



HAL
open science

Facial expression and phonetic recalibration in speech perception

Seung Kyung Kim, Sunwoo Jeong, James Sneed German

► **To cite this version:**

Seung Kyung Kim, Sunwoo Jeong, James Sneed German. Facial expression and phonetic recalibration in speech perception. Architectures and Mechanisms of Language Processing - AMLAP 2017, Sep 2017, Lancaster, United Kingdom. . hal-01794362

HAL Id: hal-01794362

<https://hal.science/hal-01794362>

Submitted on 24 Apr 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

- Perceptual learning of speech = Listeners adjust and recalibrate their phonetic boundaries based on exposure to new speech input (Norris et al., 2003)
- Social selectivity in learning
 - In general, people learn better with a model/teacher they view more positively (e.g., Chudek et al., 2012; Westfall et al., 2016)
 - People imitate the speech of a model speaker/interlocutor more when the person is more likeable/attractive (e.g., Babel, 2012; Pardo et al., 2012)

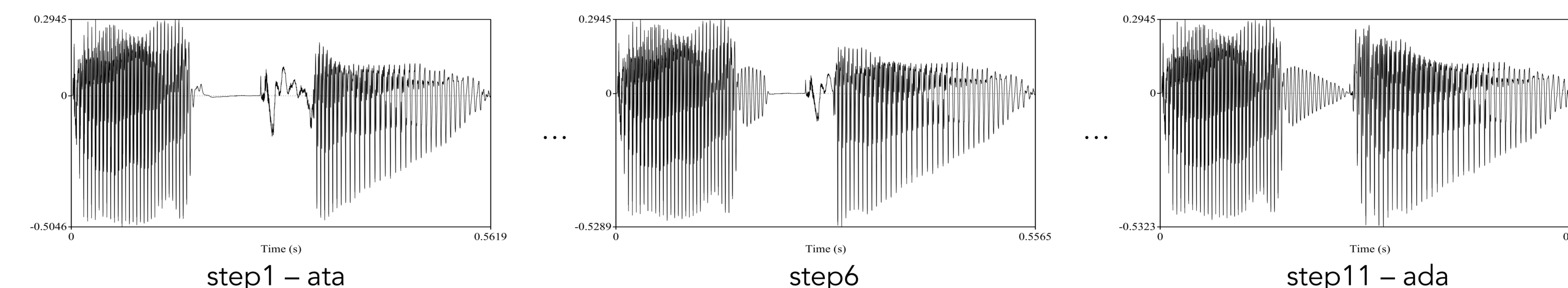
Question

Does perceptual learning of speech show social selectivity?

Do speech-external speaker characteristics (e.g., facial expressions) affect perceptual learning of speech?

11-step ata – ada continuum

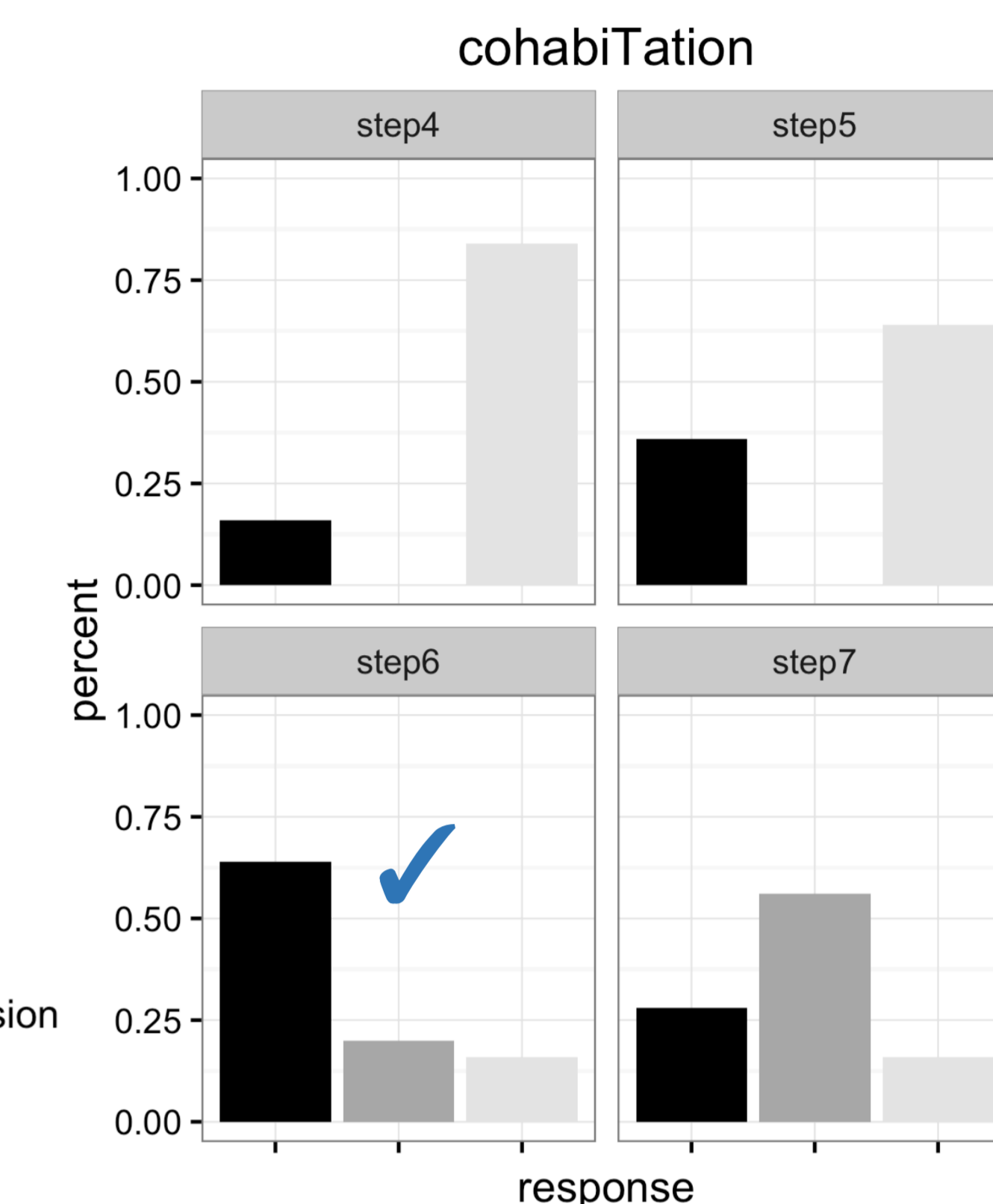
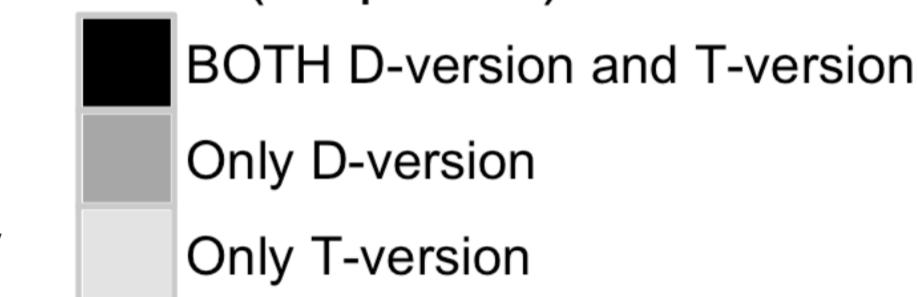
Manipulated VOT, closure duration, closure voicing, following f0



T-word manipulation

Created an 11-step continuum for each word; chose the most ambiguous step determined by a separate norming study

factor(response)



Exposure Phase – Lexical Decision Task

- T-words: 60 /t/-medial words where /t/ is the onset of the primary stressed syllable (e.g., *politician*)
- D-words: 60 /d/-medial words where /d/ is the onset of the primary stressed syllable (e.g., *academia*)
- Real-word fillers: 60 words with neither /t/ nor /d/
- Non-word fillers: 180 non-words with neither /t/ nor /d/
- /t/-sounds manipulated to be ambiguous between /t/ and /d/; /d/-sounds remain unambiguous (→ Shifting the category boundary toward /d/)
- Between-subject Image Condition:
 - No image
 - Face with no smile
 - Face with smile

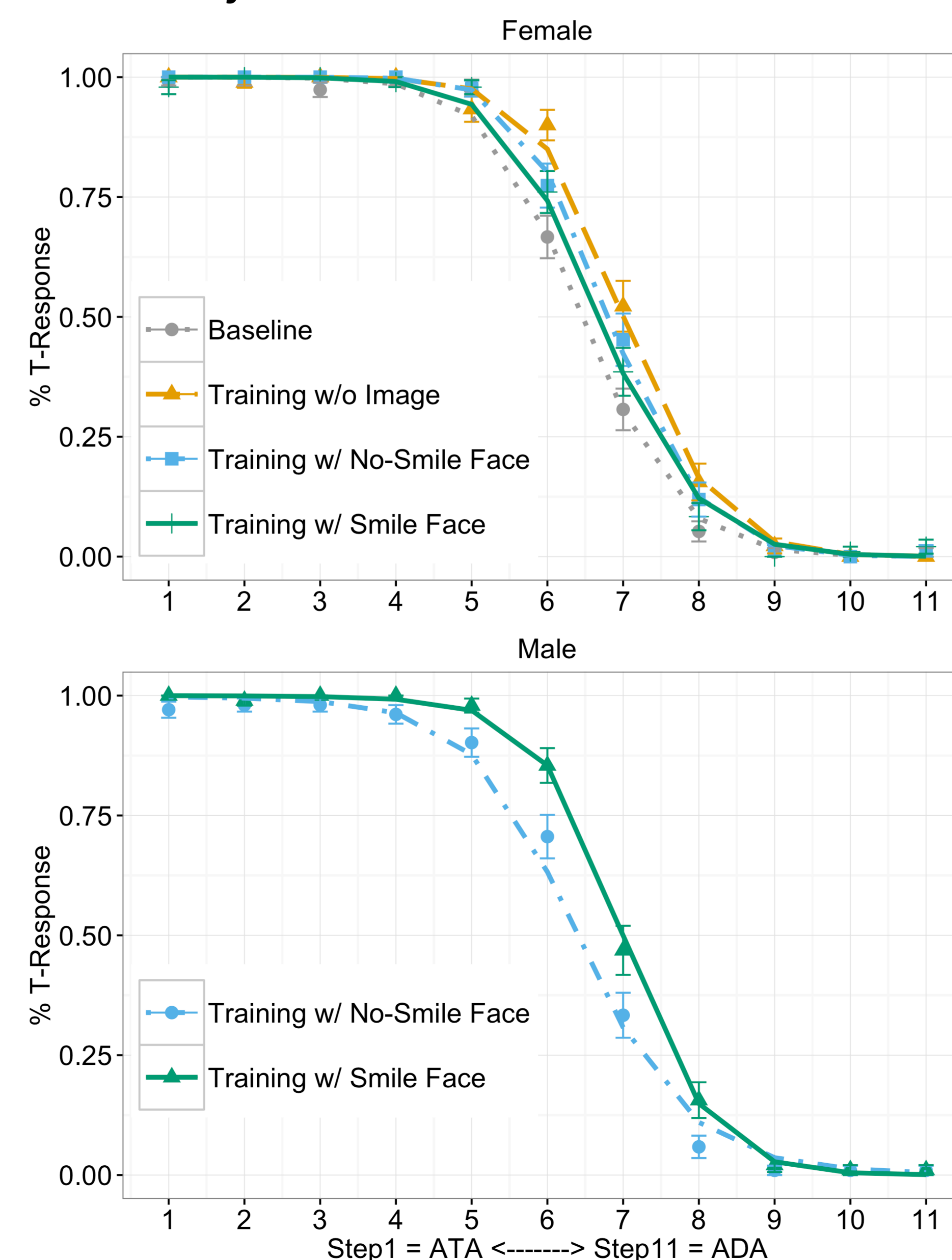


Post-Exposure Test Phase – Categorization Task on an 11-step ata – ada continuum

Implicit Social Encoding Hypothesis

Implicit social evaluations modulate encoding biases
Listeners will show a greater degree of perceptual learning when the speech is paired with a more positively viewed face

Preliminary results



Logistic mixed effect models
n = about 16 per curve
Step is centered (thus, step 6 = 0)
Random effect structure:
(1|participant) + (0+step|participant)

Female Participants

- Compared to Baseline Curve,
- No-Image Curve has a significant higher intercept ($b = 1.20, z = 2.32, p = 0.02$)
 - No-Smile Curve also has a higher intercept ($b = 1.16, z = 2.16, p = 0.03$)
 - But Smile Curve does not have a higher intercept ($b = 0.44, z = 0.9, p = 0.37$)
 - In addition, No-Smile Curve has a significant steeper slope ($b = -0.63, z = -2.44, p = 0.01$)

Compared to Smile Curve, No-Smile Curve does not show a meaningful difference

Male Participants

- Compared to No-Smile Curve,
- Smile Curve has a significant higher intercept ($b = 1.41, z = 3.28, p = 0.001$)
 - In addition, Smile Curve has a marginally significant steeper slope ($b = -0.42, z = -1.66, p = 0.097$)

Summary and conclusions

- Male participants shift their perceptual boundary according to the new input more with the smile face than the no-smile face
- Female participants don't seem to be influenced by the faces (if any, they seem to adapt more with the no-smile face)
- Gender difference observed
Different patterns may be found with male faces (cf. Babel 2012)
- Initial evidence for implicit social encoding hypothesis
- Encoding of new speech input is modulated by speech-external social conditions (cf. Sumner et al., 2014 for social encoding due to speech-internal social information)
- Different social conditions attract different degrees of attention, modulating encoding strength