

# Automatic Analysis of Singleton and Geminate Consonant Articulation using Real-time Magnetic Resonance Imaging

Christina Hagedorn<sup>1</sup>, Michael Proctor<sup>1,2</sup>, Louis Goldstein<sup>1</sup>, Shrikanth Narayanan<sup>2</sup>

<sup>1</sup>Department of Linguistics, University of Southern California, USA

<sup>2</sup>Viterbi School of Engineering, University of Southern California, USA

## Motivation

How do singletons and geminates of Italian differ in terms of the **gestural control** that their underlies their production?

## Overview

Investigating production of Italian singleton (C) and geminate (C:) consonants (e.g. 'baco' [bako] – 'bacco' [bak.o]) using Real-time MRI.

If control is primarily **temporal**, then:

- **Constriction duration** should differ for C and C:
- **Constriction location** should be invariant for C and C:
- **Constriction degree** may differ for C and C: (due to undershoot)

→ Tested using novel automatic techniques

## Background

### Real-time MRI

Provides dynamic information about all components of the vocal tract

Analyses performed directly on time functions of pixel intensities, not segmented regions [1]

### Italian length contrast

Production of singleton and geminate consonants in standard Italian differs both temporally and spatially (Bertinetto, 1981; Gill Fivela et al., 2007; Payne, 2006; Smith, 1995; Bertinetto and Loporcaro, 2005; Cerrato and Falcone, 1998; Esposito and di Benedetto, 1999)

- Geminates:
  - Longer constriction duration
  - Higher peak vertical (lower) lip position
  - More linguo-palatal contact
- Velocity, stiffness differences of singletons and geminates unclear
- Lack of dorsal data due to physiological limitations of EPG and EMA.

## Experimental method

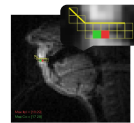
- Adult male speaker of standard Italian (Rm)
- Stimuli in carrier phrases: singleton and geminate bilabial, coronal and dorsal stops, affricates and sonorants (/p/, /m/, /t/, /d/, /n/, /ɲ/, /f/, /dʒ/, /k/ and /g/)
- Vocal tract imaged midsagittally
  - Spatial resolution: 68 x 68 pixels (200 x 200 mm.)
  - Temporal reconstruction rate: 33.8 f.p.s.

## Analytical methods & Results

### Constriction location

#### Technique:

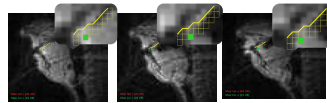
- Defined **dynamically**: site of **maximal constriction change**
- Pixel for which intensity changes most systematically
- Determined by:
  - Pixel(s) whose intensity is most correlated with neighboring pixels' intensities
  - Pixel(s) with greatest dynamic range of intensity



[apa]

**Automatic location of bilabial constriction:**  
Pixels of **maximal correlation** and **maximal dynamic range of intensity**

### Validation of automatic technique:

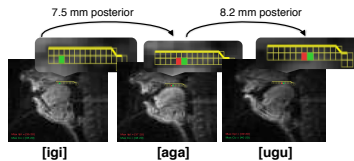


[iti]

[ata]

[utu]

**Coronal stop constriction location invariant across different vowel contexts**

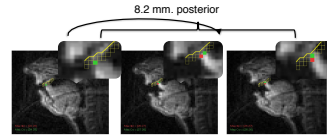


[igi]

[aga]

[ugu]

**Dorsal stop constriction location varies across different vowel contexts**



[ata]

[ane]

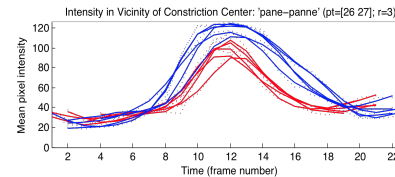
[ale]

**Automatic location of coronal constriction**

**Singleton and geminates do not differ in constriction location for all places of articulation**

### Constriction Degree

- High pixel intensity: compression and movement
- Cohort of pixels: falls within radius of 3 pixels from optimal pixel of maximum dynamic intensity

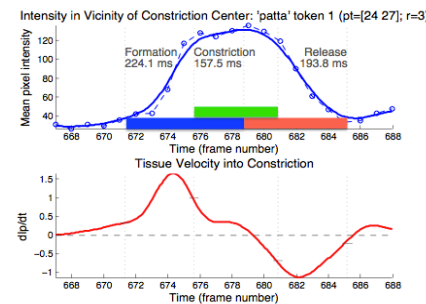


**Intensity Function:** Cohort of pixels centered at constriction location during production of [ane] and [an:e]

**Intensity: C < C: (p<.001) for all places of articulation**

### Constriction kinematics

- First difference of intensity function: tissue velocity into and out of specified region
- Identify thresholds in the velocity function to estimate salient kinematic events



**Intensity and Velocity Functions:** Constriction formation (blue bar), constriction duration (green bar), release (red bar) in the production of [atta]

**Formation duration**  
**Constriction duration**  
**Consonant duration** } **C < C: (p<.05) for all places of articulation**

**Peak velocity: inconsistent across places of articulation**  
**When significant, C < C: (p<.05)**

## Conclusion

- rMRI and analytical methods evaluated are capable of capturing and quantifying
  - Constriction location
  - Constriction degree (compression)
  - Salient kinematic events
- Italian singletons and geminates differ in:
  - Duration
  - Constriction degree
- Italian singletons and geminates do *not* differ in place of articulation
- Future work
  - Investigate possible differences in stiffness (measured by time to peak velocity)
  - Stiffness differences could account for differences in peak velocity despite singletons and geminates having the same virtual target.

## Selected References

- [1] A. Lammert, M. I. Proctor, and S. S. Narayanan, "Data-driven analysis of realtime vocal tract mri using correlated image regions," in Proc. Interspeech, Makuhari, Japan, 2010.
- [2] A. Lófqvist, "Lip kinematics in long and short stop and fricative consonants," JASA, vol. 117 no. 2, pp. 858-878, 2005.
- [3] S. Narayanan, K. Nayak, S. Lee, A. Sethy, and D. Byrd, "An approach to real-time magnetic resonance imaging for speech production," JASA, vol. 115, no. 4, pp. 1771-1776, 2004.
- [4] E. Bresch, J. Nielsen, K. Nayak, and S. Narayanan, "Synchronized and noise-robust audio recordings during real-time magnetic resonance imaging scans," JASA, vol. 120, no. 4, pp. 1791-1794, 2006.

## Acknowledgments

Research supported by NIH Grants R01 DC007124-01 and DC008780-01.