

# TIME TO TARGET ATTAINMENT IN SINGING AND SPEECH

Reed Blaylock | reed.blaylock@gmail.com | University of Southern California

## 1 INTRODUCTION

How do the goals of different speech tasks—like talking and singing—affect speech motor control?

### DIFFERENT TYPES OF SPEECH HAVE DIFFERENT GOALS

Singing and other varieties of musical speech have properties we don't observe in research about talking.

#### Talking

- Vowels are as long as they need to be
- Isochrony not imposed
- Primarily communicative goals

#### Singing

- Very long vowel durations
- Isochronous rhythm
- Aesthetic and communicative goals

These different goals have measurable consequences for speech kinematics (e.g., phrase-final lengthening). Once we know them, we can model speech motor control beyond talking.

### DO SINGING AND TALKING REACH TARGETS DIFFERENTLY?

#### Hypothesis 1

An aesthetic goal for long, clear vowels results in vowel constrictions being made at least as early as in speech. There would be **no difference** between singing and talking.

#### Hypothesis 2

Because sung vowels are so long, vowel gestures can afford to take more time to reach their targets. It would take **longer in singing** for vowel constrictions to reach their targets.

Phrase-final lengthening in singing: Ramanarayanan et al. 2011

## 2 METHOD

I measured how long it takes for the tongue body to reach target vowel constrictions in talking and singing real-time MRI videos.

### DATA & MEASUREMENTS

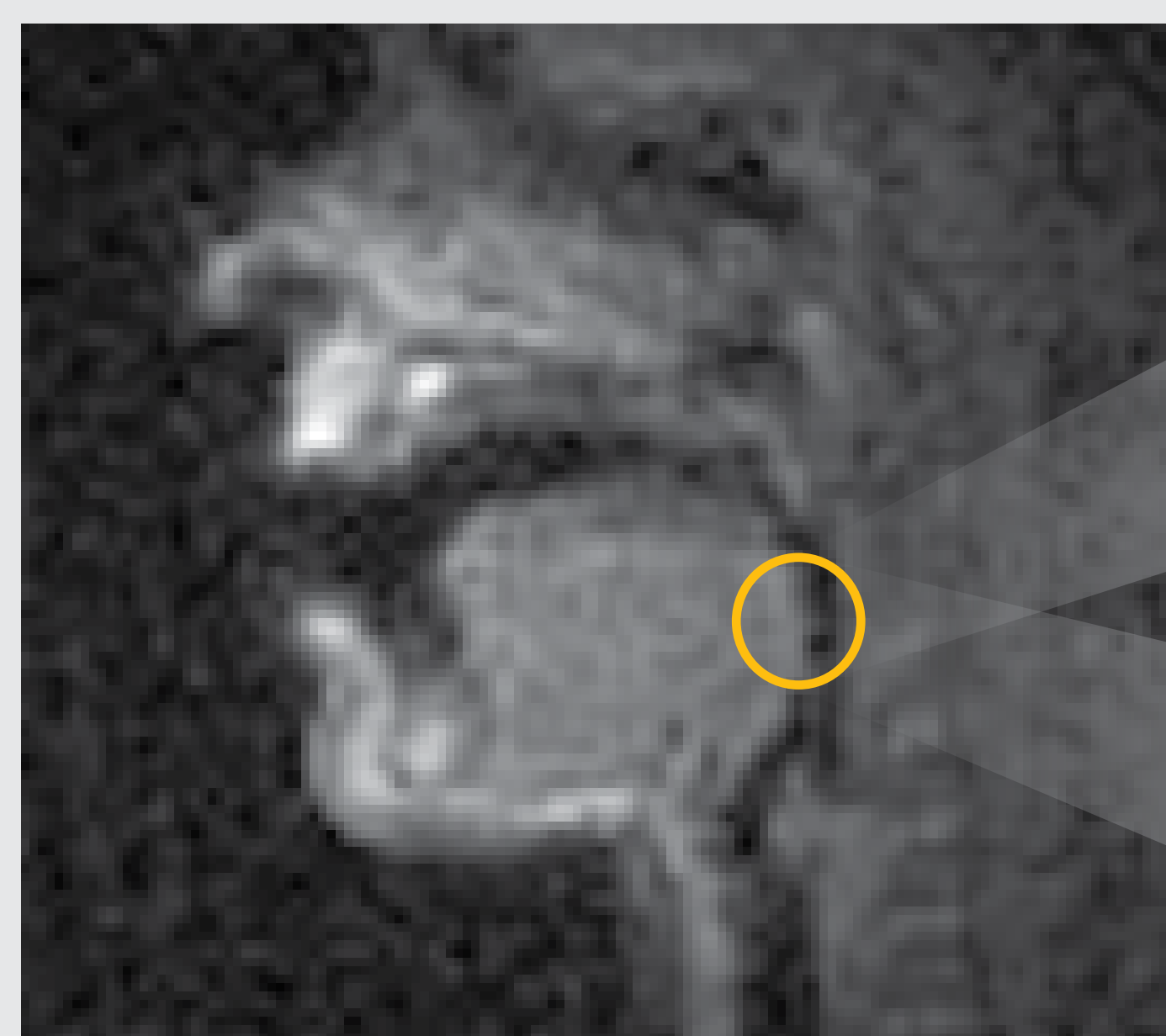
One trained soprano from USC School of Music

Measured tongue body movement for spoken and sung low and mid-low vowels:

- Spoken vowels from TIMIT sentences
- Sung vowels from songs the singer had memorized

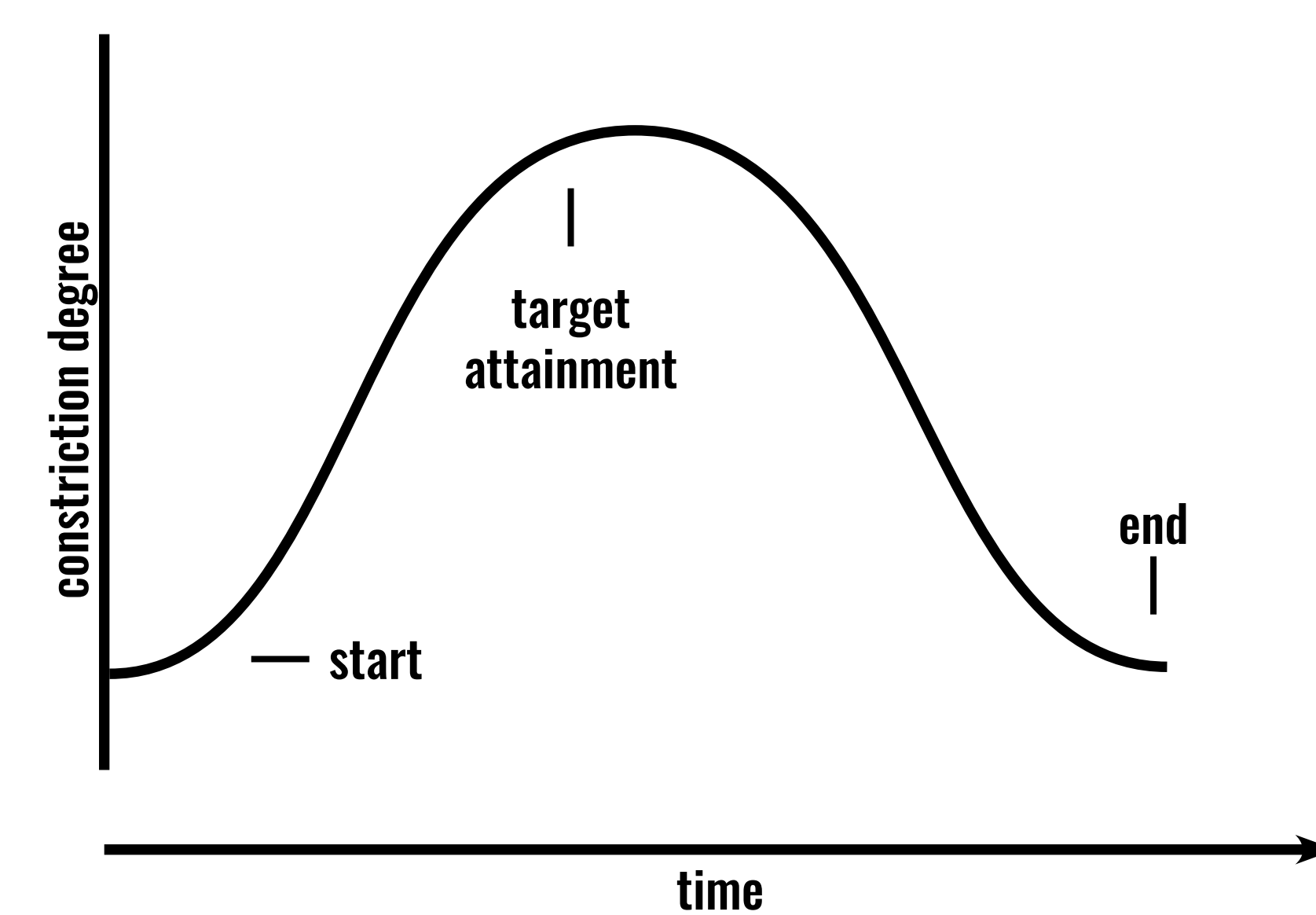
Region of interest analysis with automatic (hand-corrected) "target attainment" landmarks

### REAL-TIME MRI REGION OF INTEREST MEASURES TONGUE BODY MOVEMENTS



narrow

wide

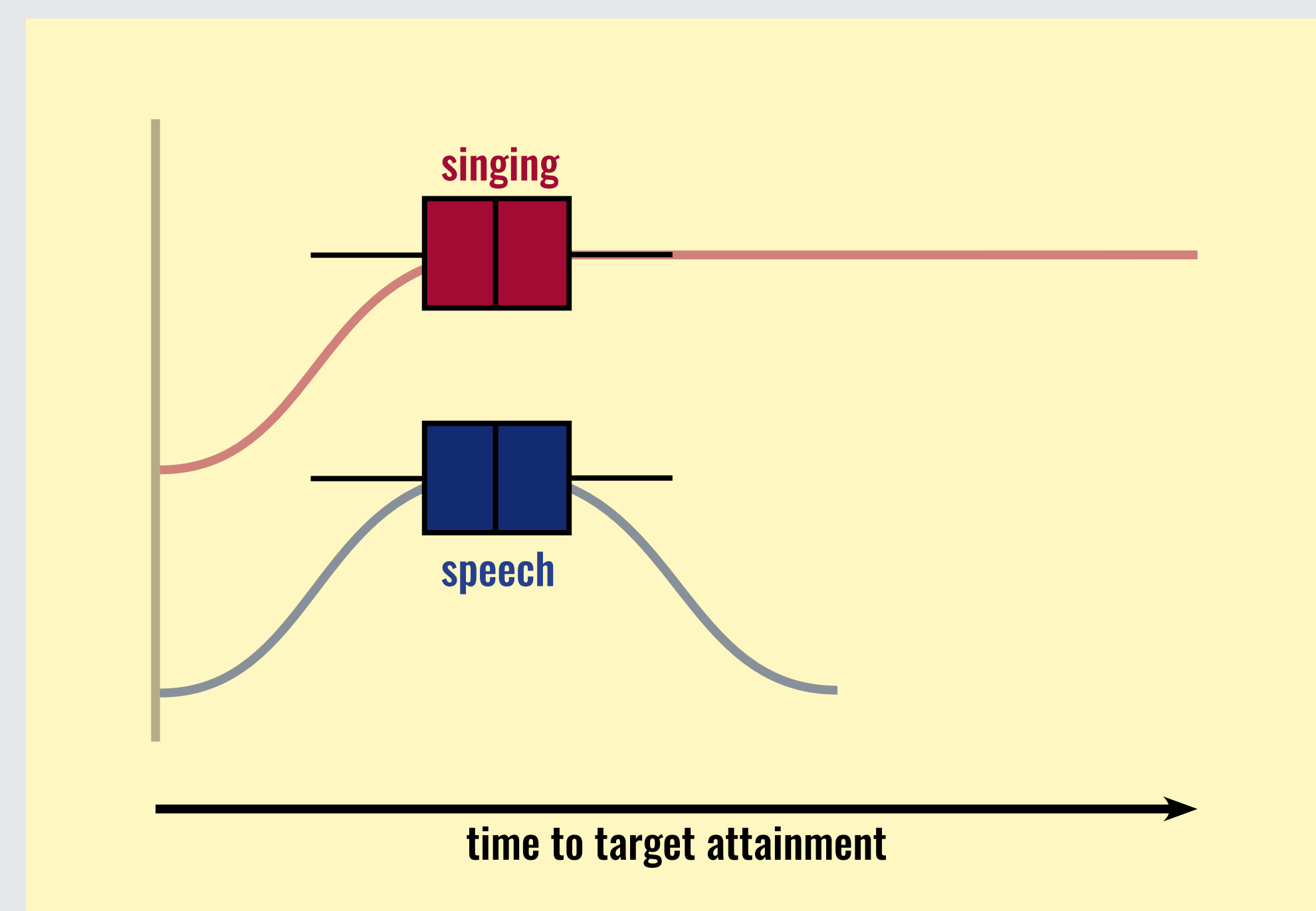


Soprano real-time MRI: Bresch & Narayanan 2010. Region of interest technique: e.g., Lammert et al. 2013. Articulatory landmark finding: Tiede 2010.

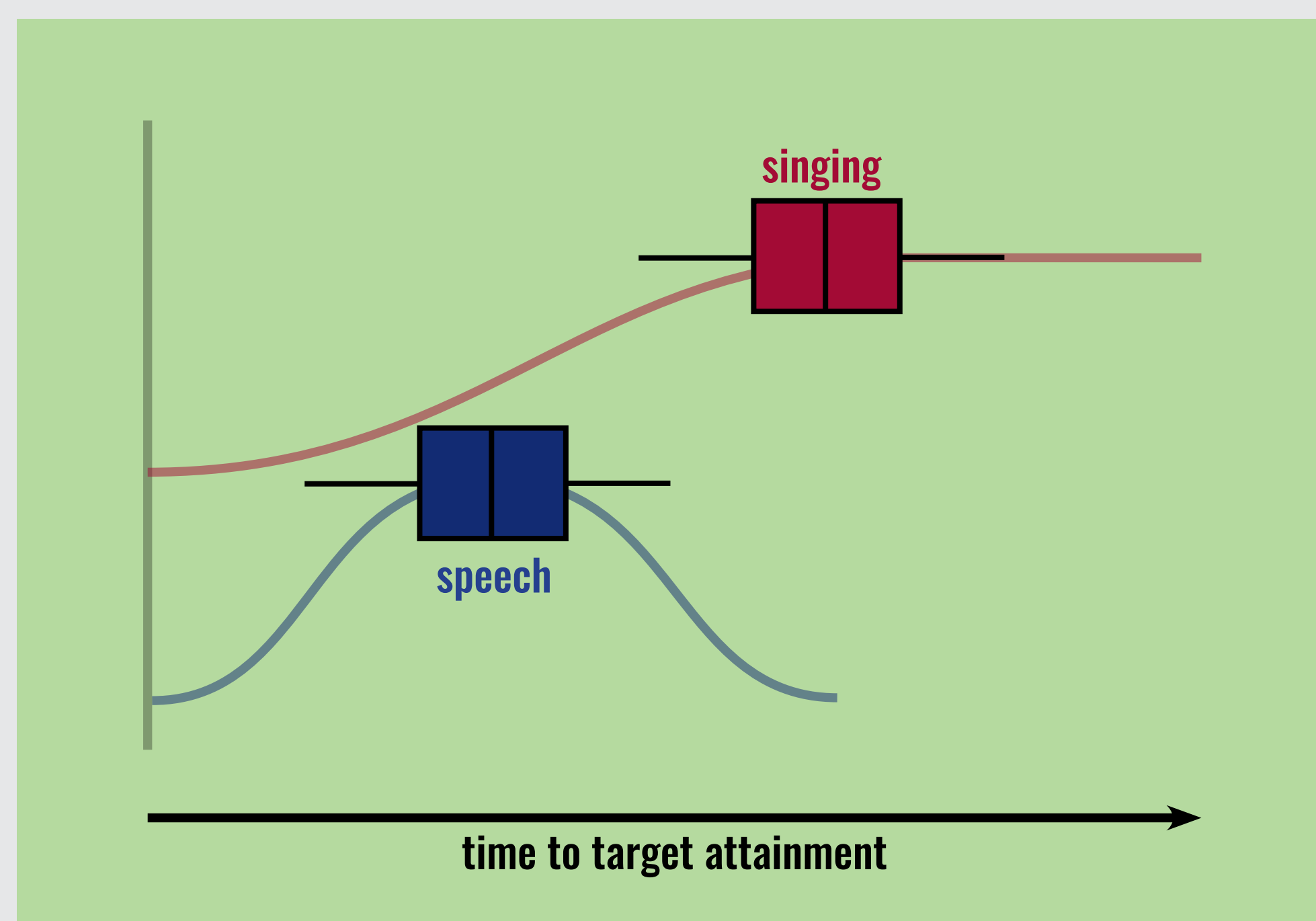
## 3 RESULT

In this study, it took longer for sung tongue body movements to reach their goal than it took spoken tongue body movements.

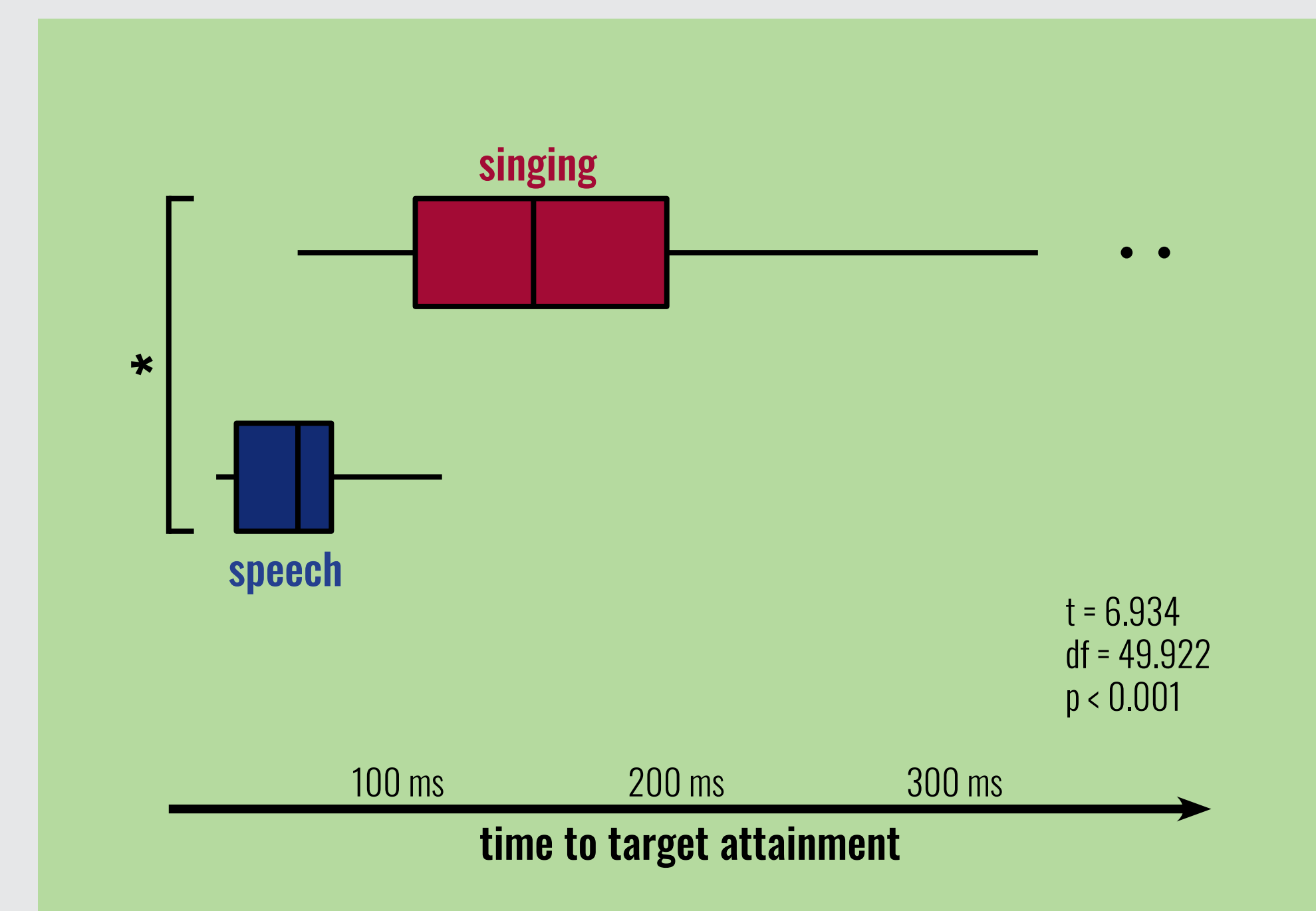
### HYPOTHESIS 1: NO DIFFERENCE



### HYPOTHESIS 2: LONGER IN SINGING



### RESULT: LONGER IN SINGING



## 4 CONCLUSION

Speech tasks are complex, and we need to learn more about musical speech to understand how speech actions adapt.

### SUNG SPEECH MOTOR CONTROL

How could we model sung speech in task dynamics?

- **Stiffness:** Gestures with lower stiffness take longer to reach their goals.
- **Targets:** Sung movements are often larger than speech movements, and it takes longer to move larger distances.

In either case, gesture parameters need to adapt to the broader singing or talking task requirements (see e.g., H&H Theory).

Task dynamics: Saltzman & Munhall 1989. H&H Theory: Lindblom 1983, 1990.

### FINDING TRUTH IN VARIETY

Varieties of speech have different planning and kinematics depending on their aesthetic and communicative goals.

But historically, linguists have tended to ignore musical speech and its relationship with non-musical speech.

To understand and model the cognitive units of speech more effectively, we need to expand our scope of inquiry to include musical and under-studied speech varieties.

Putting musical speech aside: e.g., Hockett 1955, Ladefoged 1989, cf Lindblom 1989.

### REFERENCES

Visit <https://reedblaylock.com> for full references.  
Bresch, E., & Narayanan, S. (2010). Real-time magnetic resonance imaging investigation of resonance tuning in soprano singing. *The Journal of the Acoustical Society of America*, 128(5), EL335-EL341.  
Hockett, C. F. (1955). *A manual of phonology* (Vol. 21). Indiana University Publications in Anthropology and Linguistics.  
Ladefoged, P. (1989). *Representing Phonetic Structure* (No. 73: Working Papers in Phonetics). Phonetics Laboratory, Department of Linguistics, UCLA.  
Lammert, A. C., Ramanarayanan, V., Proctor, M. I., & Narayanan, S. S. (2013, August). Vocal tract cross-distance estimation from real-time MRI using region-of-interest analysis. In *Interspeech* (pp. 959-962).  
Lindblom, B. (1983). Economy of speech gestures. In *The production of speech* (pp. 217-245). Springer, New York, NY.  
Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In *Speech production and speech modelling* (pp. 403-439). Springer, Dordrecht.  
Lindblom, B. (1990). On the notion of "possible speech sound." *Journal of Phonetics*, 18(2), 135-152.  
Ramanarayanan, V., Lammert, A., Byrd, D., Goldstein, L., Narayanan, S. (2011). "Planning and execution in soprano singing and speaking behavior: An acoustic/articulatory study using real-time MRI." *International Seminar on Speech Production (ISSP)*, Montreal, Canada.  
Saltzman, E. L., & Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. *Ecological psychology*, 1(4), 333-382.  
Tiede, M. (2010). *MVIEW: Multi-channel visualization application for displaying dynamic sensor movements*. development.

Supported by NIH. Thanks to Tasia Dedenbach for helping with poster design.

USC University of Southern California

