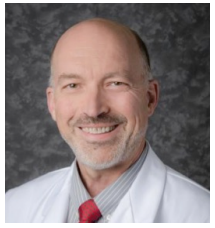
Talk presented at the Madonna Motor Speech Conference, Savanah, GA, February 2018

A hierarchical task-based control model of speech incorporating sensory feedback

Ben Parrell*
University of Wisconsin
Communication Sciences & Disorders

Vikram Ramanarayanan*
Educational Testing Service
University of California, San Francisco

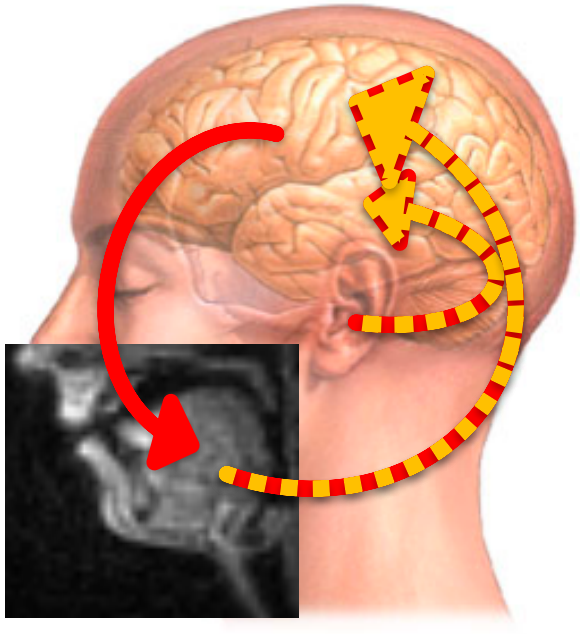
Sri Nagarajan, John Houde
University of California, San Francisco

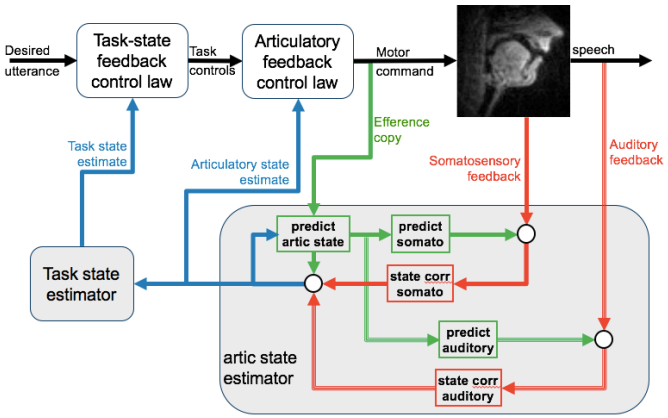


*equal contribution

The sensory system in speech production

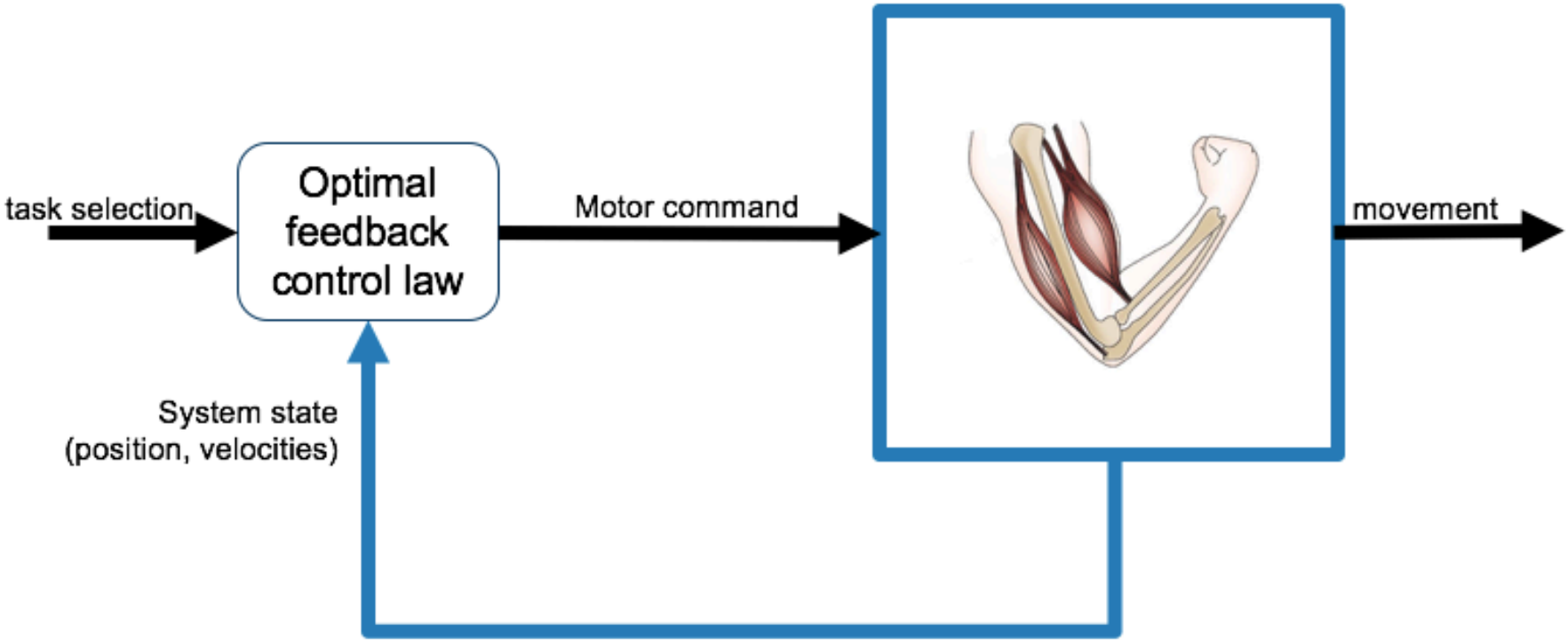
- Speech is responsive to sensory feedback
  - Delayed auditory feedback disrupts fluent speech
    (Lee, 1950; Yates, 1963)
  - Speakers compensate for all manner of auditory and somatosensory feedback perturbations
    (e.g. Houde 1998; Burnett 1998; Tremblay 2003;…)
- But speech is not dependent on sensory feedback
  - Post-lingually deafened adults can be intelligible for decades
    (Waldstein, 1990; Cowie and Douglas-Cowie, 1992)
  - Speech is highly intelligible during oral sensory and auditory deprivation (though articulatory precision is affected)
    (Scott and Ringel, 1971; Ringel and Steer, 1963)





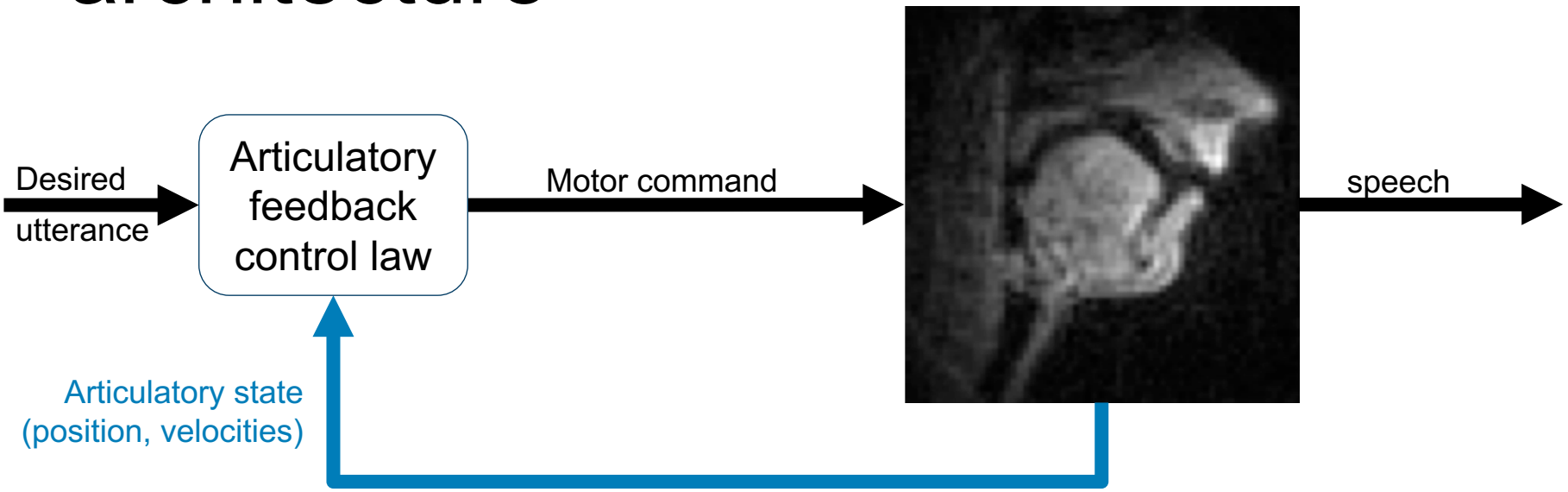
Feedback-Aware Control of Tasks in Speech (FACTS)

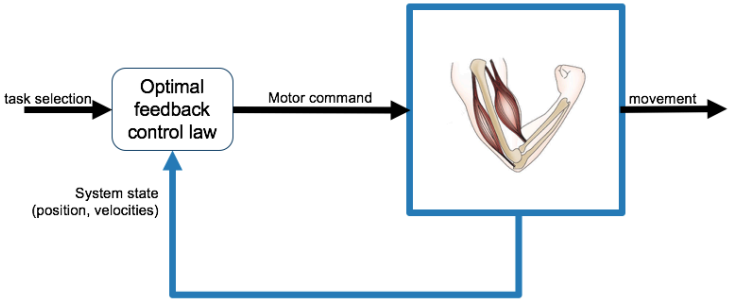
FACTS is based on a domain-general State Feedback Control architecture



Scott, S. H. (2004): *Nature Reviews Neuroscience, 5*(7), 534-546.

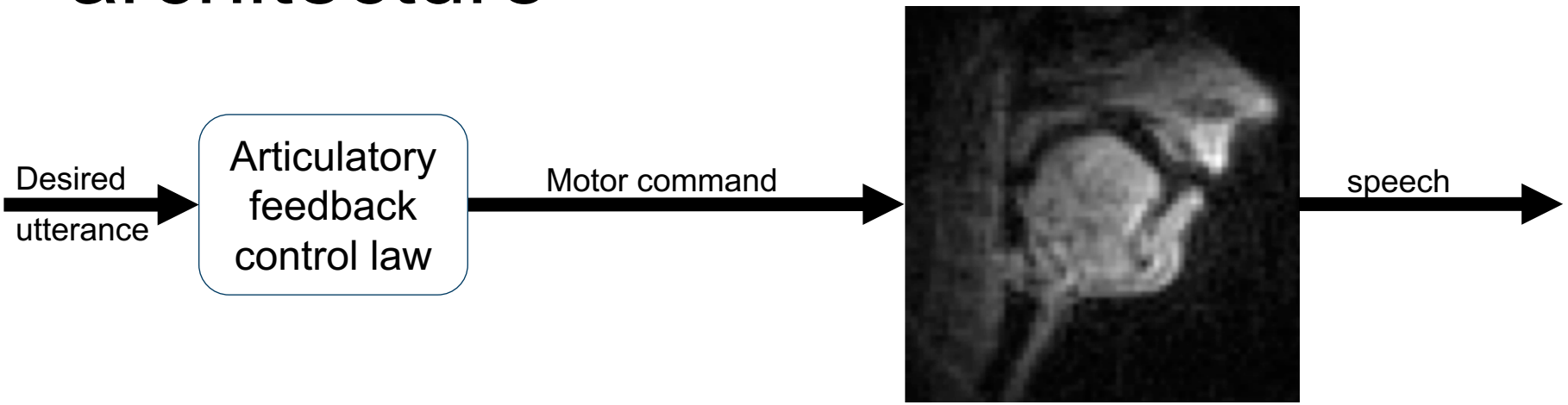
FACTS is based on a domain-general State Feedback Control architecture



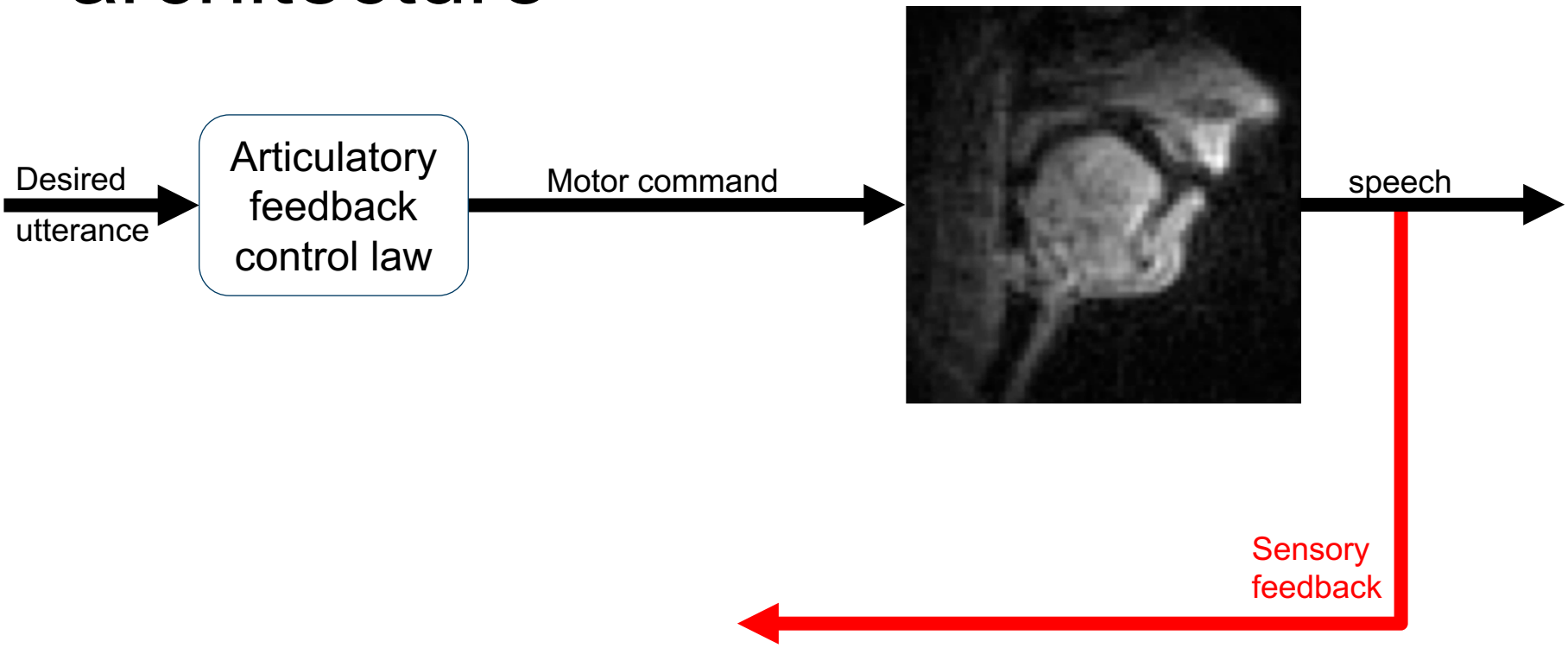


Scott, S. H. (2004): *Nature Reviews Neuroscience, 5*(7), 534-546.

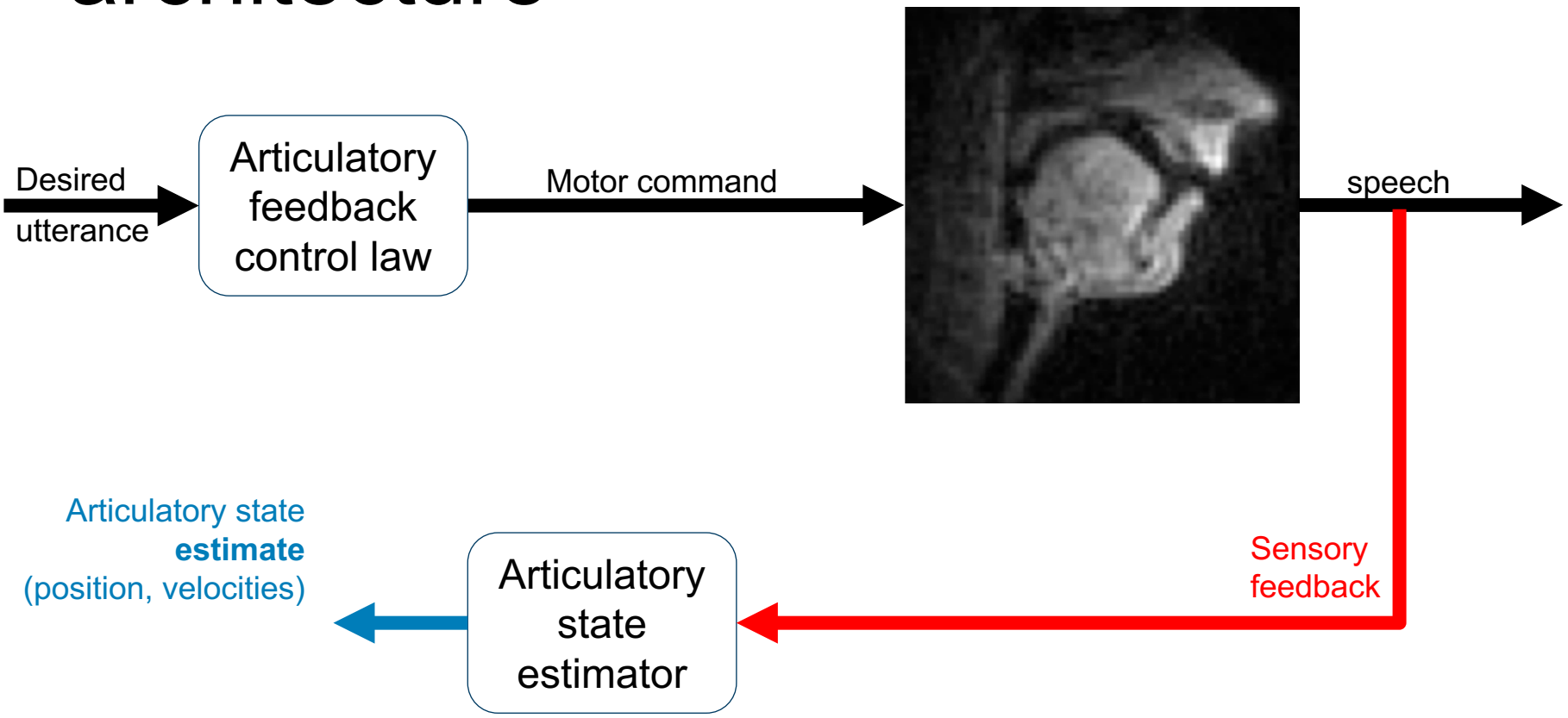
FACTS is based on a domain-general State Feedback Control architecture



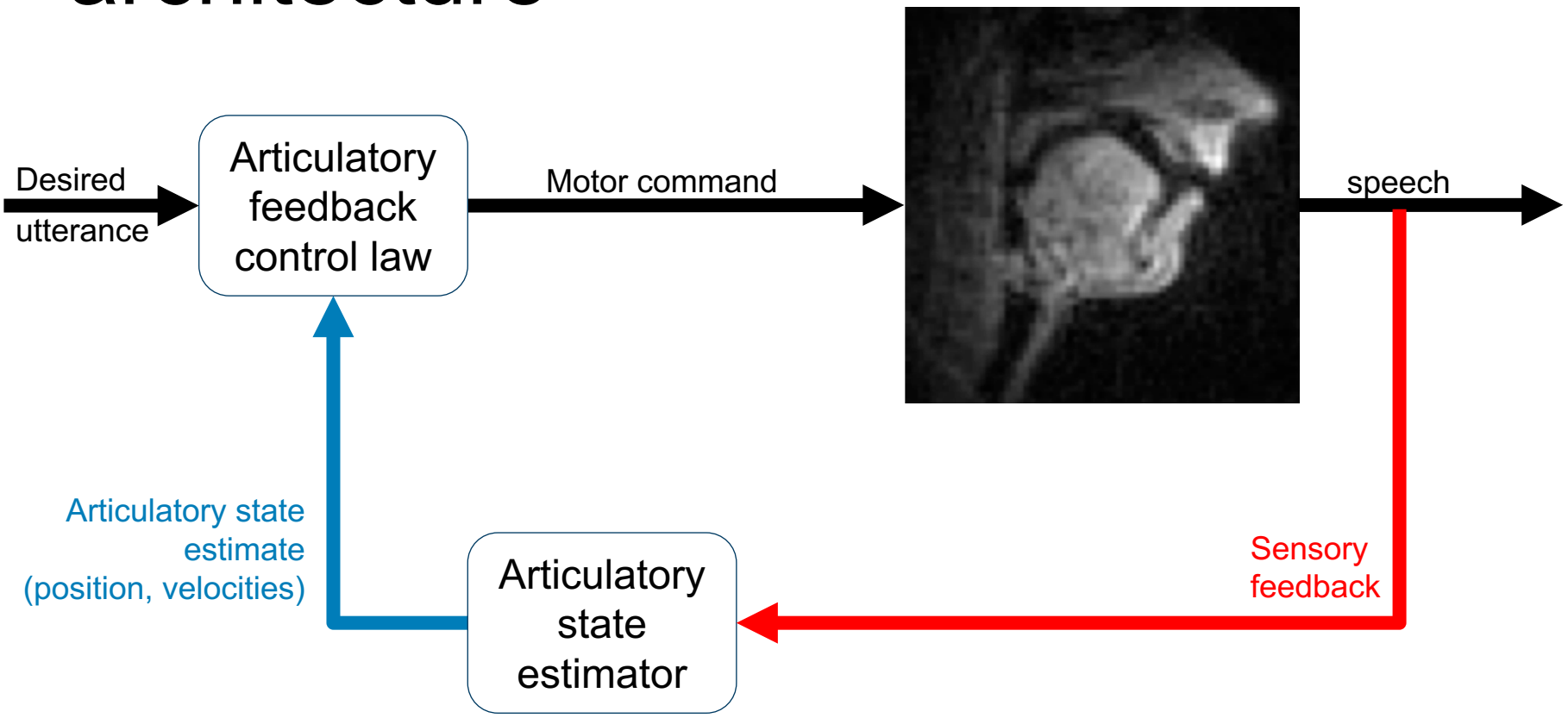
FACTS is based on a domain-general State Feedback Control architecture



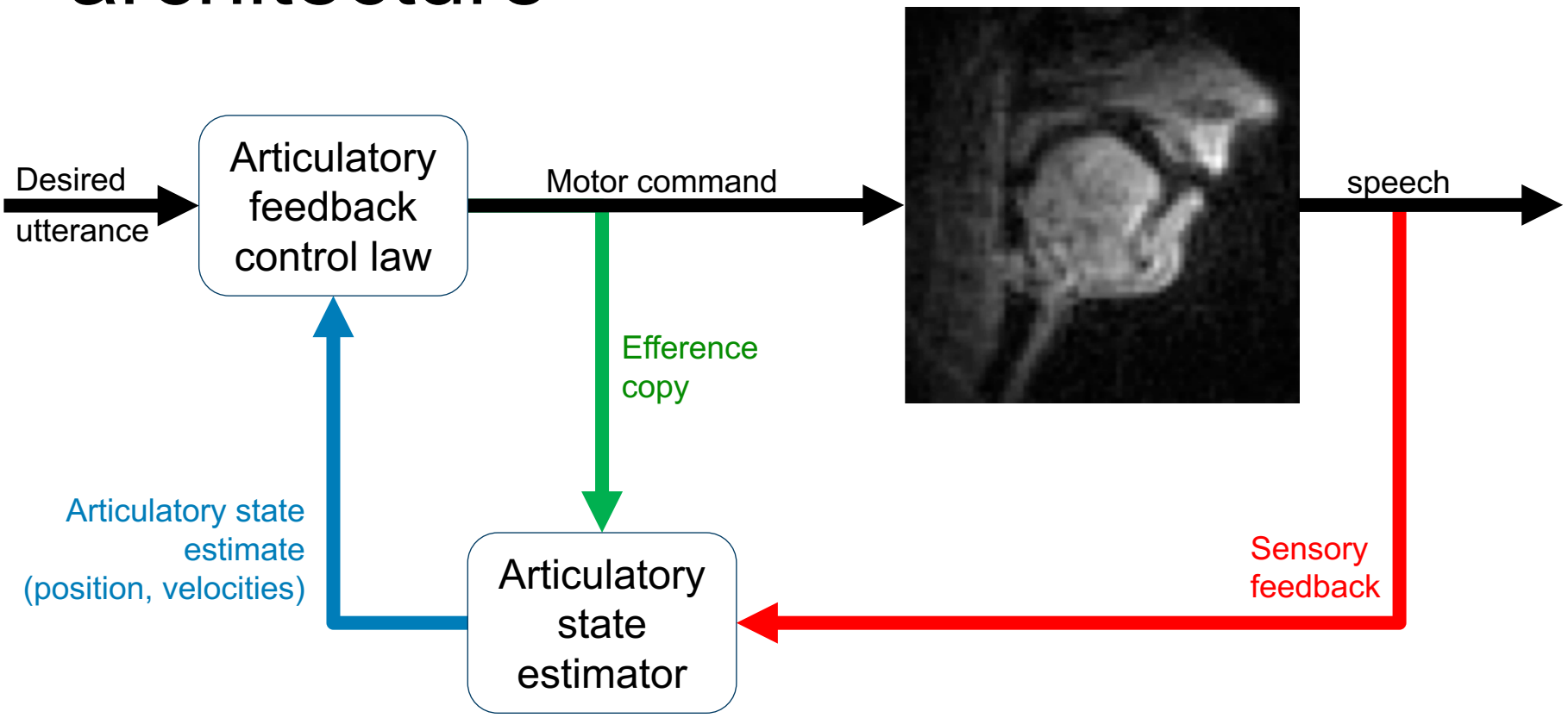
FACTS is based on a domain-general State Feedback Control architecture



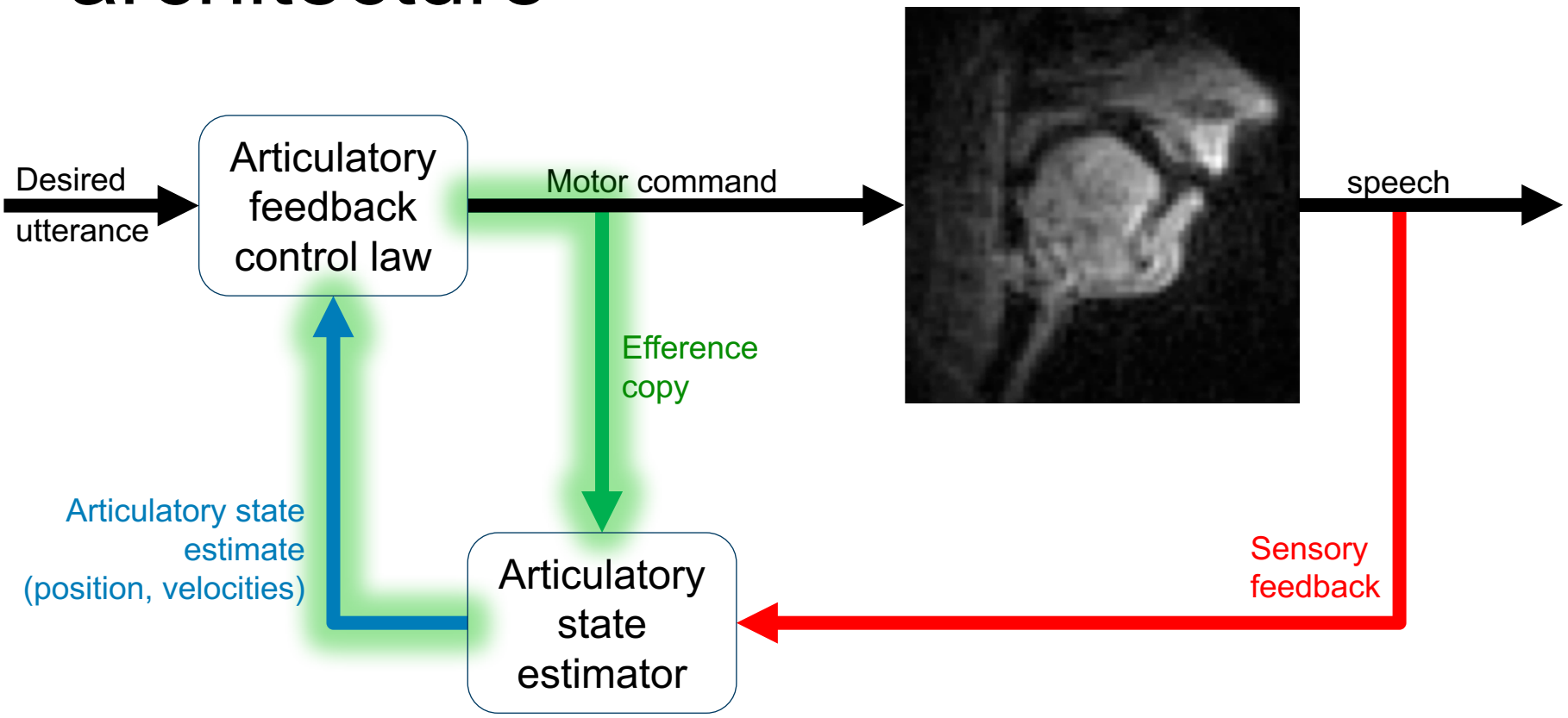
FACTS is based on a domain-general State Feedback Control architecture



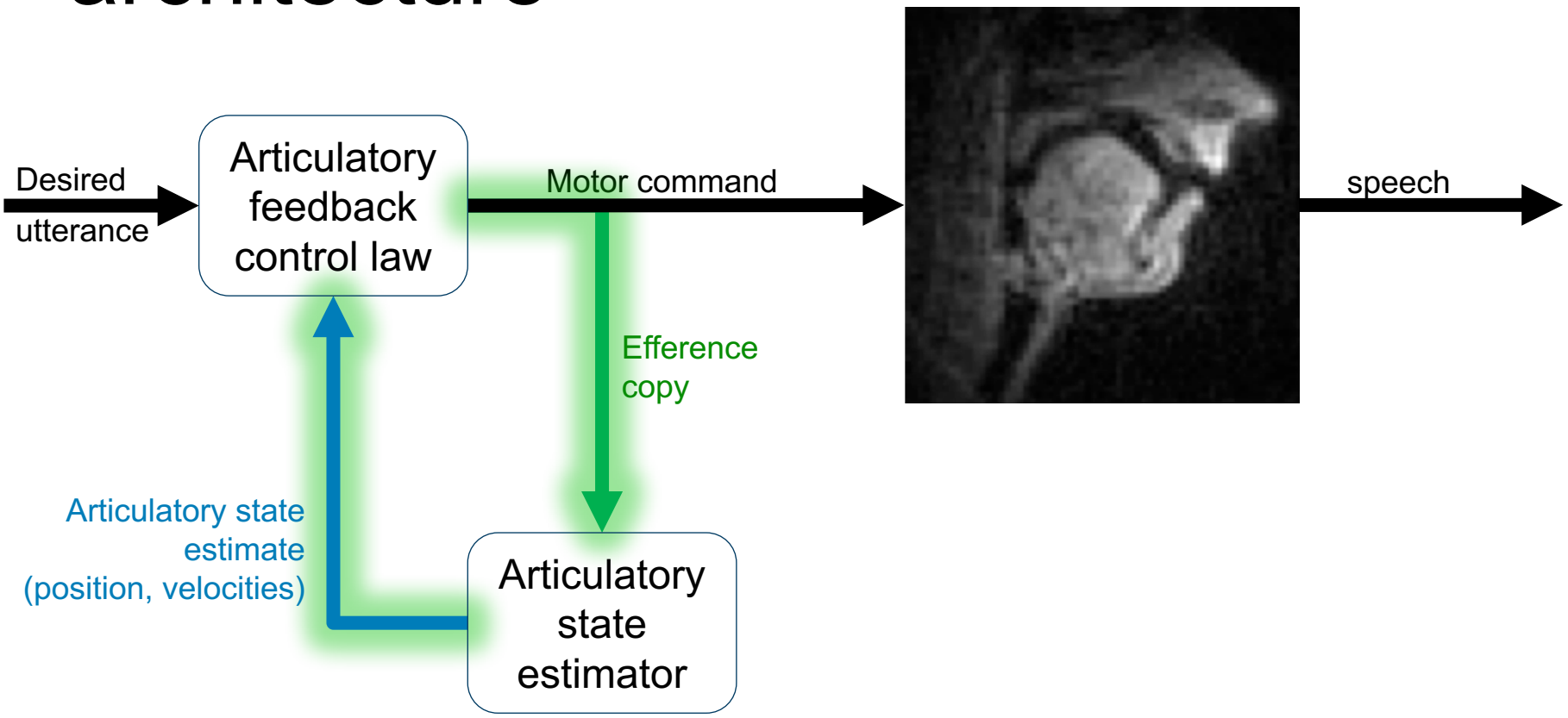
FACTS is based on a domain-general State Feedback Control architecture



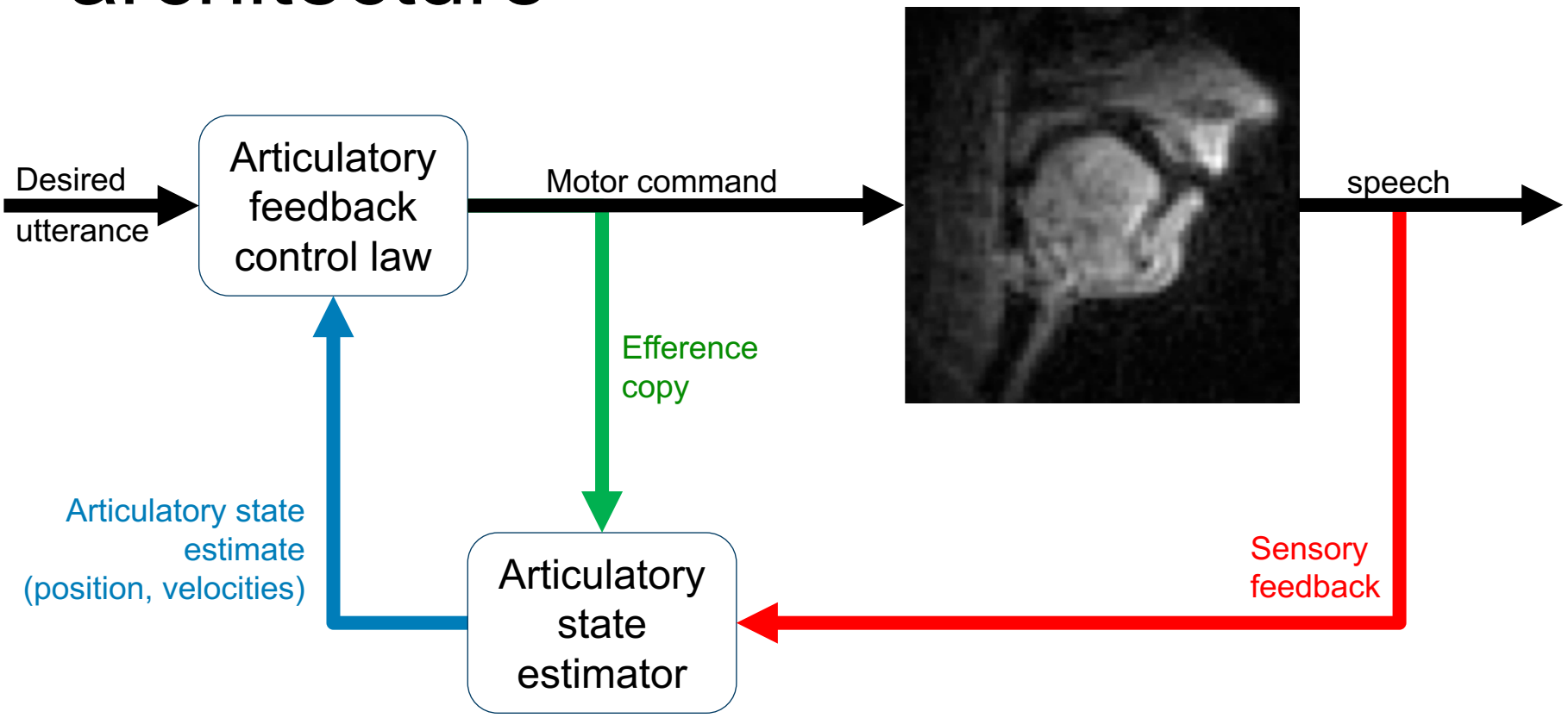
FACTS is based on a domain-general State Feedback Control architecture



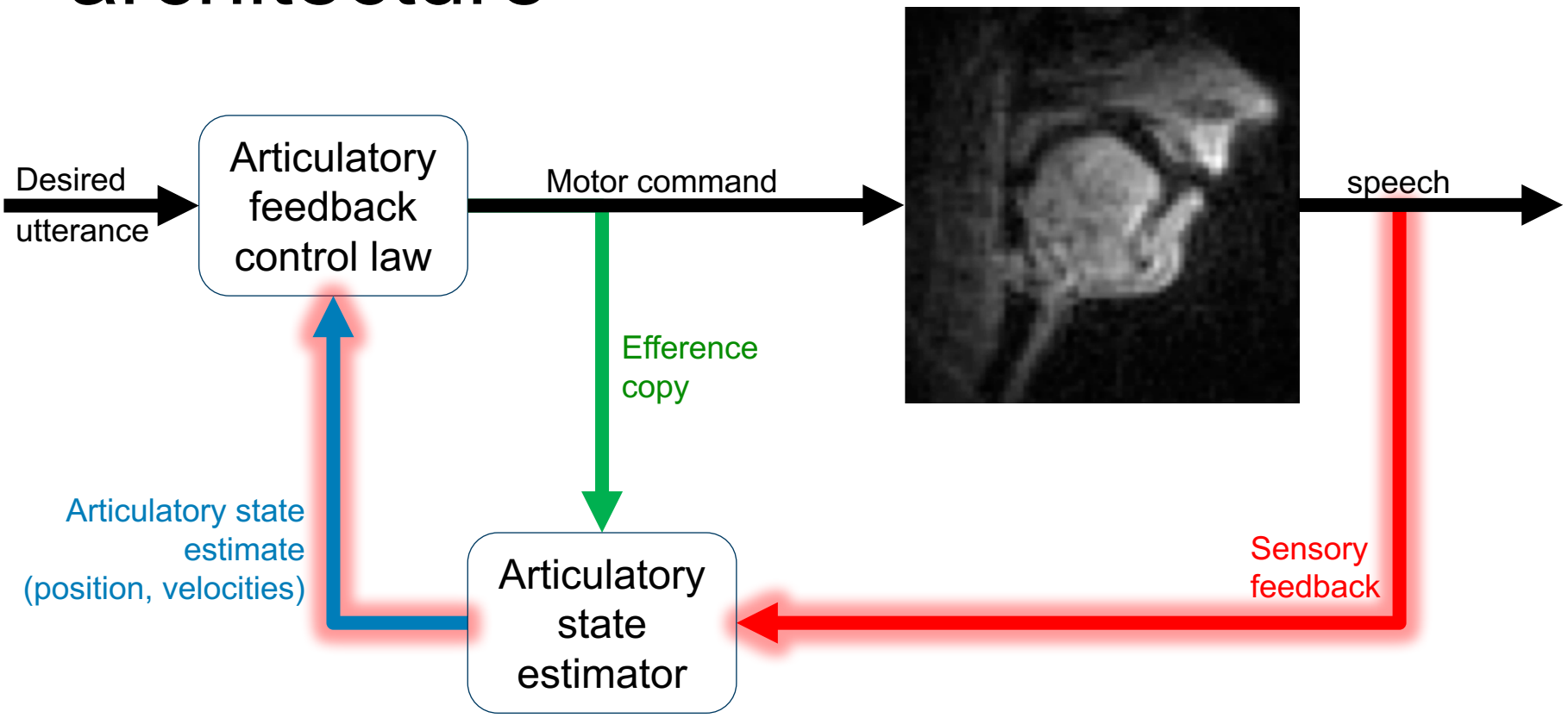
FACTS is based on a domain-general State Feedback Control architecture



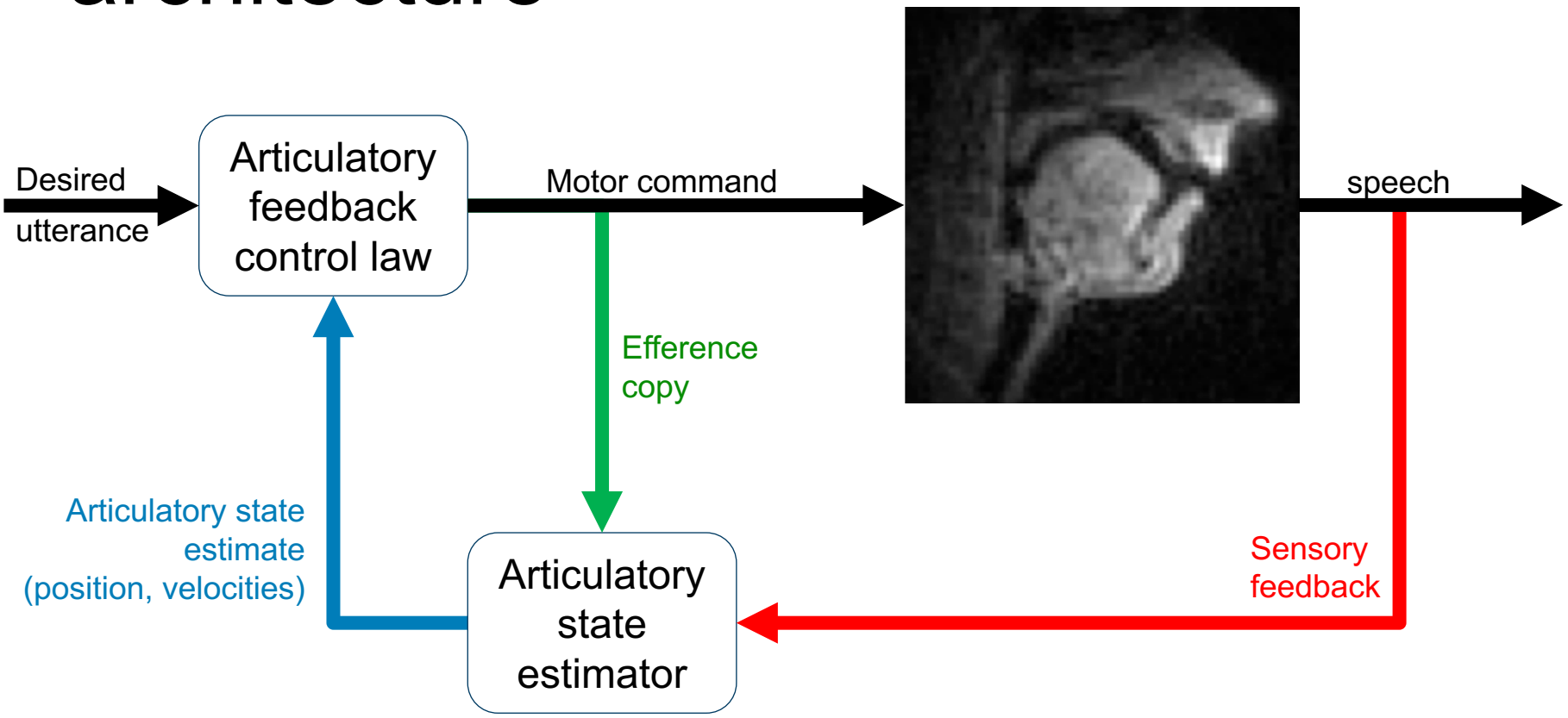
FACTS is based on a domain-general State Feedback Control architecture



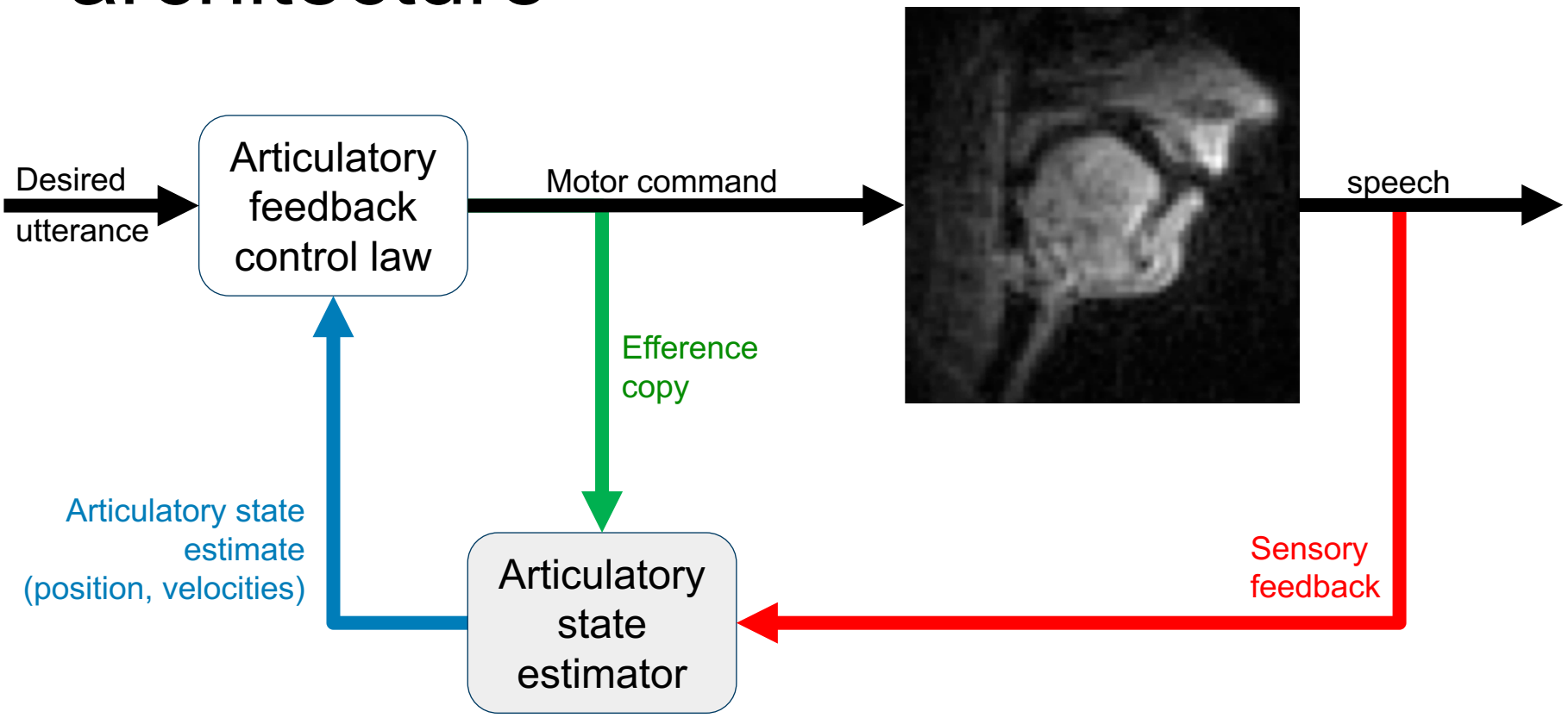
FACTS is based on a domain-general State Feedback Control architecture



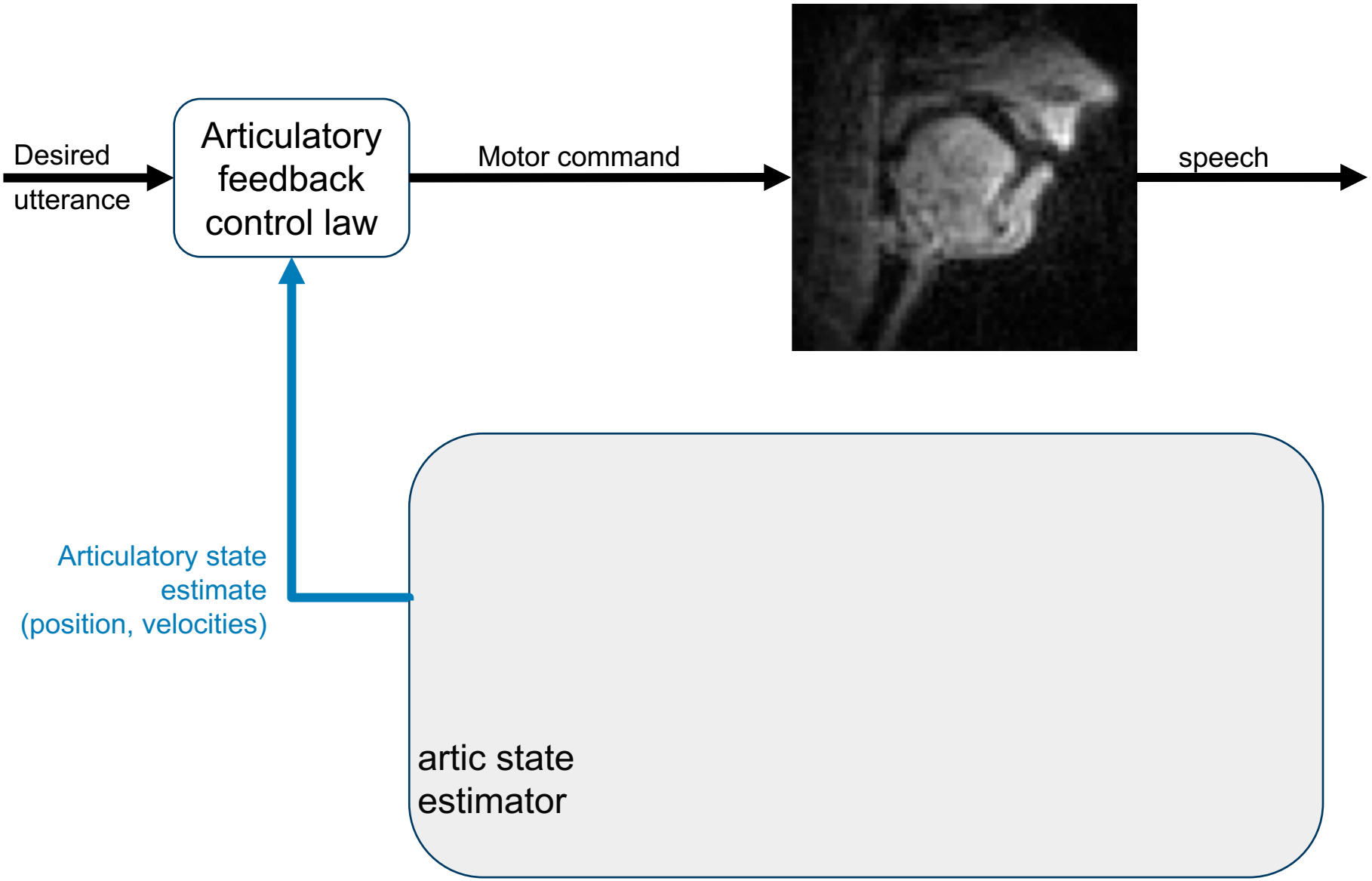
FACTS is based on a domain-general State Feedback Control architecture



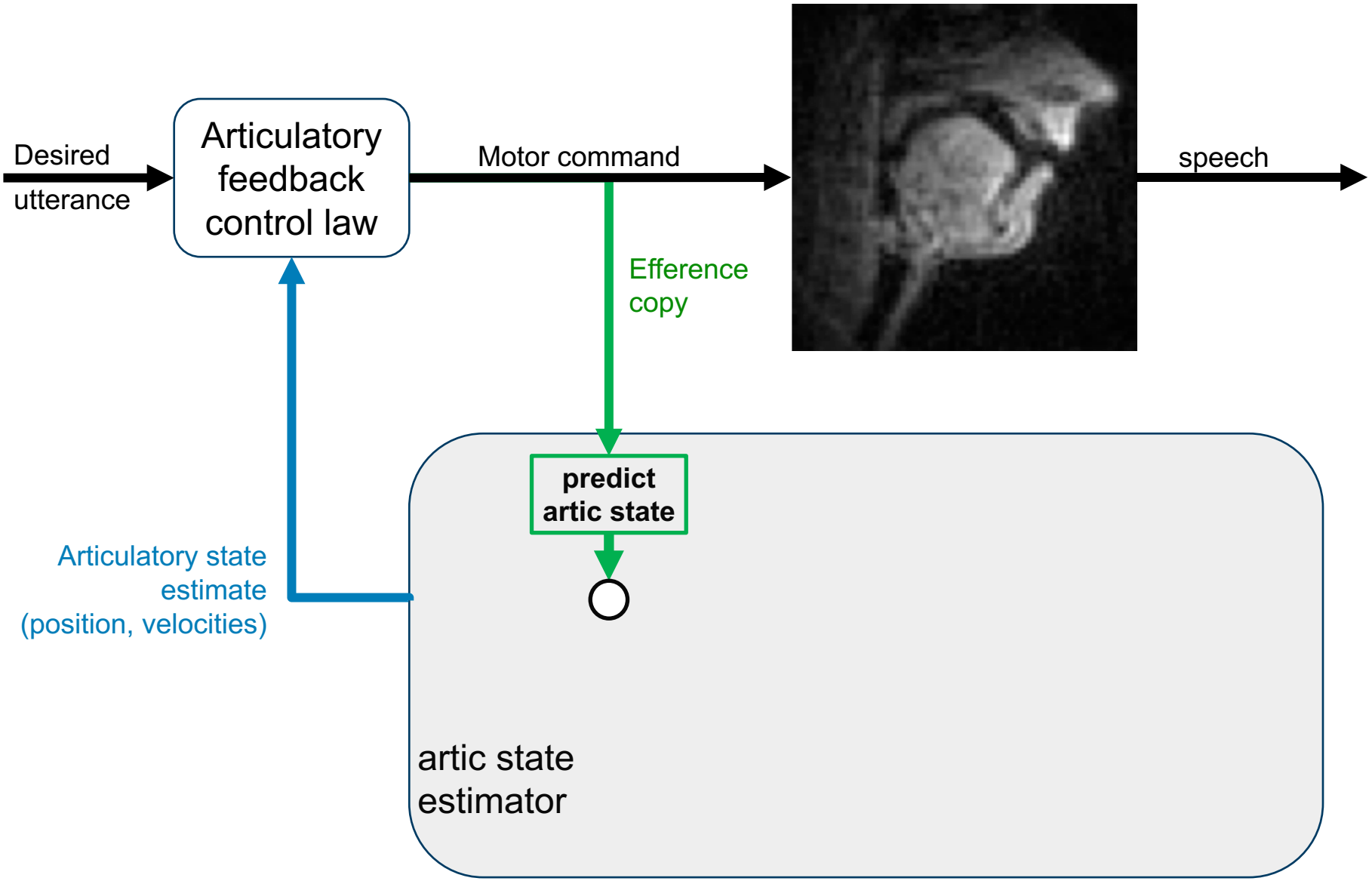
FACTS is based on a domain-general State Feedback Control architecture



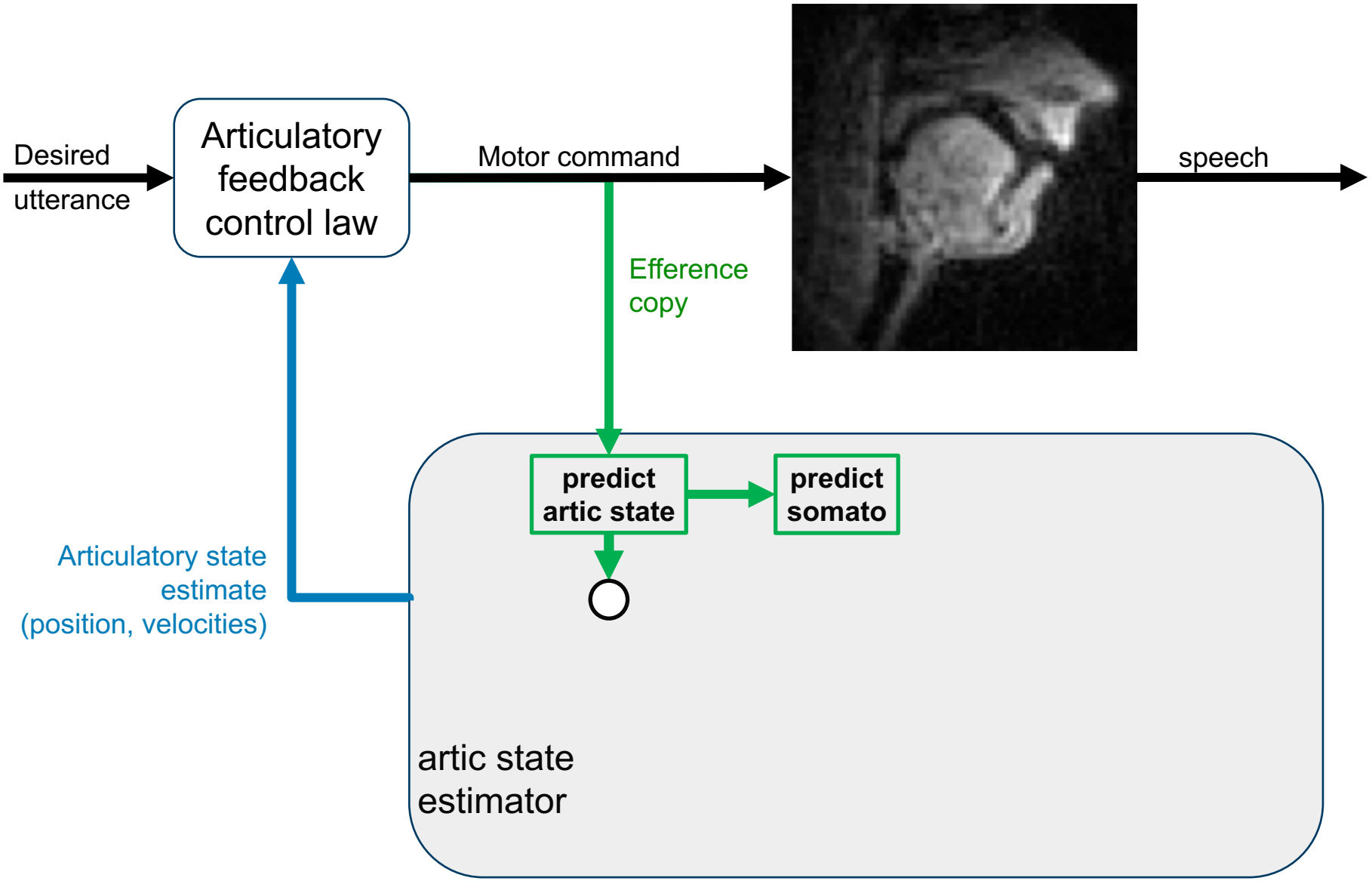
Estimation of the articulatory state



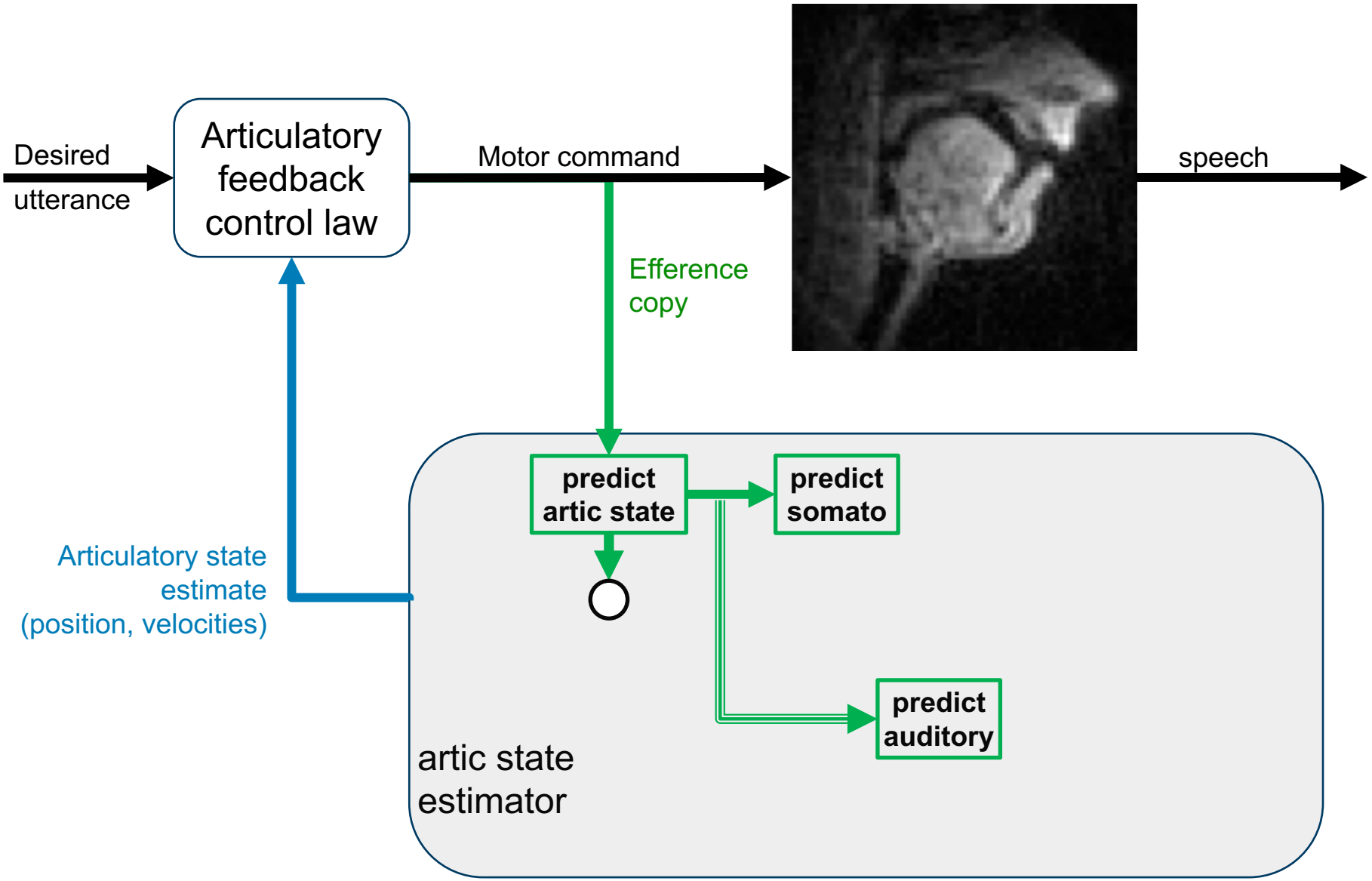
Estimation of the articulatory state



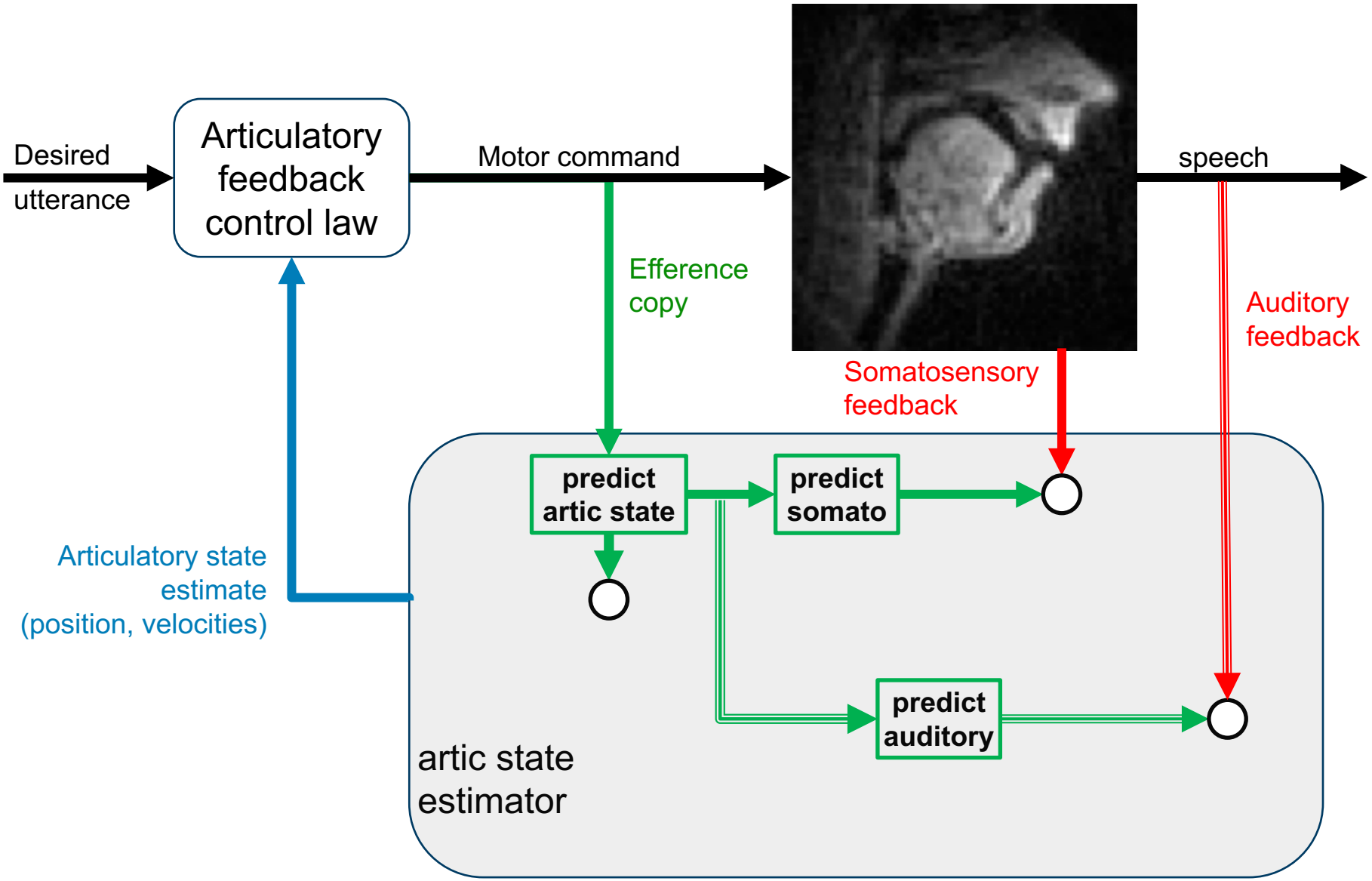
Estimation of the articulatory state



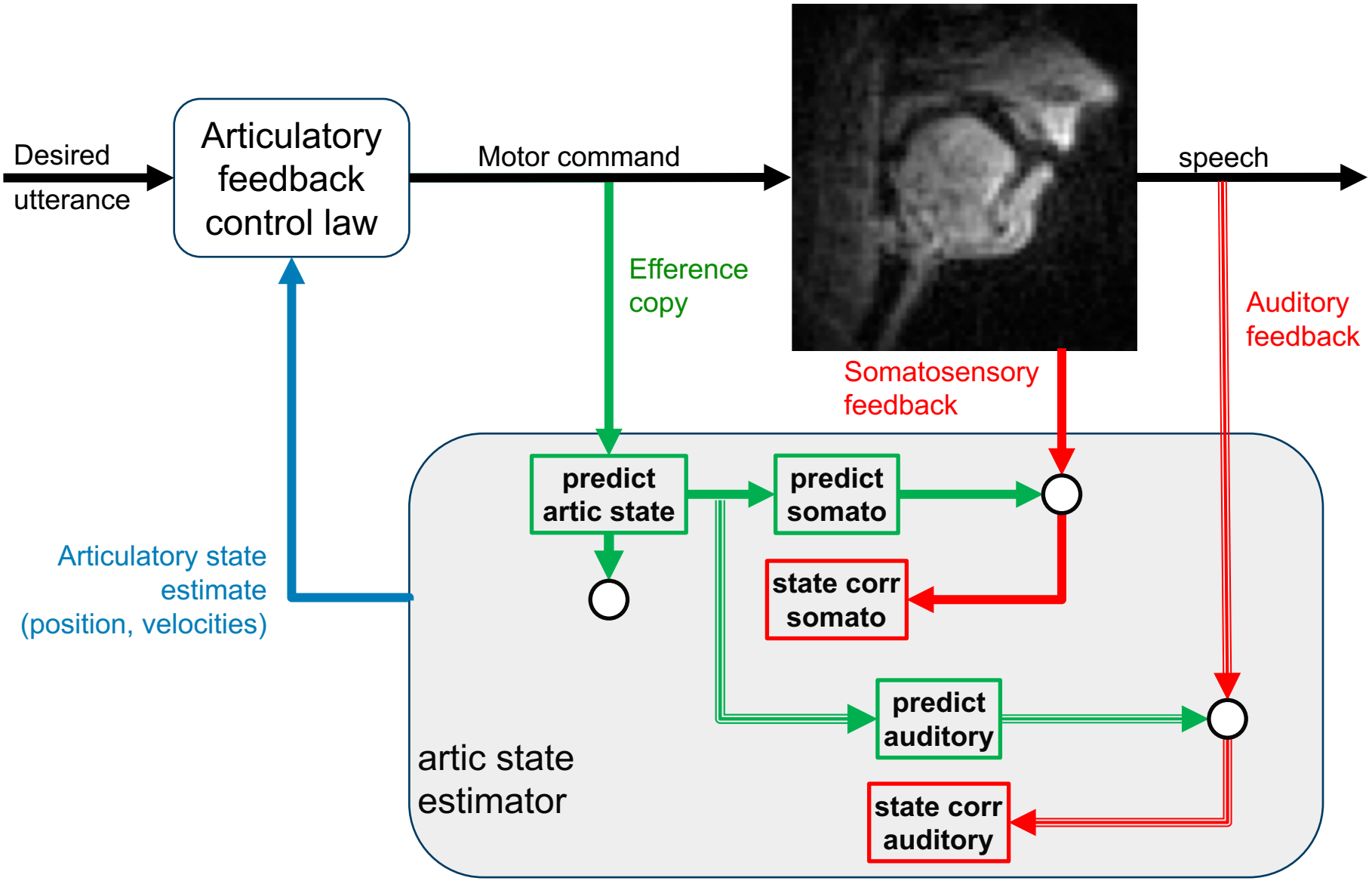
Estimation of the articulatory state



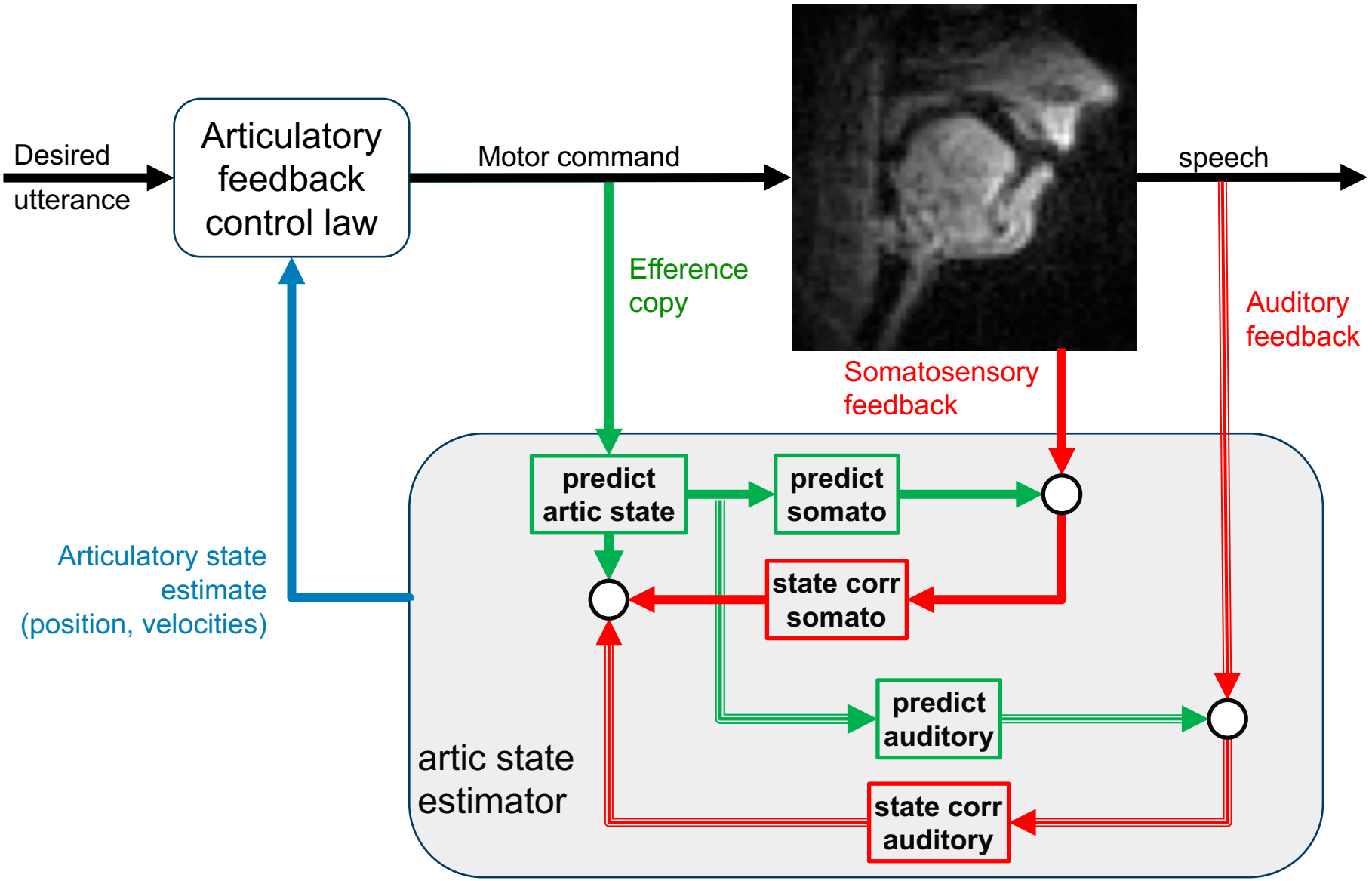
Estimation of the articulatory state



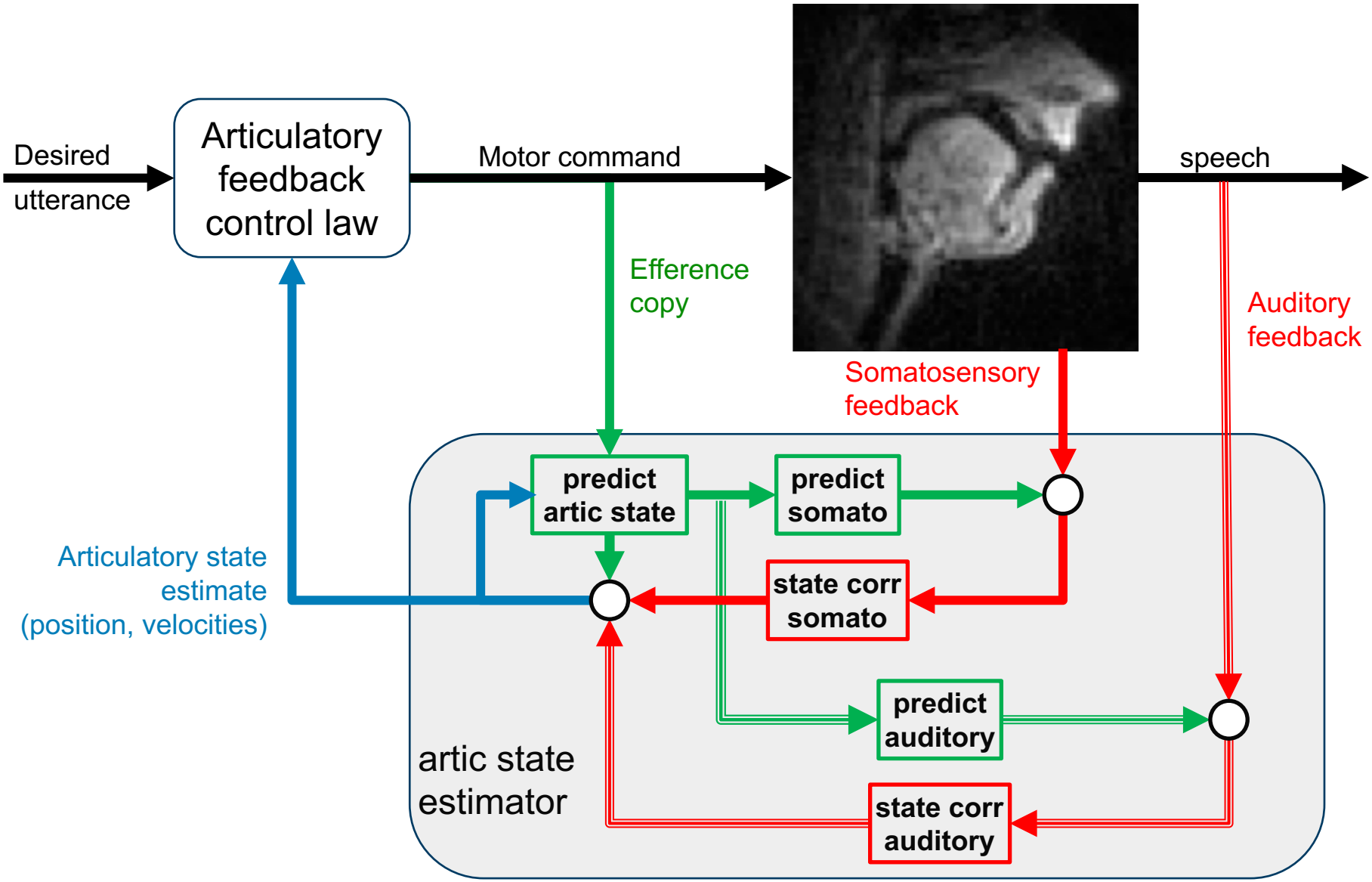
Estimation of the articulatory state



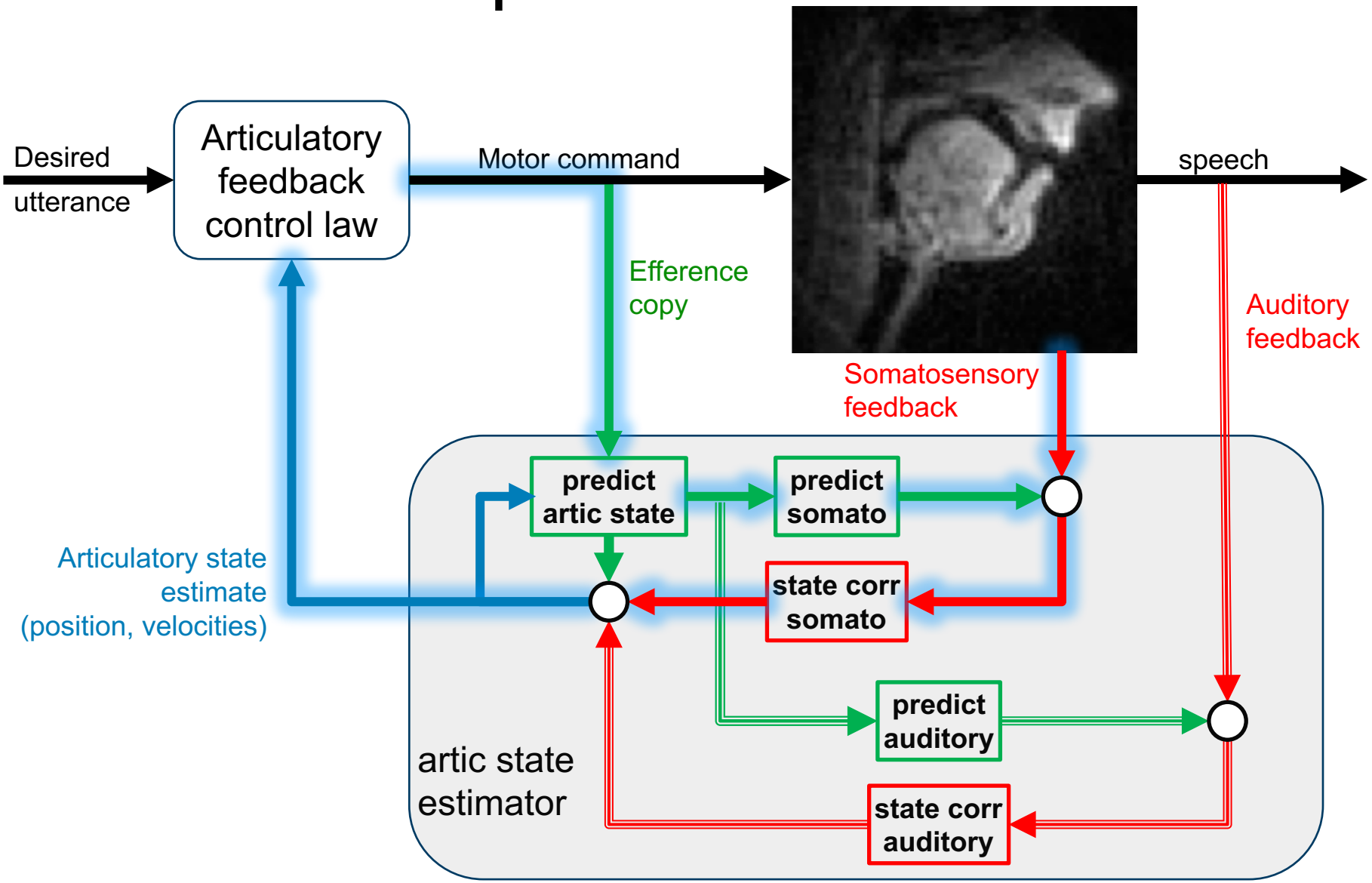
Estimation of the articulatory state



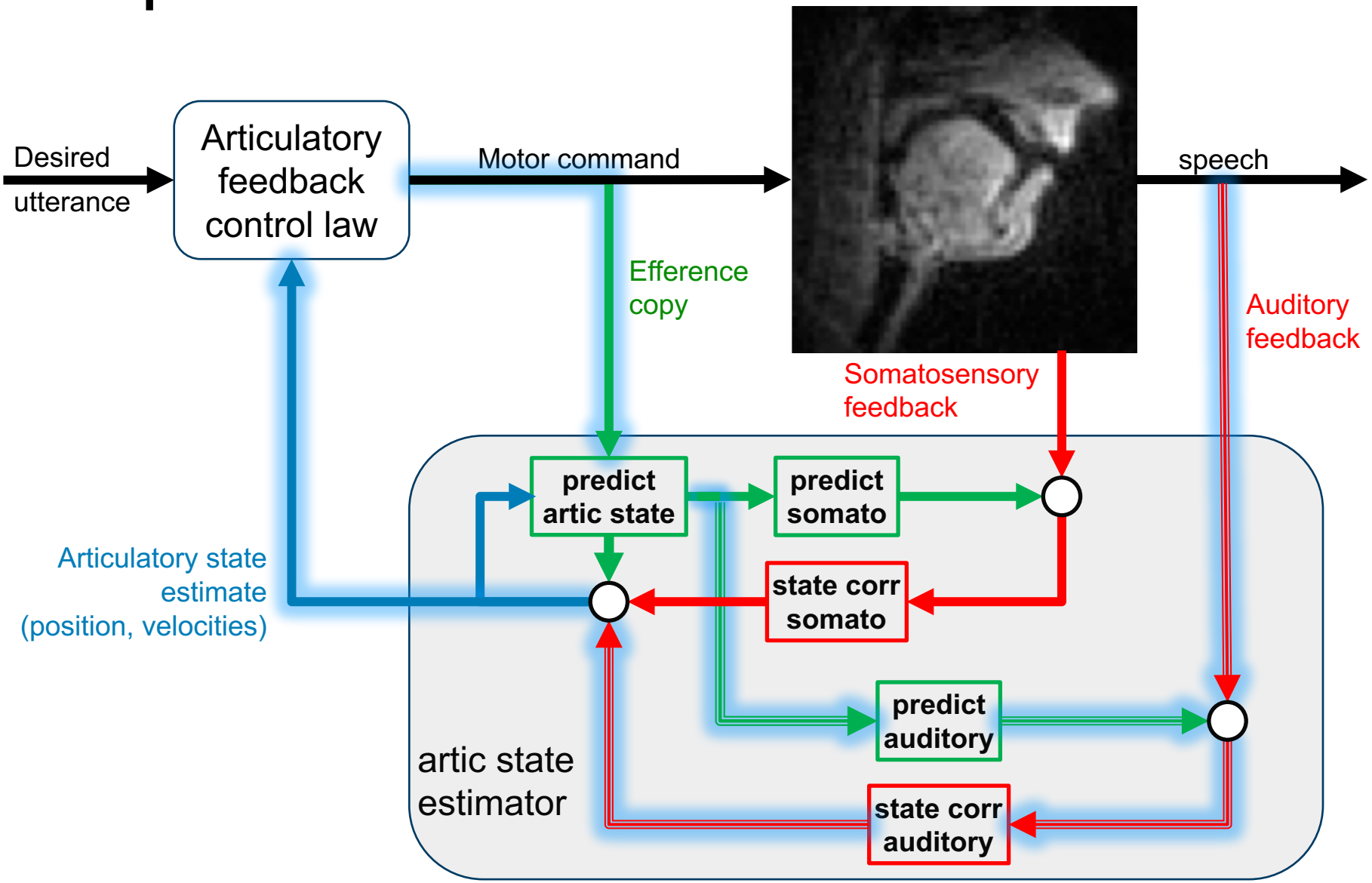
Estimation of the articulatory state



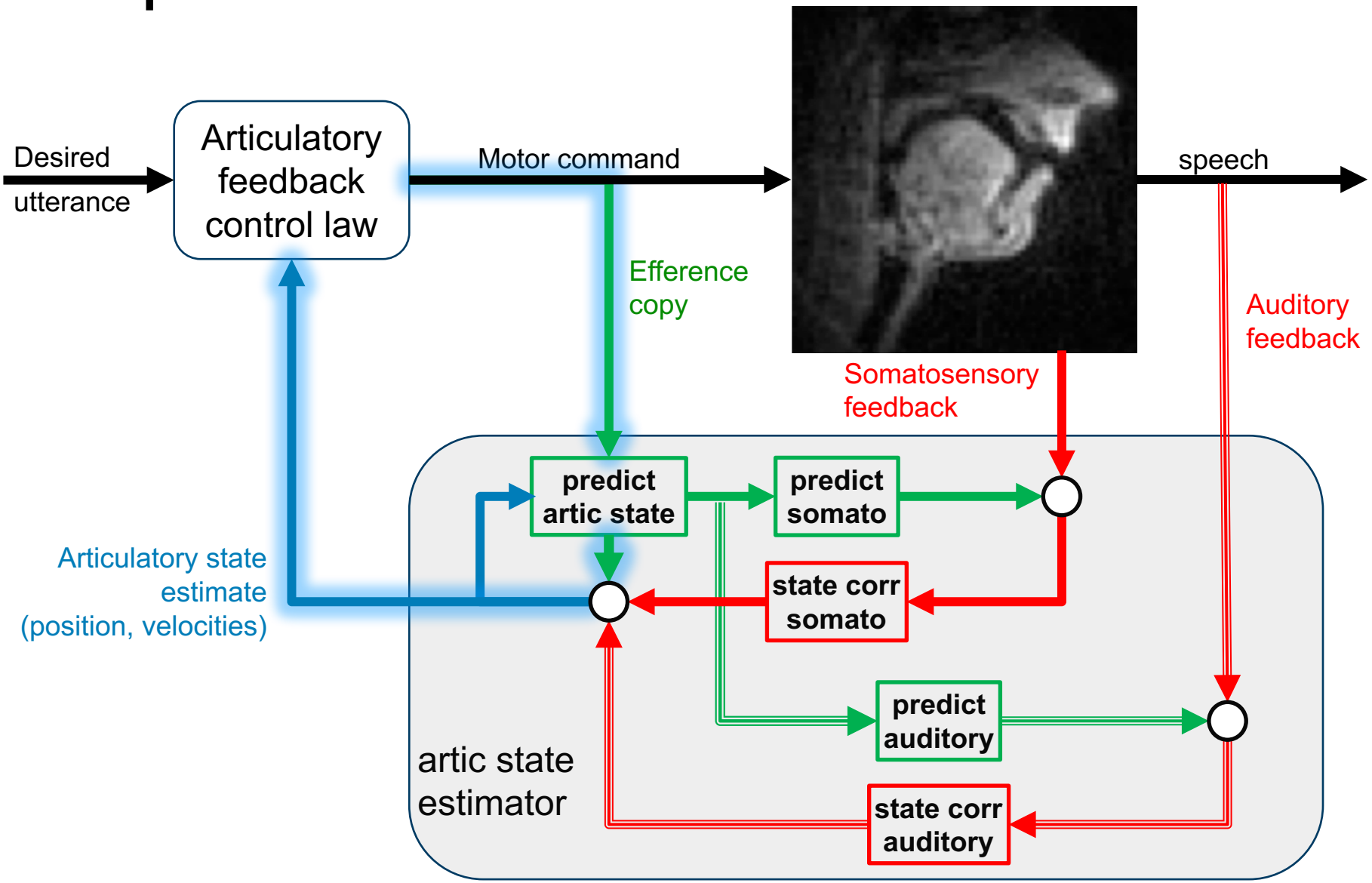
FACTS has a somatosensory feedback loop



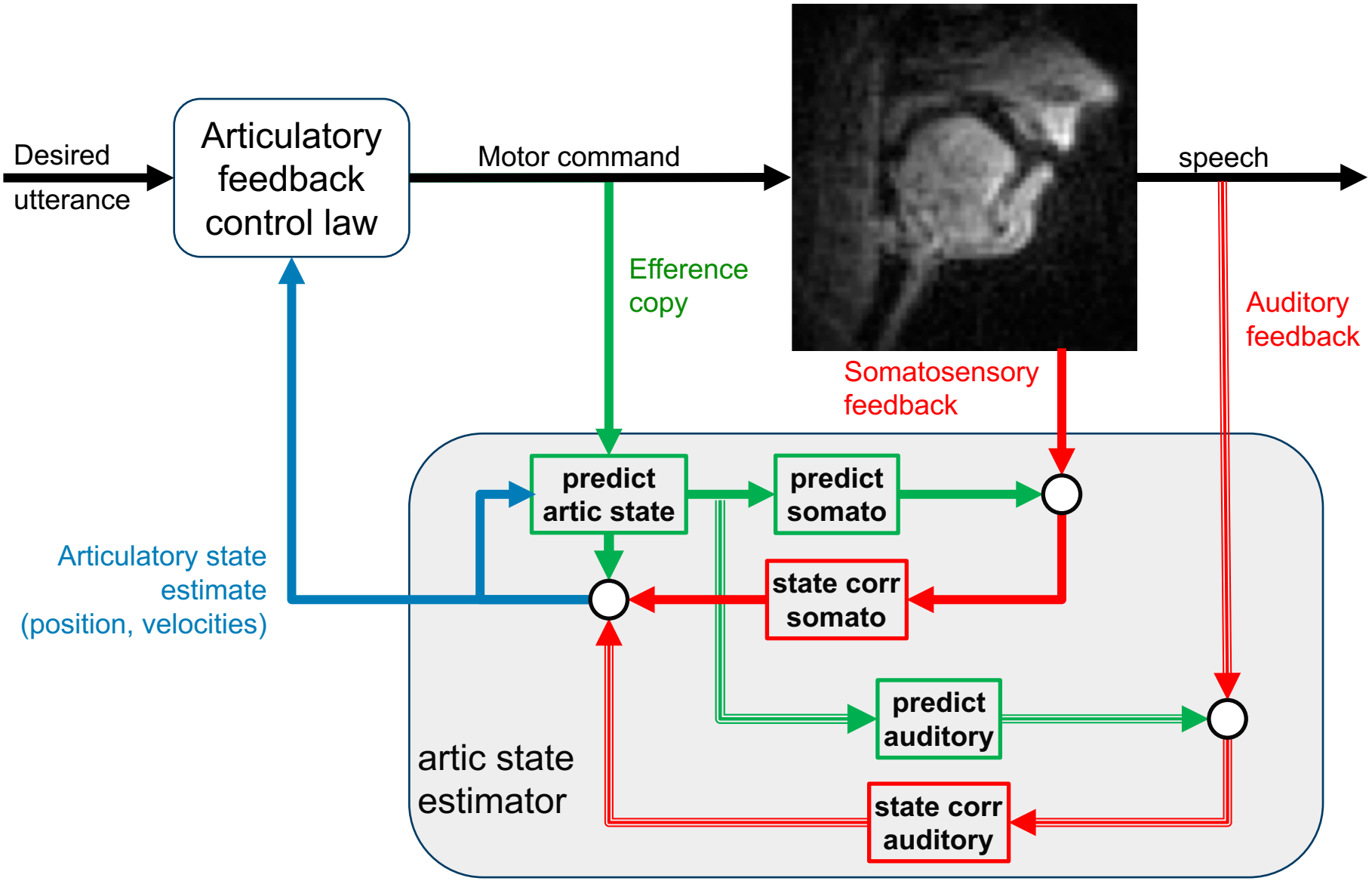
FACTS has an auditory feedback loop



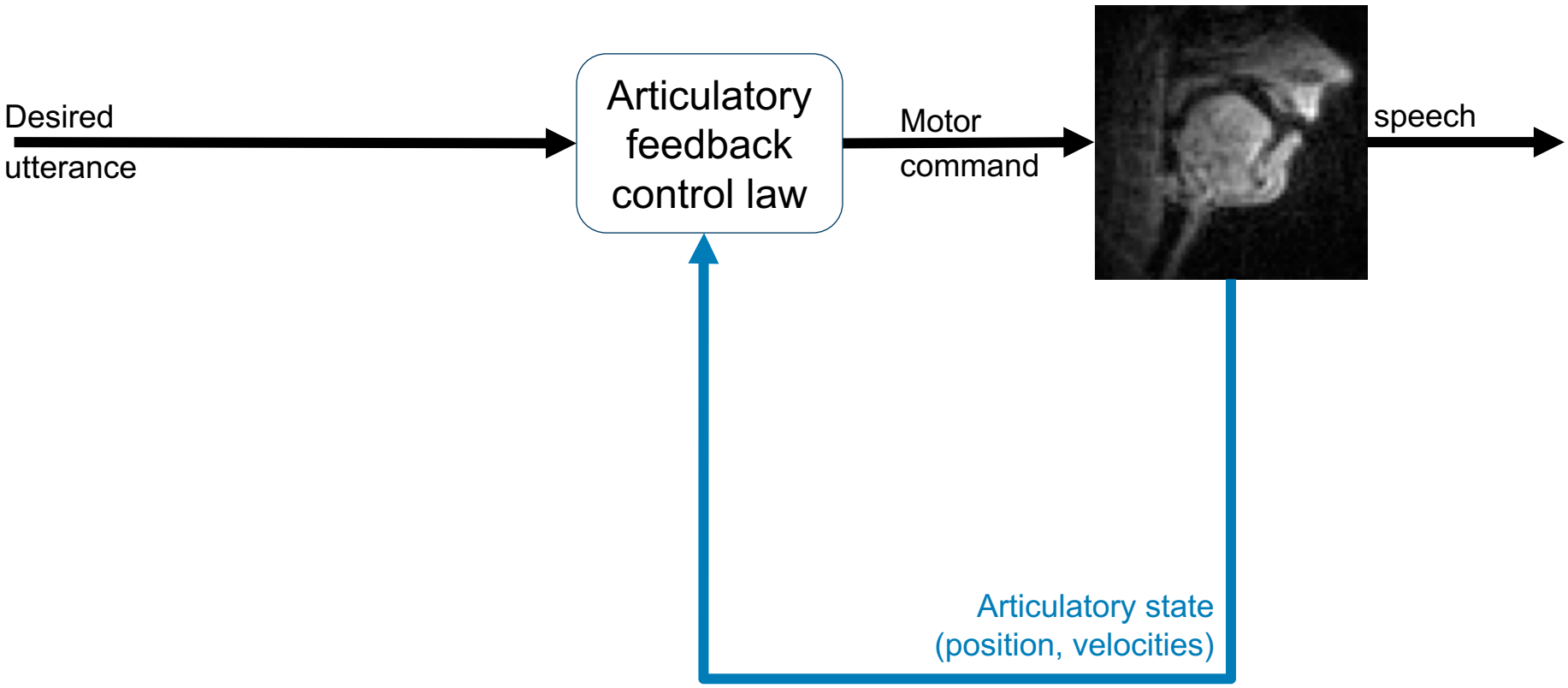
FACTS has a predictive control loop



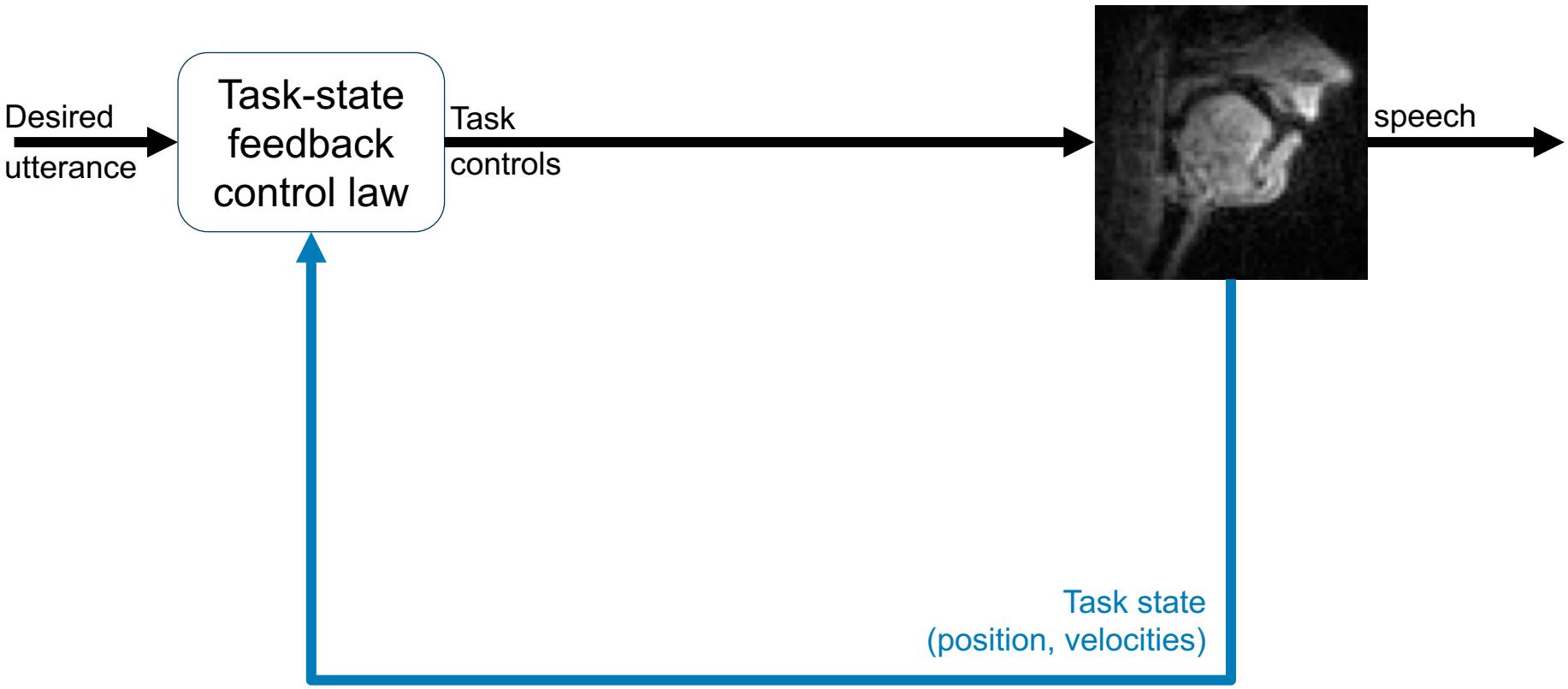
Modeling control of speech



Moving to a task-level controller



Moving to a task-level controller



Task Dynamics: a task-state feedback control system

Desired utterance

Task-state feedback control law

"bad"

Lip Aperture — CLO

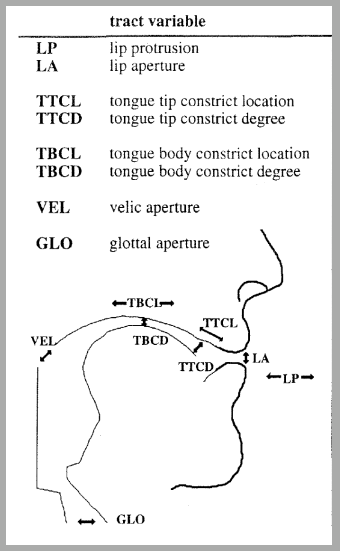
Tongue Tip Constriction Degree — CLO

Tongue Body Constriction Degree — WIDE

(pos

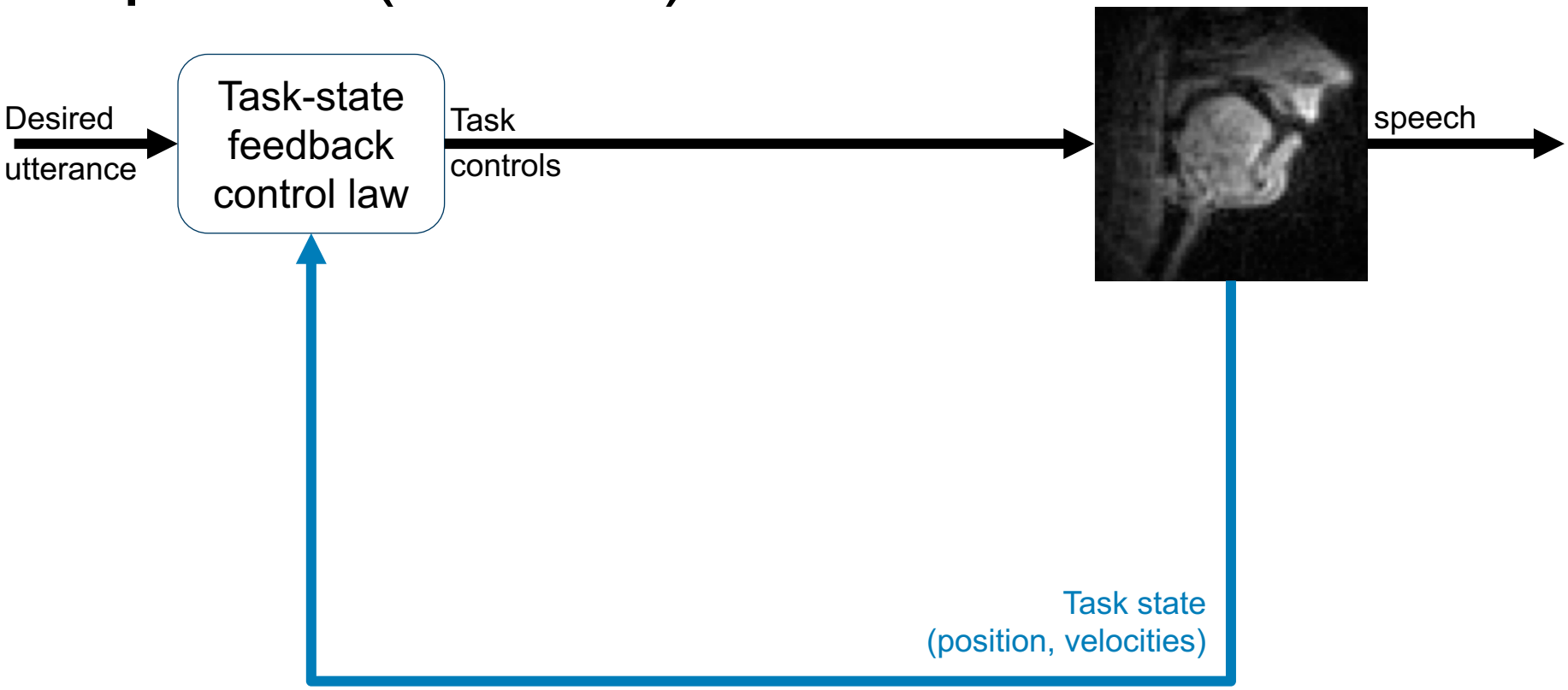
speech

Tasks: constrictions in the vocal tract

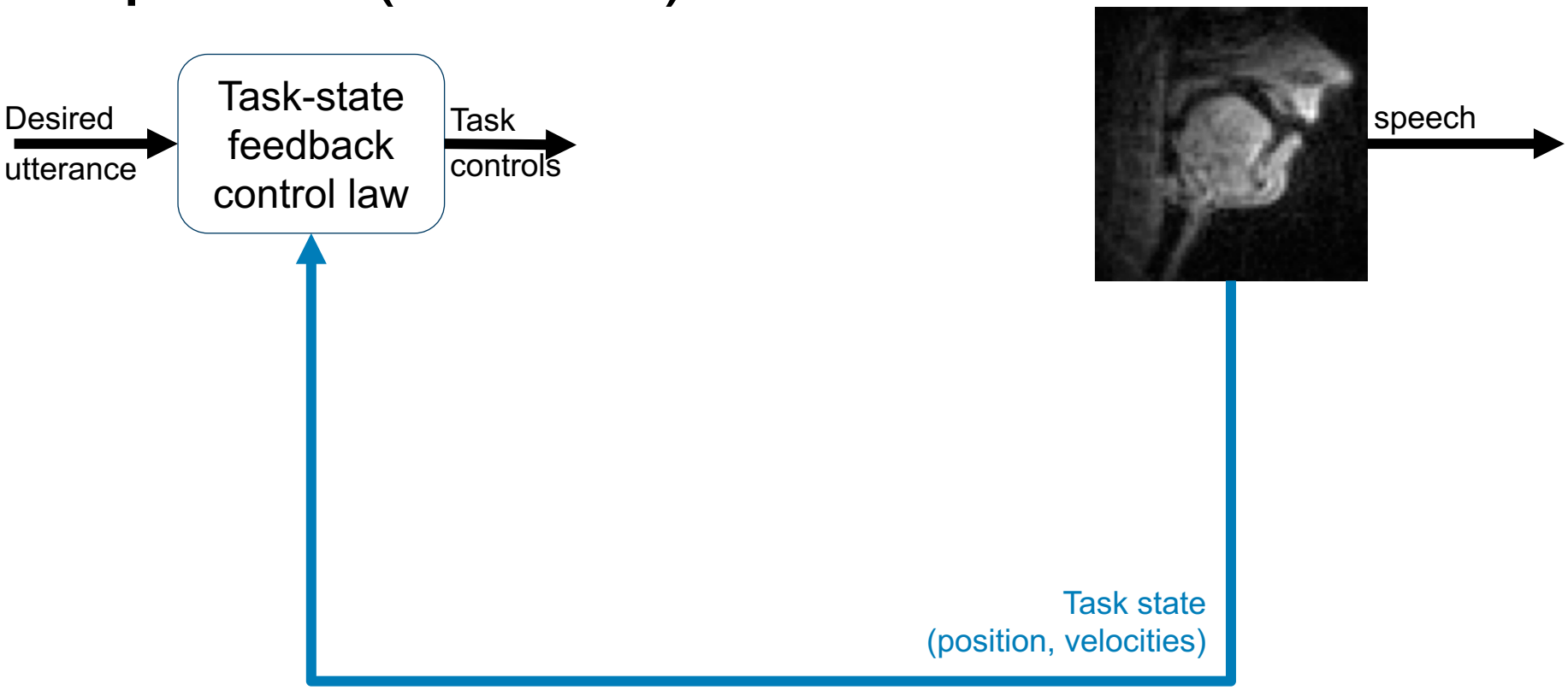


Saltzman & Munhall. (1989): *Ecological Psychology, 1*(4), 333-382.

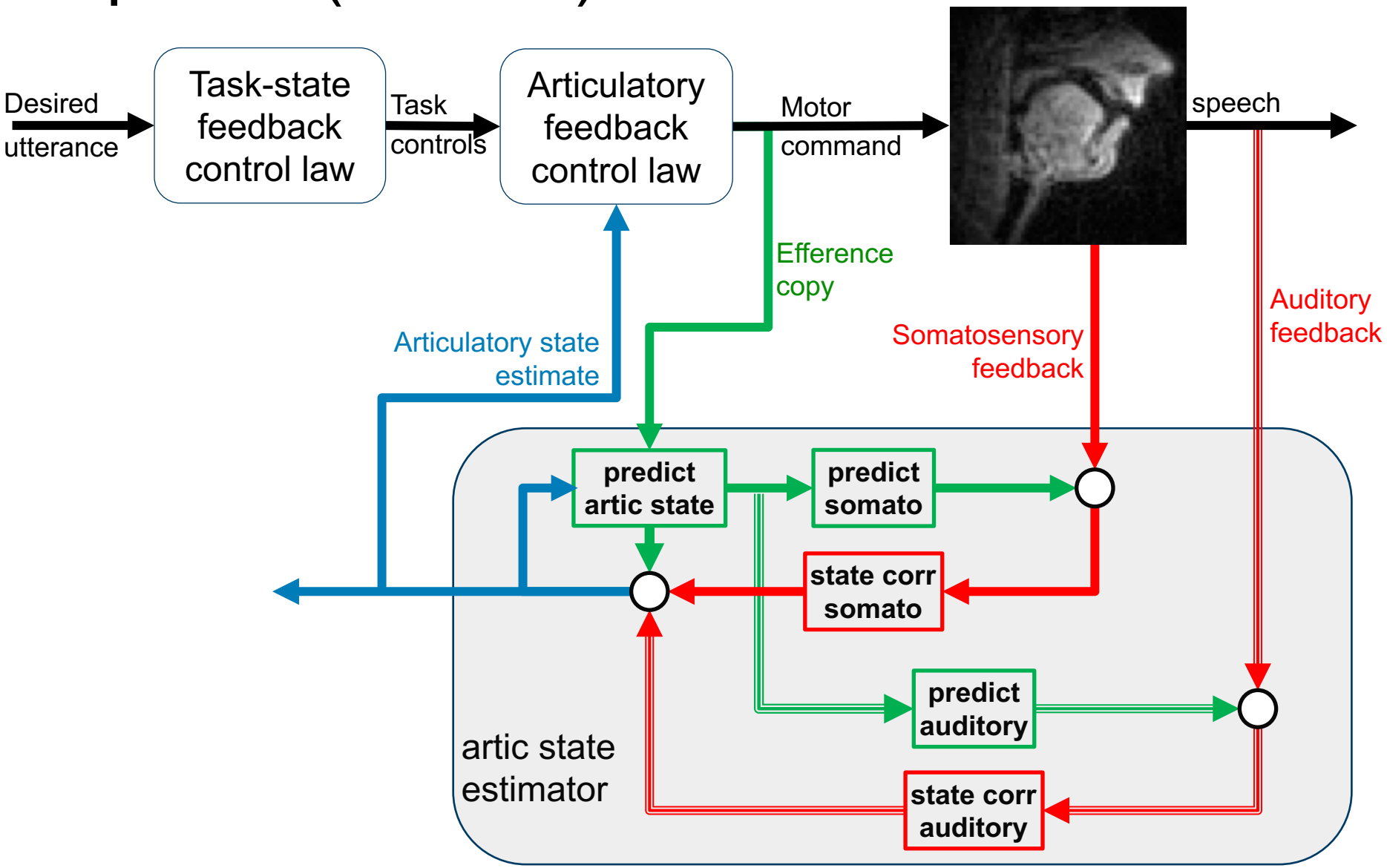
Feedback-Aware Control of Tasks in Speech (FACTS) model



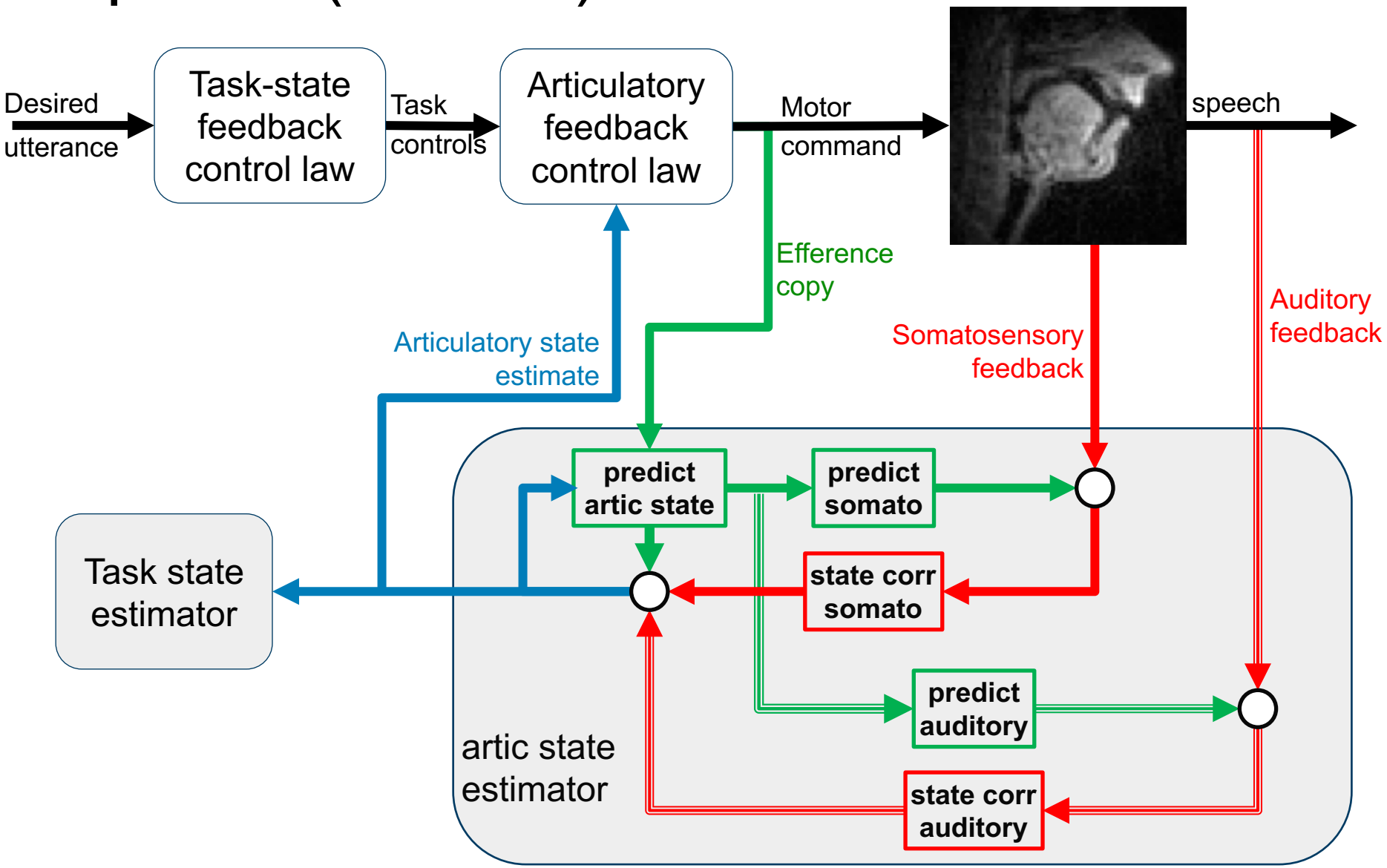
Feedback-Aware Control of Tasks in Speech (FACTS) model



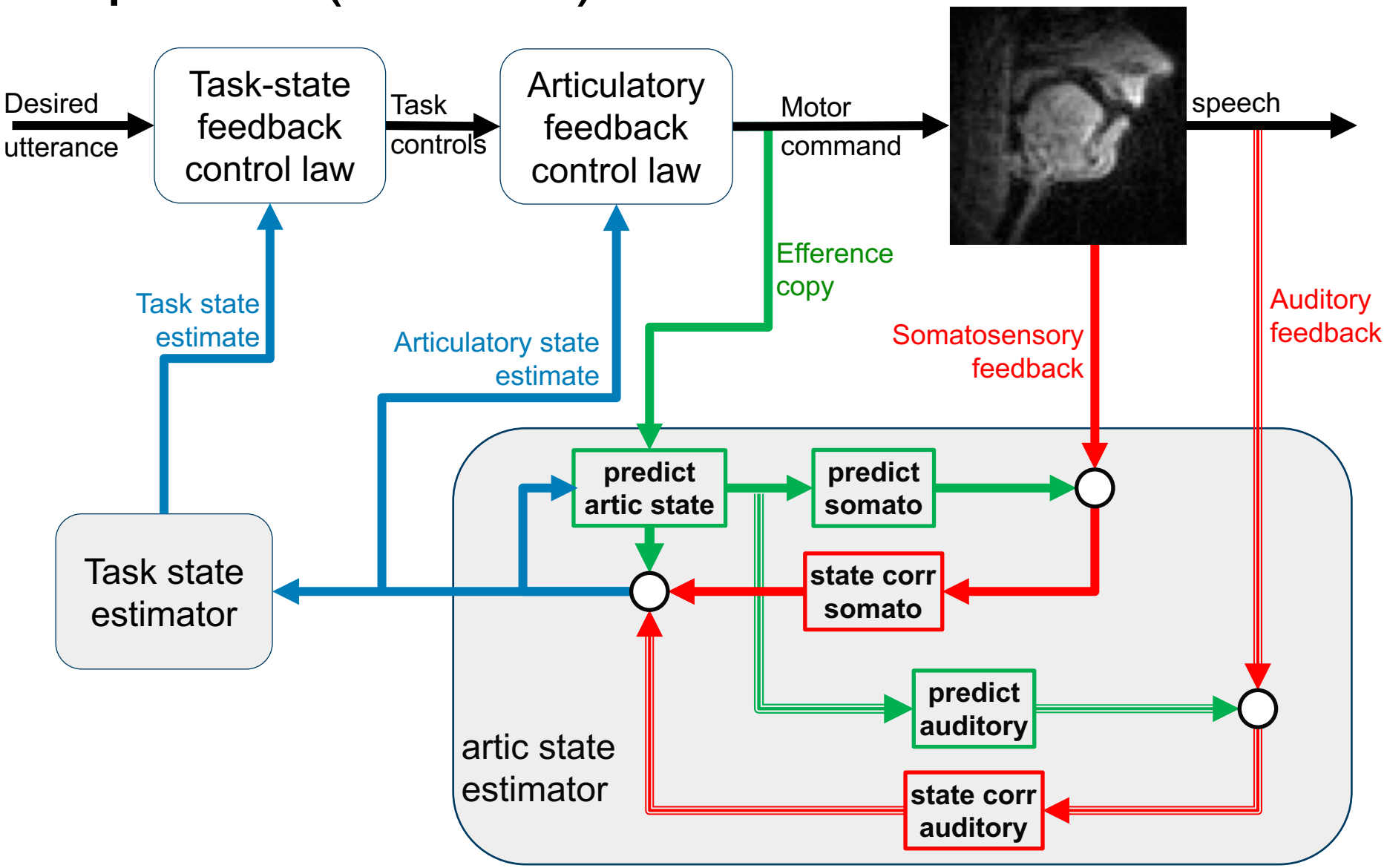
Feedback-Aware Control of Tasks in Speech (FACTS) model



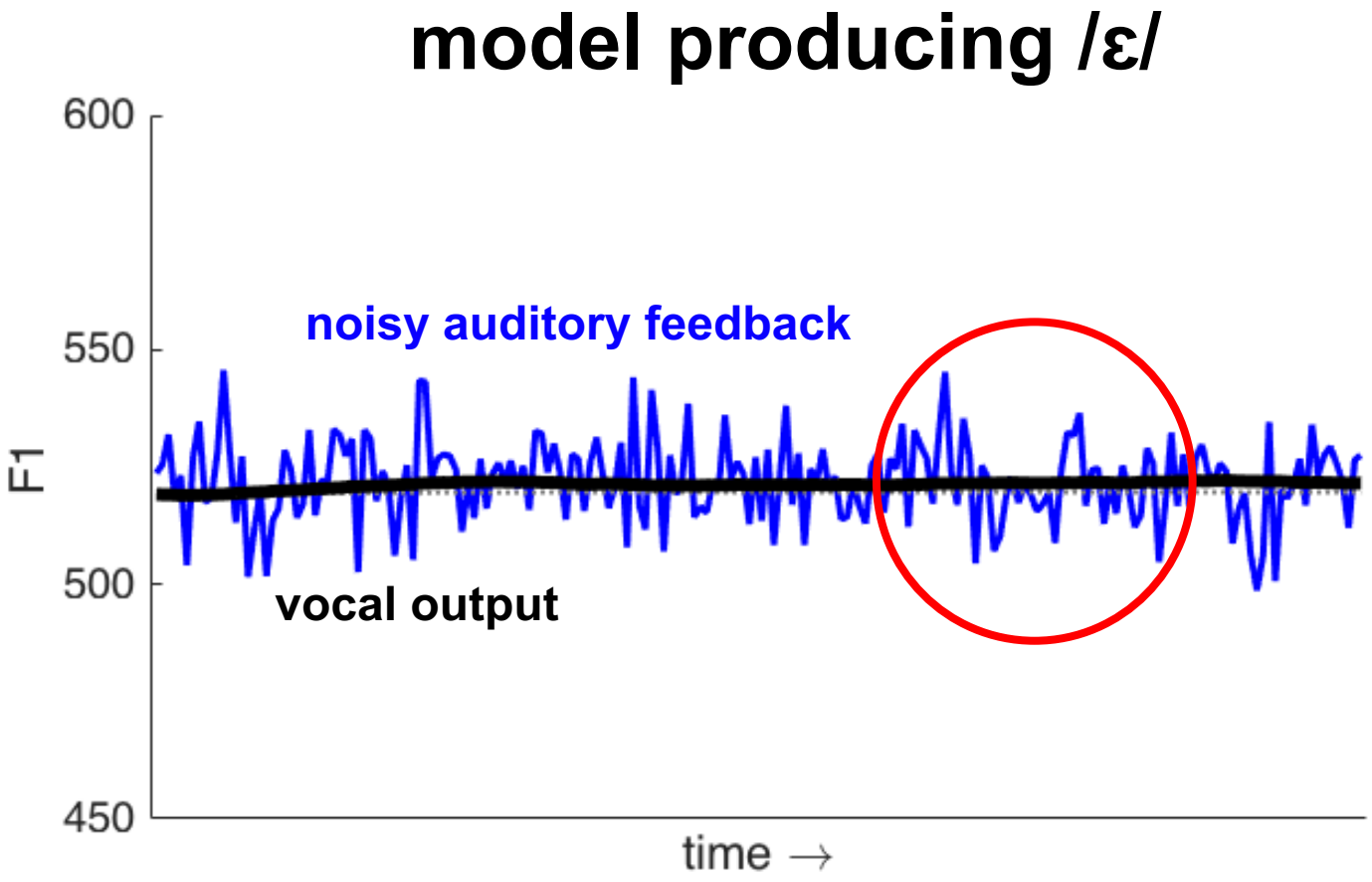
Feedback-Aware Control of Tasks in Speech (FACTS) model



Feedback-Aware Control of Tasks in Speech (FACTS) model

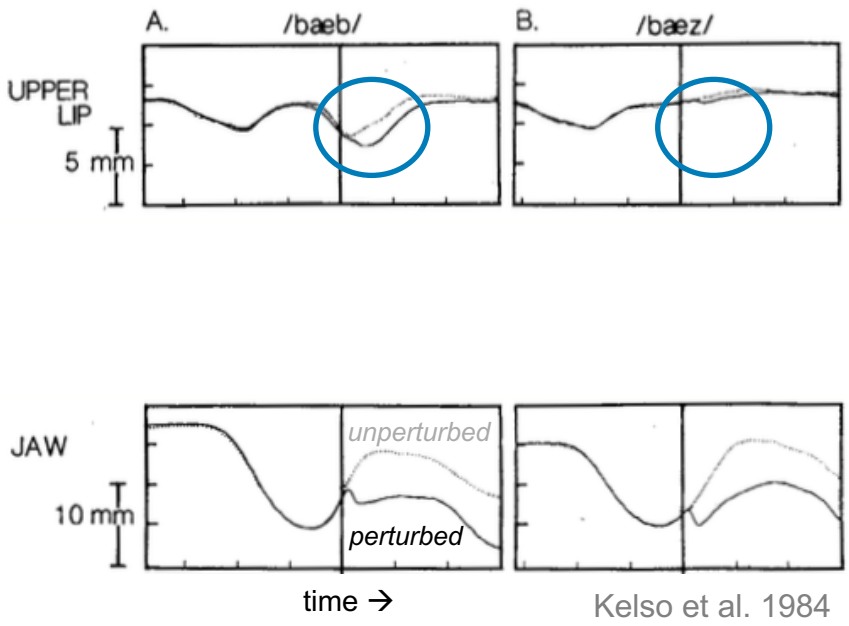


FACTS produces stable speech in the presence of sensory noise



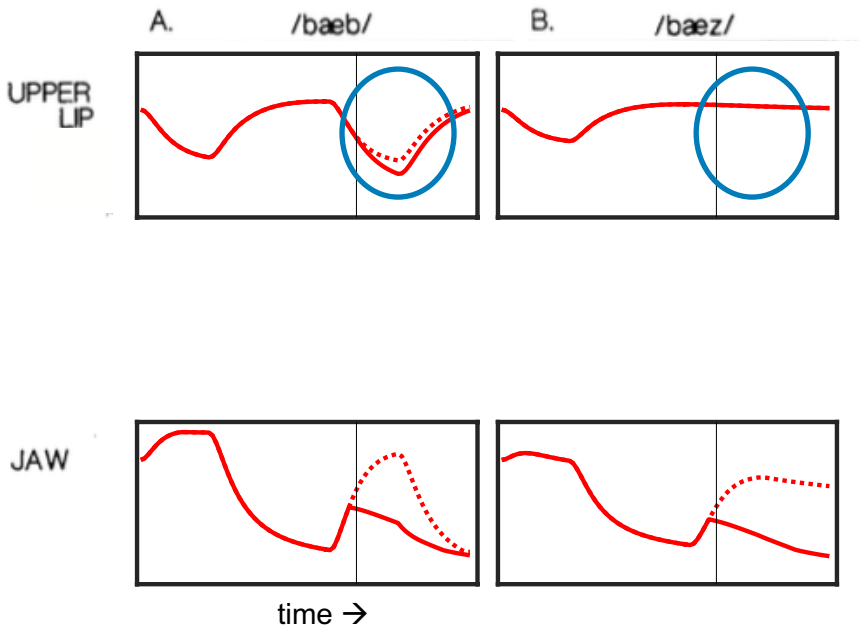
FACTS simulates task-specific responses to mechanical perturbations

Human data

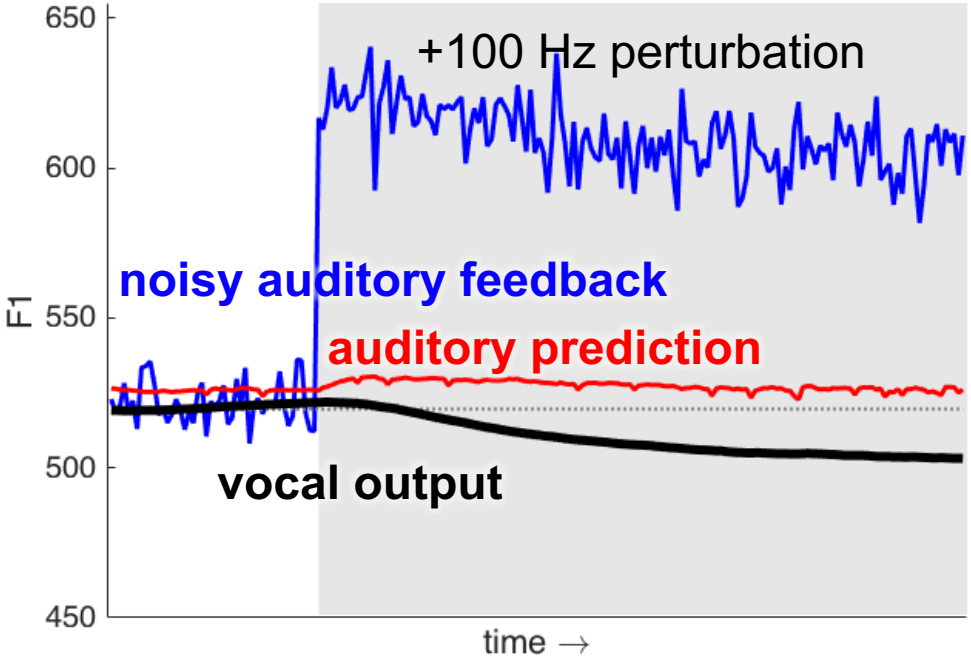


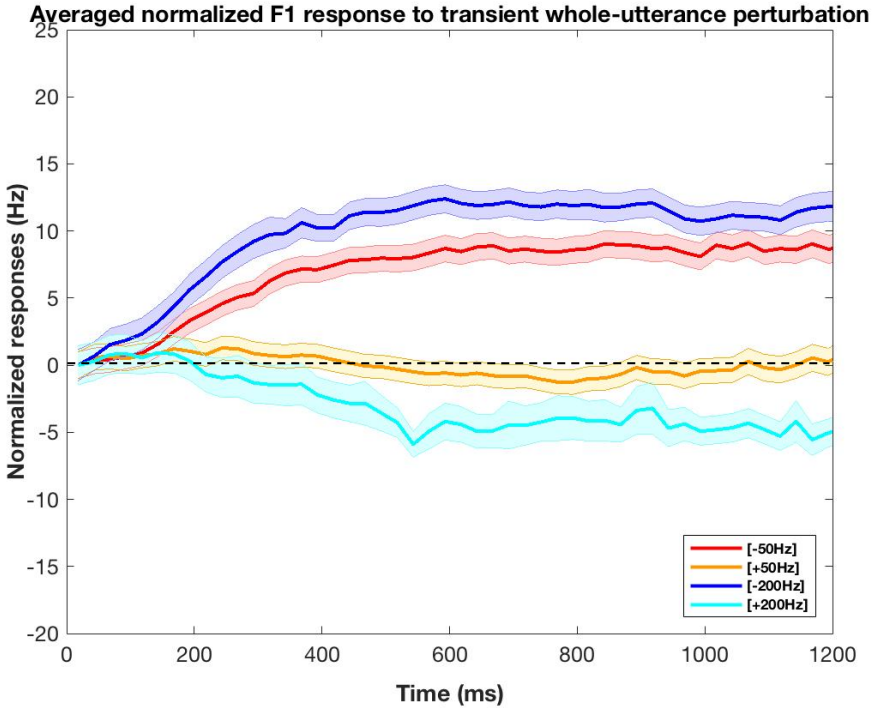
Kelso et al. 1984

Model data

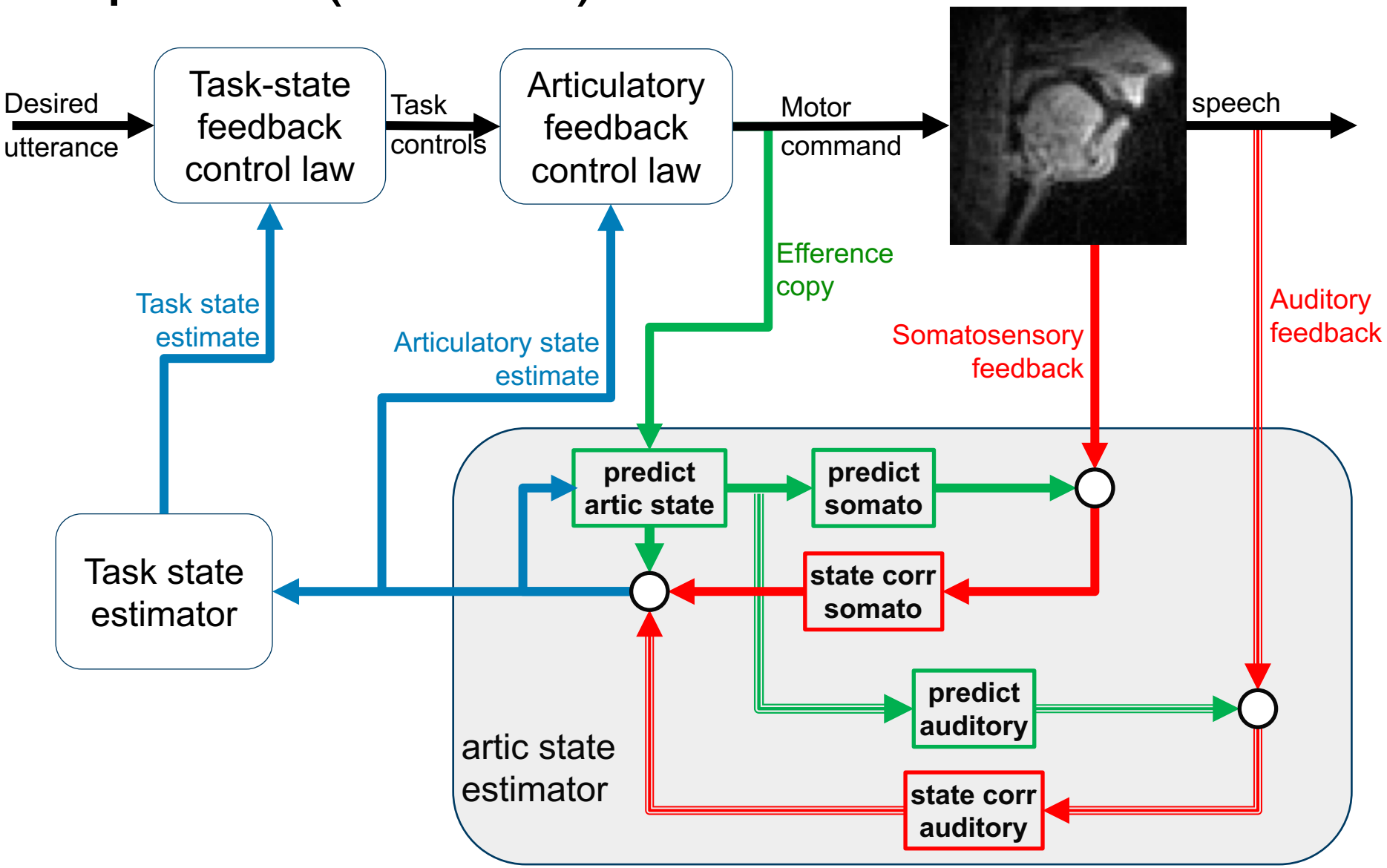


FACTS simulates partial compensation for auditory perturbations

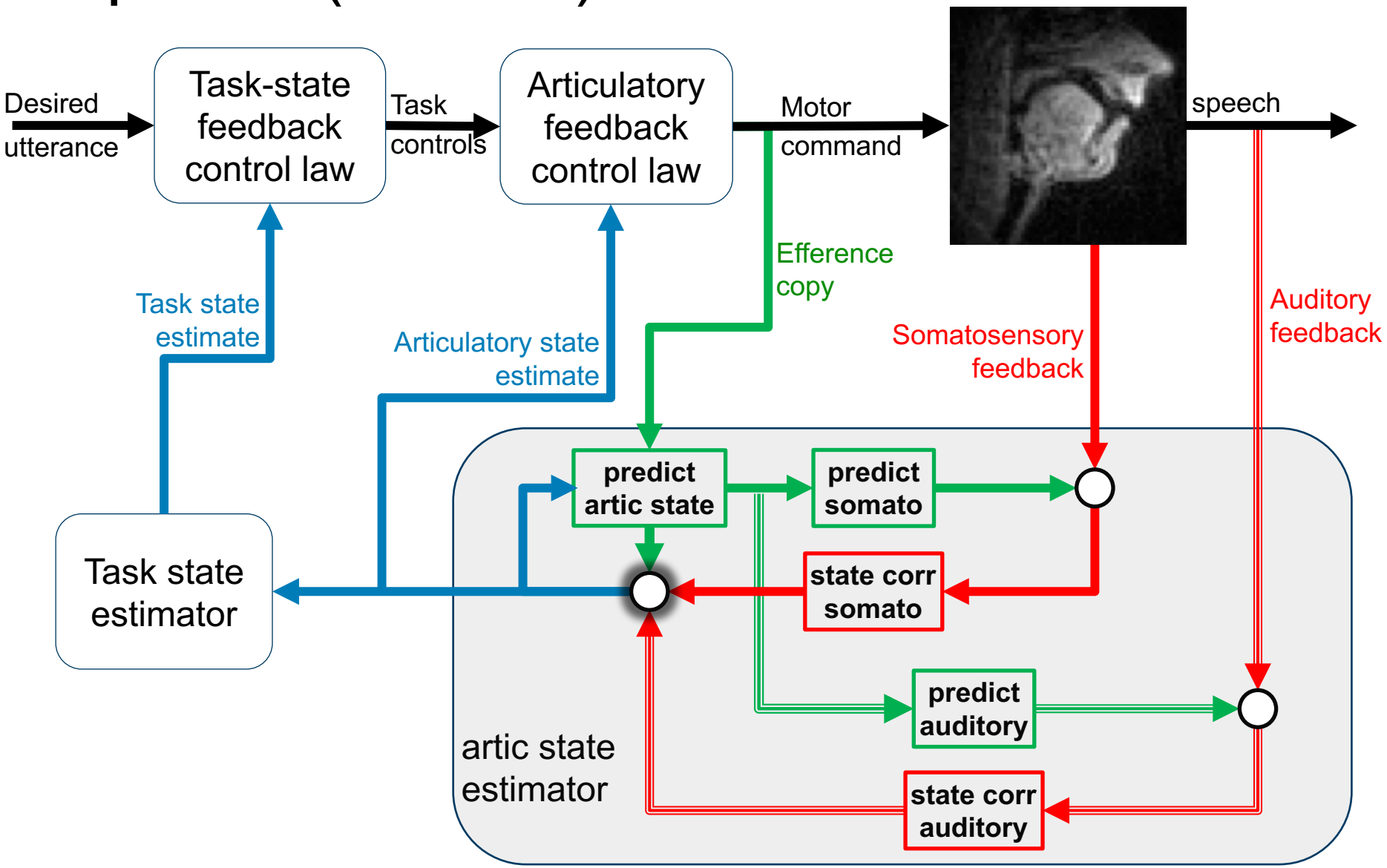




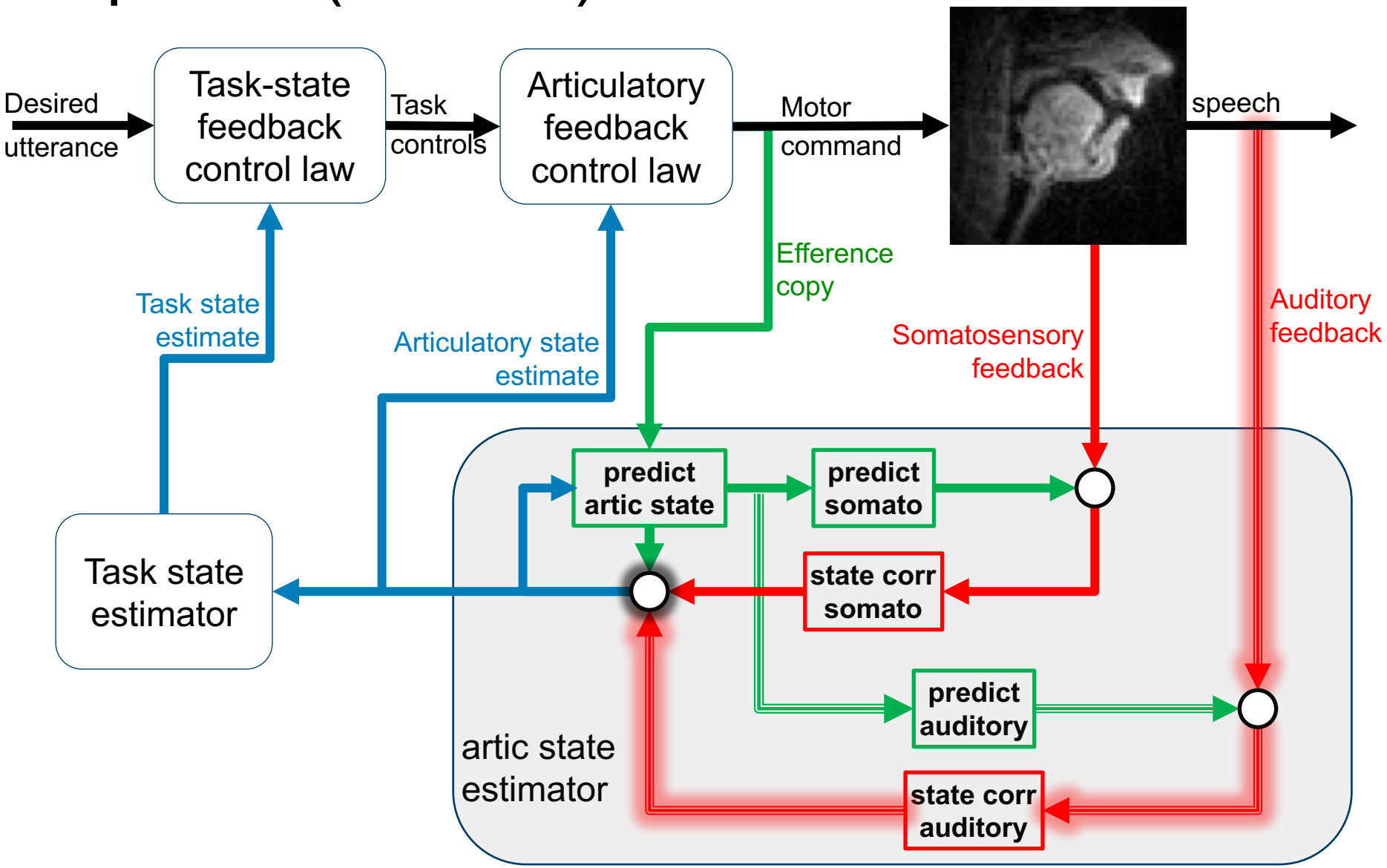
Feedback-aware Control of Tasks in Speech (FACTS) model



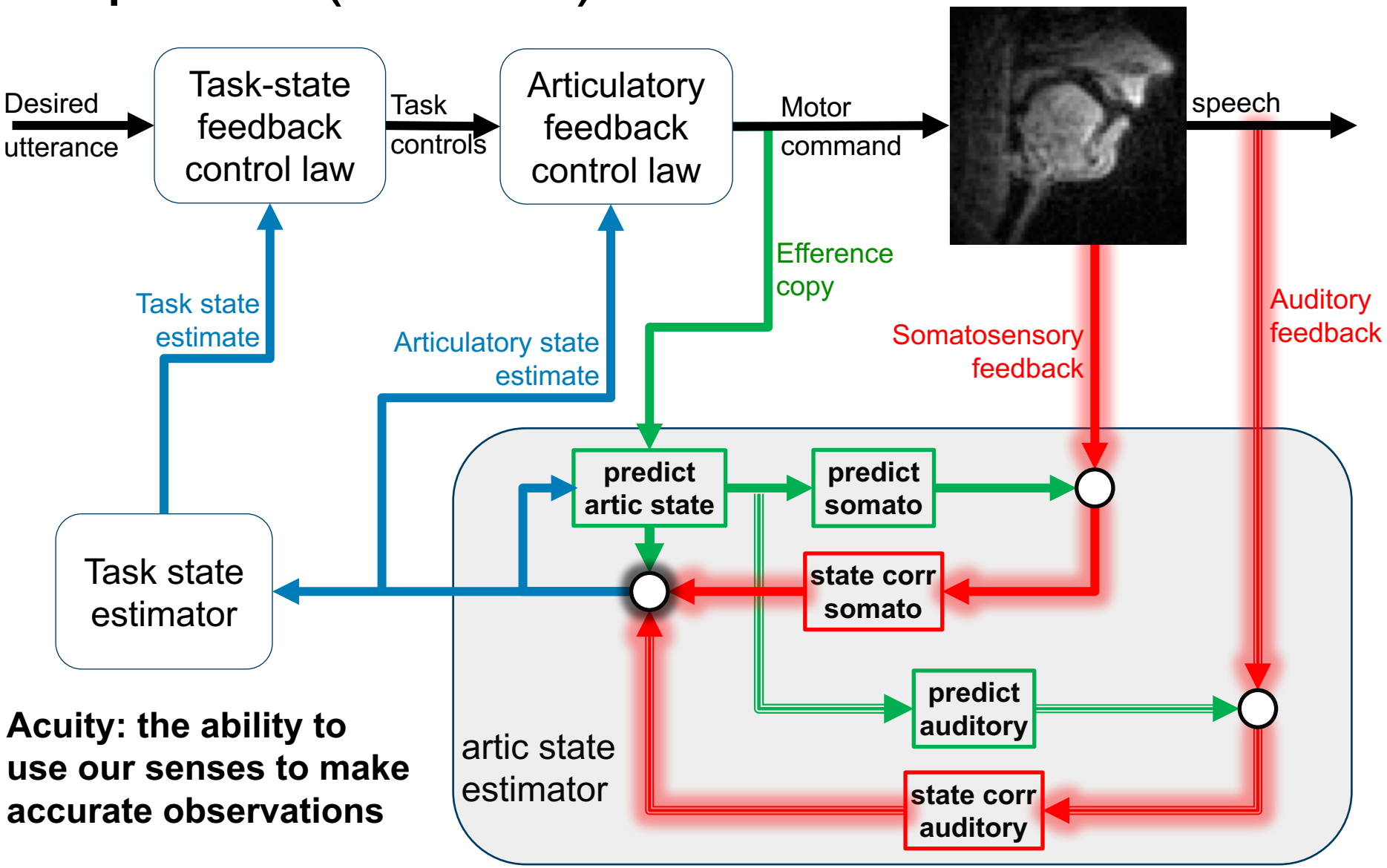
Feedback-aware Control of Tasks in Speech (FACTS) model



Feedback-aware Control of Tasks in Speech (FACTS) model

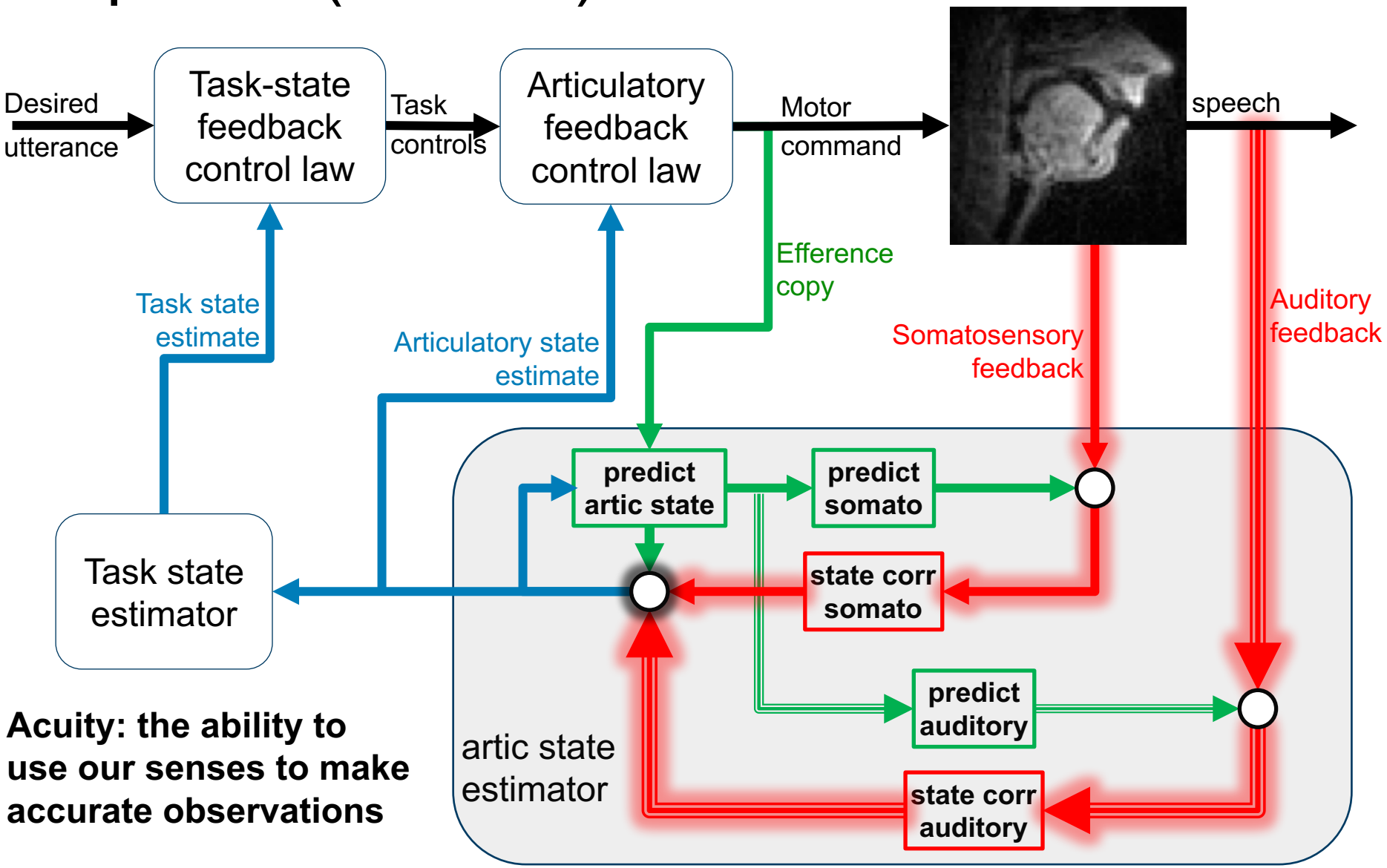


Feedback-aware Control of Tasks in Speech (FACTS) model



Acuity: the ability to use our senses to make accurate observations

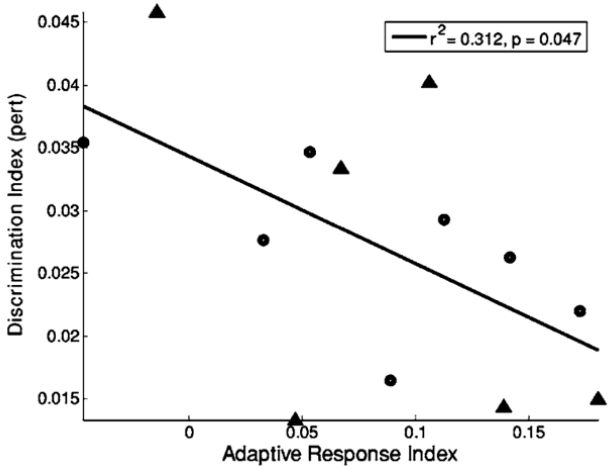
Feedback-aware Control of Tasks in Speech (FACTS) model



Acuity: the ability to use our senses to make accurate observations

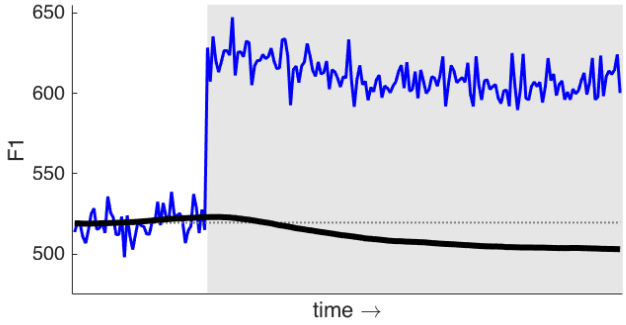
Some evidence for acuity-response relationship

Speakers with better auditory acuity adapt more to auditory perturbations

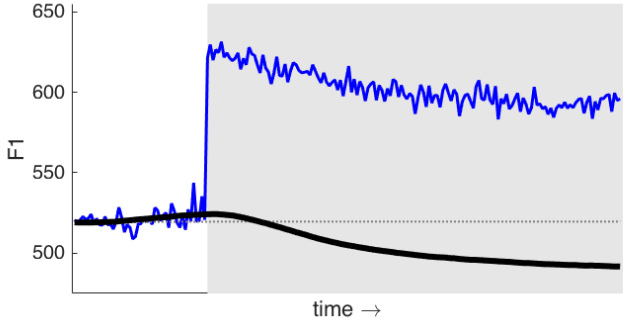


Villacorta et al. 2007

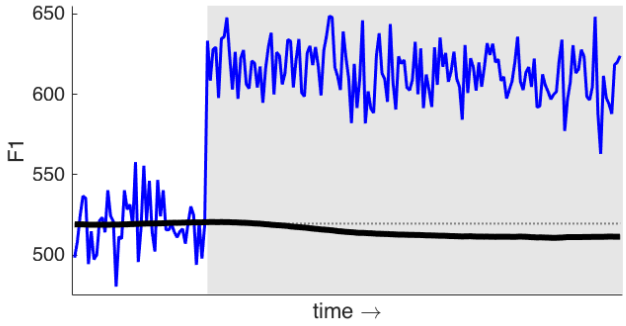
Baseline response



⇧ **auditory acuity**

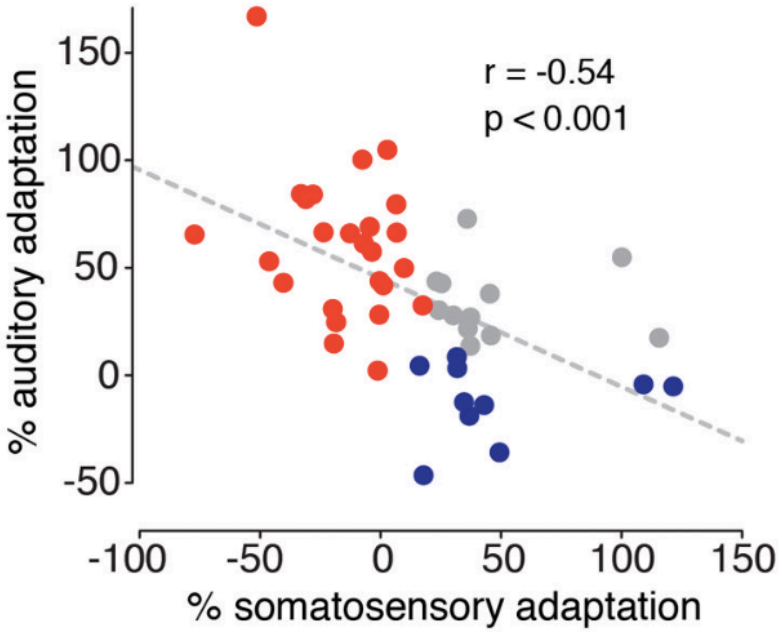


⇩ **auditory acuity**



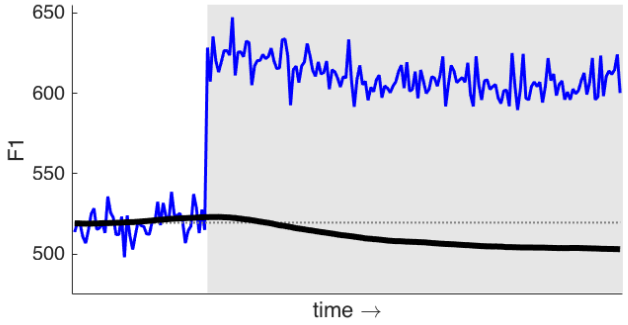
Some evidence for acuity-response relationship

Speakers' response to auditory perturbations is *negatively* correlated with their response to somatosensory perturbations



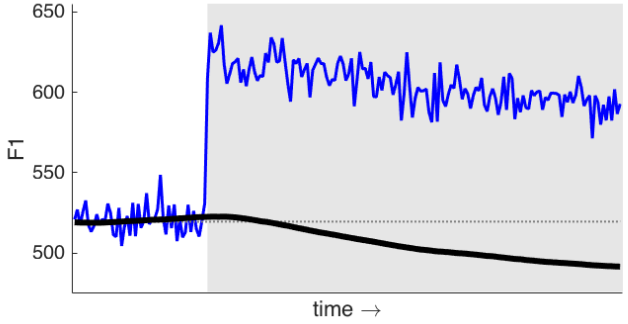
Lametti et al. 2012

Baseline response



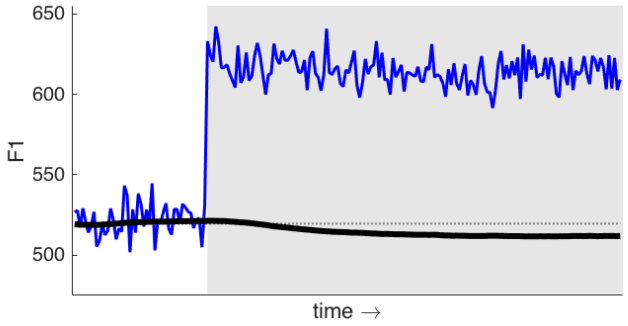
⇧ auditory acuity

⇩ somato-sensory acuity



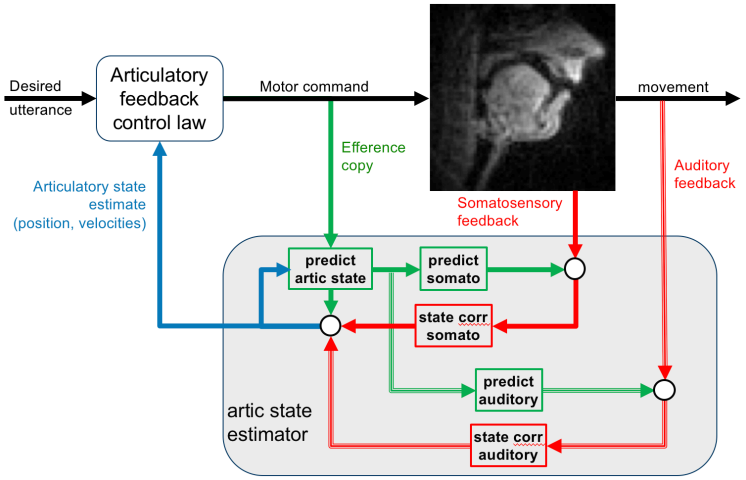
⇩ auditory acuity

⇧ somato-sensory acuity



FACTS model

- Is responsive to, but not reliant on, sensory feedback
- Produces stable speech in the presence of noise
- Reproduces full task-specific compensation for mechanical perturbations
- Reproduces partial compensation for auditory perturbations
- Models trade-off between auditory and somatosensory perturbation responses



What's next?

- Model neural system
- How is the prediction process learned and updated?
- Is multisensory feedback combined to estimate a single state or are auditory and somatosensory signals treated separately?

Computational implementation