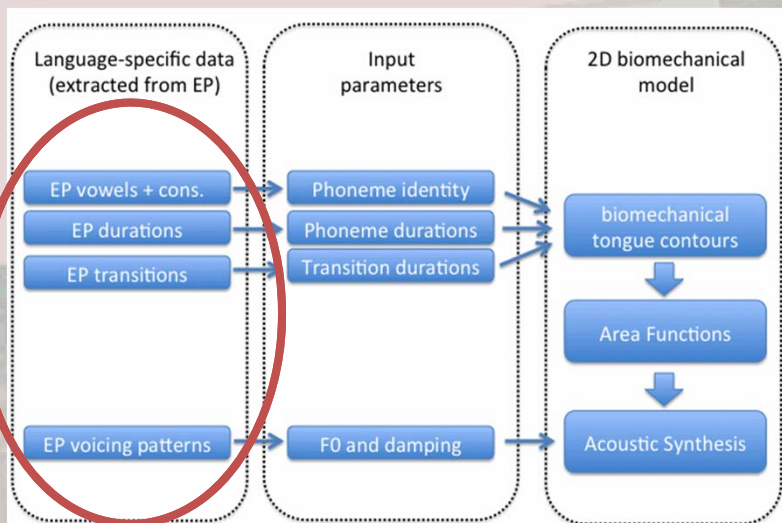
- Provide “physically correct” transitions between phonemes
- Display the dynamics of muscle movements
- Have been shown to respect physical laws of human movements (Fuchs & Perrier 2008), similar to hand and joint movements (Viviani 1992)

Demo biomechanical model

Using the biomechanical model with EP data



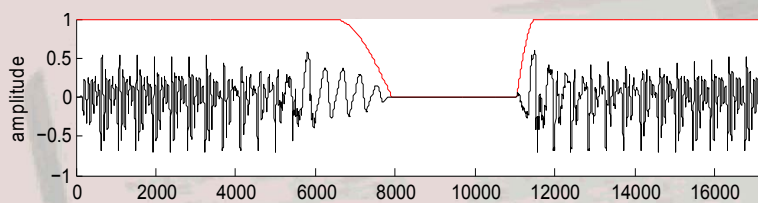
Using the biomechanical model with EP data



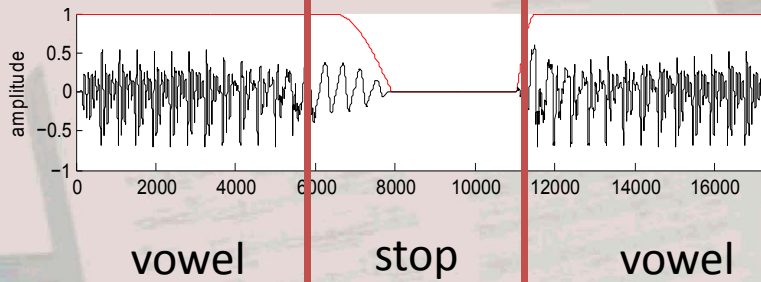
EP Database recording – Durations and voicing maintenance

- EGG and acoustic recordings
- 6 EP speaker (Aveiro)
- Vowels /a e i o/
- All EP stops and fricatives
- Initial and medial position CVCV in frame sentence
- 10 repetitions
- Analysis:
 - (contextual) vowel and consonant durations
 - obtain the voicing status for 10 equidistant landmarks throughout the consonant duration, allowing for a devoicing curve

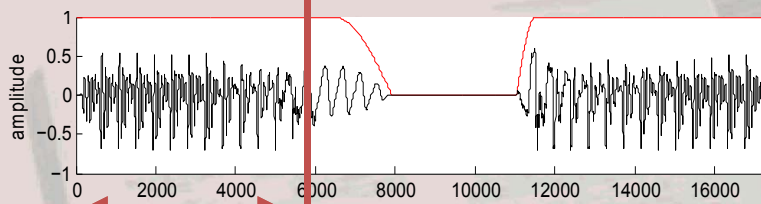
Perception experiments: Stimuli generation



Perception experiments: Stimuli generation



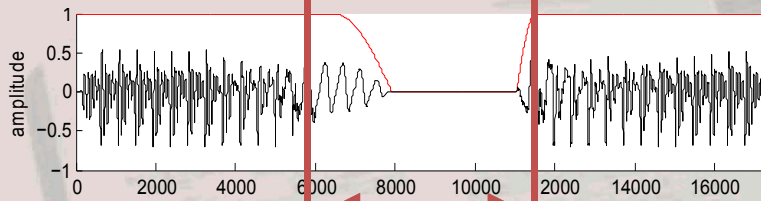
Generate the first continuum



steps	vowel duration (ms)
1	70
2	100
3	130
4	
5	

(1)

Generate the second continuum

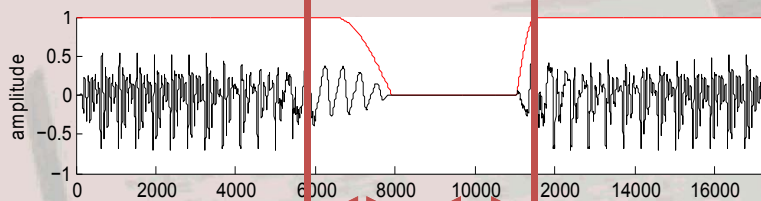


(2)

consonant duration
(ms)

50
75
100
125
150

Generate the third continuum



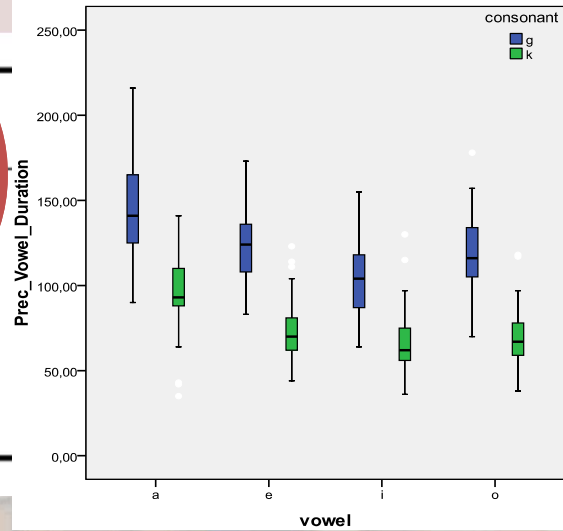
(3)

consonant voicing (%)

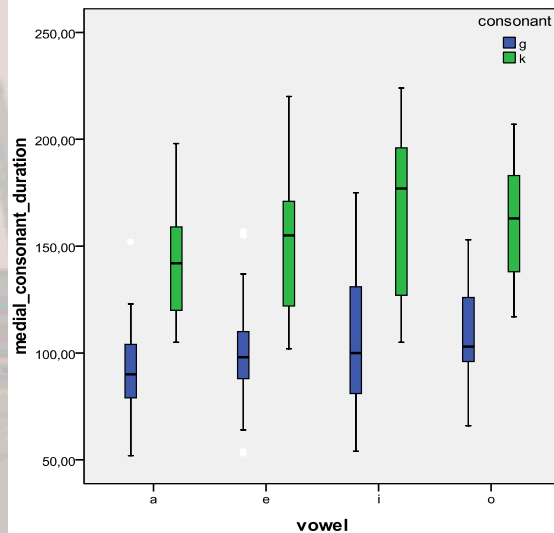
0 (fully devoiced)
12.5
25
37.5
50
75
100 (fully voiced)

(1): Contextual vowel durations

steps	vowel duration (ms)
1	70
2	100
3	130
4	
5	
6	
7	

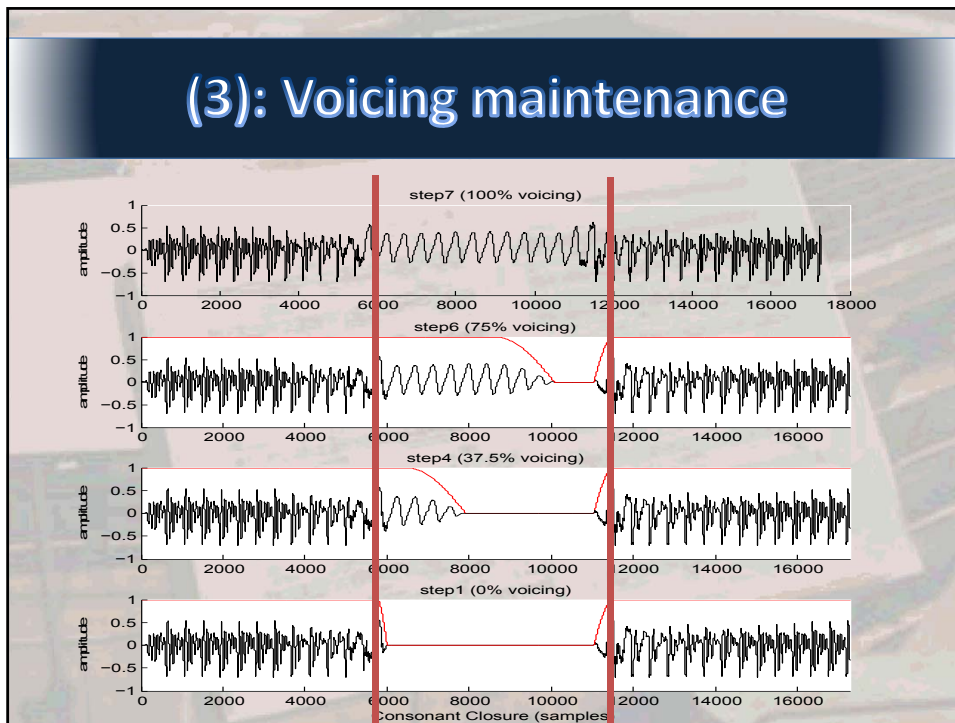


(2): Stop durations



consonant duration (ms)
50
75
100
125
150

(3): Voicing maintenance



Summary: three stop voicing continua

(1)

steps	vowel duration (ms)	consonant duration (ms)	consonant voicing (%)
1	70	50	0 (fully devoiced)
2	100	75	12.5
3	130	100	25
4		125	37.5
5		150	50
6			75
7			100 (fully voiced)

Summary: three stop voicing continua

(2)

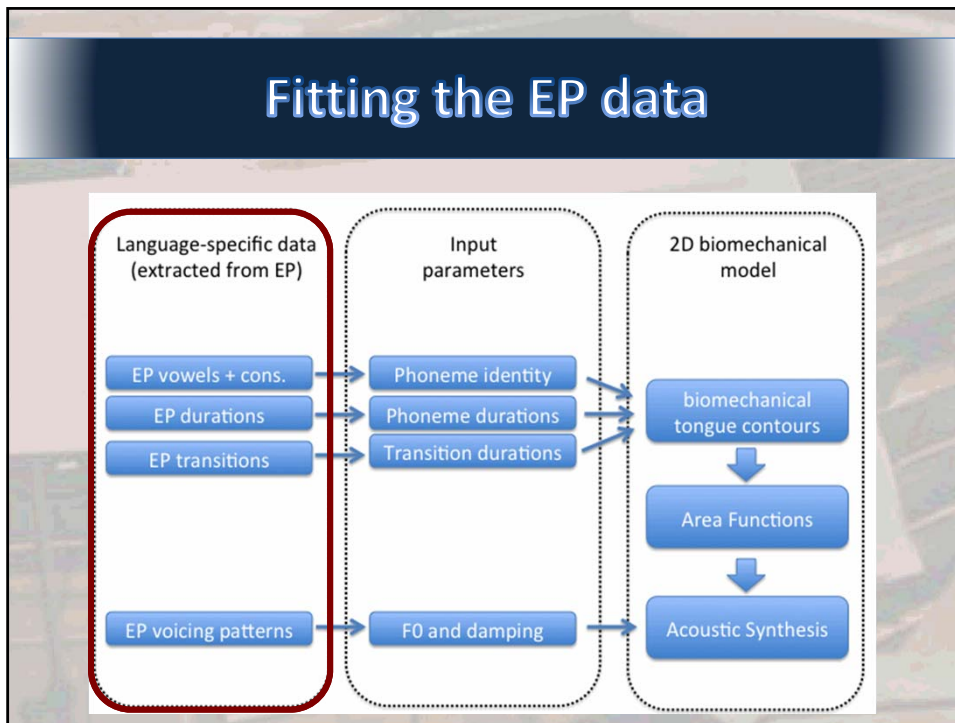
steps	vowel duration (ms)	consonant duration (ms)	consonant voicing (%)
1	70	50	0 (fully devoiced)
2	100	75	12.5
3	130	100	25
4		125	37.5
5		150	50
6			75
7			100 (fully voiced)

Summary: three stop voicing continua

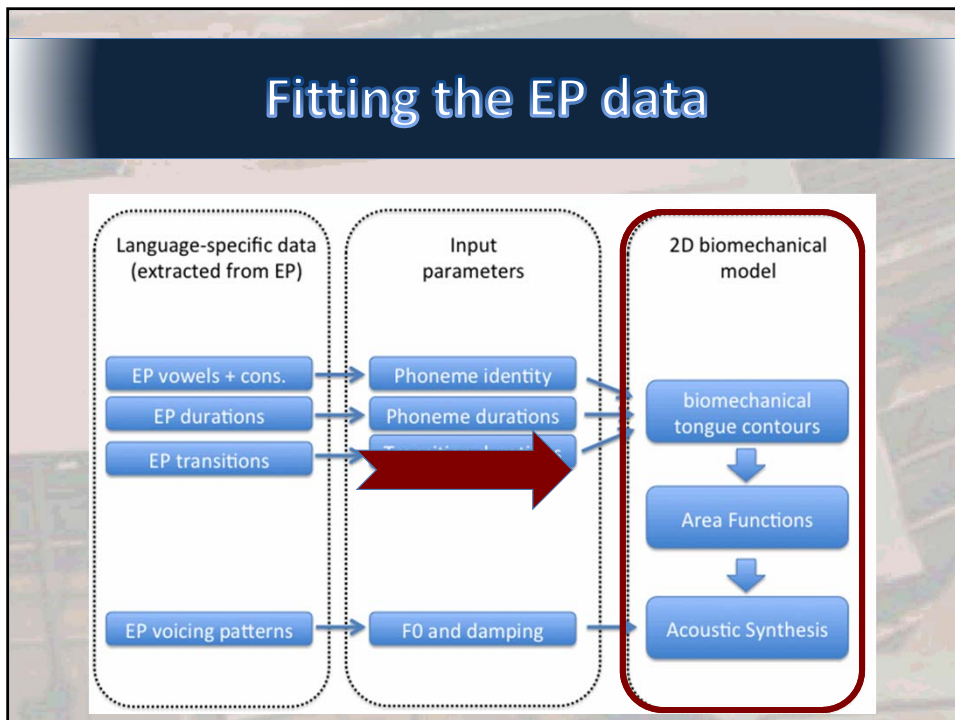
(3)

steps	vowel duration (ms)	consonant duration (ms)	consonant voicing (%)
1	70	50	0 (fully devoiced)
2	100	75	12.5
3	130	100	25
4		125	37.5
5		150	50
6			75
7			100 (fully voiced)

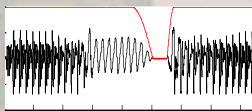
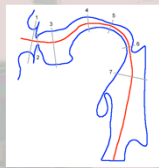
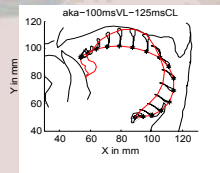
Fitting the EP data



Fitting the EP data



Overview: Synthesis procedure

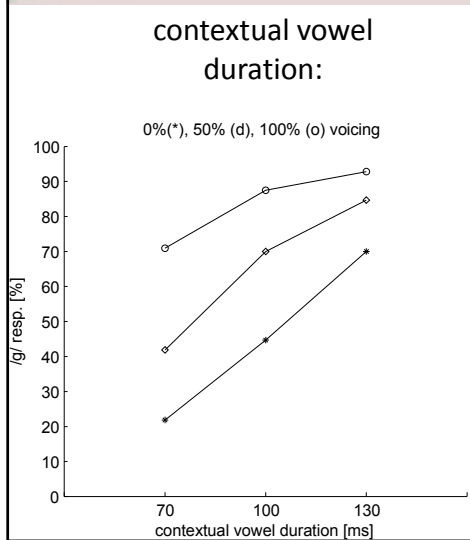


- Biomechanical model:
 - Generation of muscle commands
 - Controlled by the parameters **hold time** and **transition time**
 - Output is a two-dimensional tongue contour in a standard vocal tract
- Computation of time-dependent area functions
- Acoustic synthesis with a four mass vocal fold model (Story 2004, Story 2005)

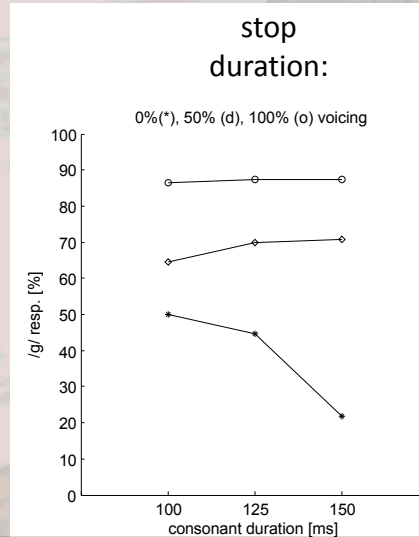
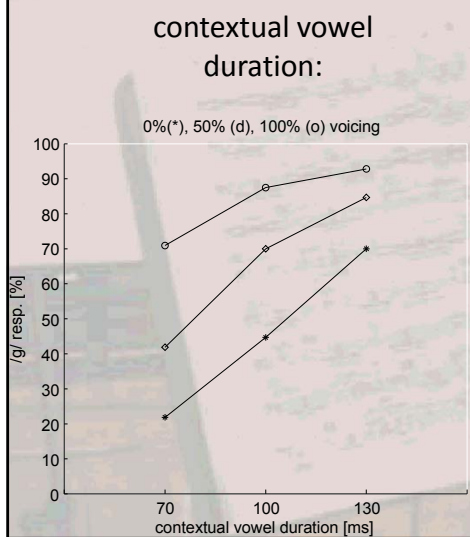
Thank you!

- This work was partially funded by a post-doctoral fellowship from FCT (Portugal) grant **SFRH/BPD/48002/2008** to Daniel Pape.
- This work was partially funded by FEDER through the Operational Program Competitiveness Factors - COMPETE and by Portuguese National Funds through FCT - Foundation for Science and Technology in the context of the project FCOMP-01-0124-FEDER-022682 (FCT reference **PEst-C/EEI/UI0127/2011**) to IEETA (Aveiro).

Preliminary data: Voicing perception in dependence of durations (n=15)



Preliminary data: Voicing perception in dependence of durations (n=15)



Preliminary data: Voicing perception in dependence of durations (n=15)

