

Learning non-native speech stop voicing as a result of exposure to foreign film

BACKGROUND

- Previous research on the acquisition of non-native speech overwhelmingly supports the need for authentic, native-speaker input in that language in order to develop target-like linguistic competence (Flege, 2007; Flege & Liu, 2001; MacKay et al., 2001)
 - What constitutes as “native-speaker” input?
 - Only face-to-face interaction?
- Foreign film may serve as a form of high variable phonetic input (i.e. input that exposes participants to natural productions by multiple talkers and a wide range of tokens) (Bradlow, 2008; Bradlow et al., 1997, 1999; Lively et al., 1993, 1994; Logan et al., 1991; inter alia)
- Foreign film benefits from having both an audio and visual component
 - Audiovisual perceptual training seems to be more effective than auditory training (Hardison, 1999, 2003; Hazan et al., 2005, 2006; Lively et al., 1993; Ortega-Llebaria et al., 2001; Shen, 2019; Wang et al., 2008, 2009; inter alia)

Main Research Question

Can foreign film contribute to more target-like non-native stop voicing production?

SHADOWING METHODS

Talkers

- 74 monolingual native speakers of American English (15 male, 57 female, 1 non-binary, and 1 declined to answer; mean age 21.17 y.o, SD=3.27)

Stimuli

- Recorded by a native speaker of French (male, 22 y.o.) in a sound-attenuated booth
- Monosyllabic CV or CVC French word containing at least one of the target sounds: word-initial bilabial stops (e.g. *pouce* [pus], *bout* [bu])
- Represented across 6 stimulus items and repeated three times per session

	[+voice]	[-voice]
French	Prevoiced (VOT <0 ms)	Short lag (VOT <30 ms)
English	Short lag/prevoiced	Long lag (VOT >30 ms)

Testing procedures (Pre-test and Post-test)

- Talkers heard a word through headphones and were asked to repeat it back into a microphone
- Data collection took place in a sound-attenuated booth using PsychoPy

Film Exposure

- Season 1, Episode 1 of Chef’s Table: France (45 min)
- Episode was viewed at a computer station in a quiet room equipped with a sound dampening tri-fold stall
- Participants completed a while-watching activity based on their group assignment (Audio-Only, Subtitled, Control)

Data Processing

- Annotated manually in Praat; VOT was annotated from the release of the stop to the onset of voicing
- A custom Praat script extracted duration measurements

VOT DURATION SHADOWING RESULTS

Duration Analysis

- Three linear mixed effects models: voiceless, positive voiced, and negative voiced
- **Dependent variable:** Duration (ms)
- **Fixed effects:** Group (Audio-Only, Subtitled, and Control) and Session (pre-test and post-test)
- **Random effects:** Subject and Item (Intercepts)

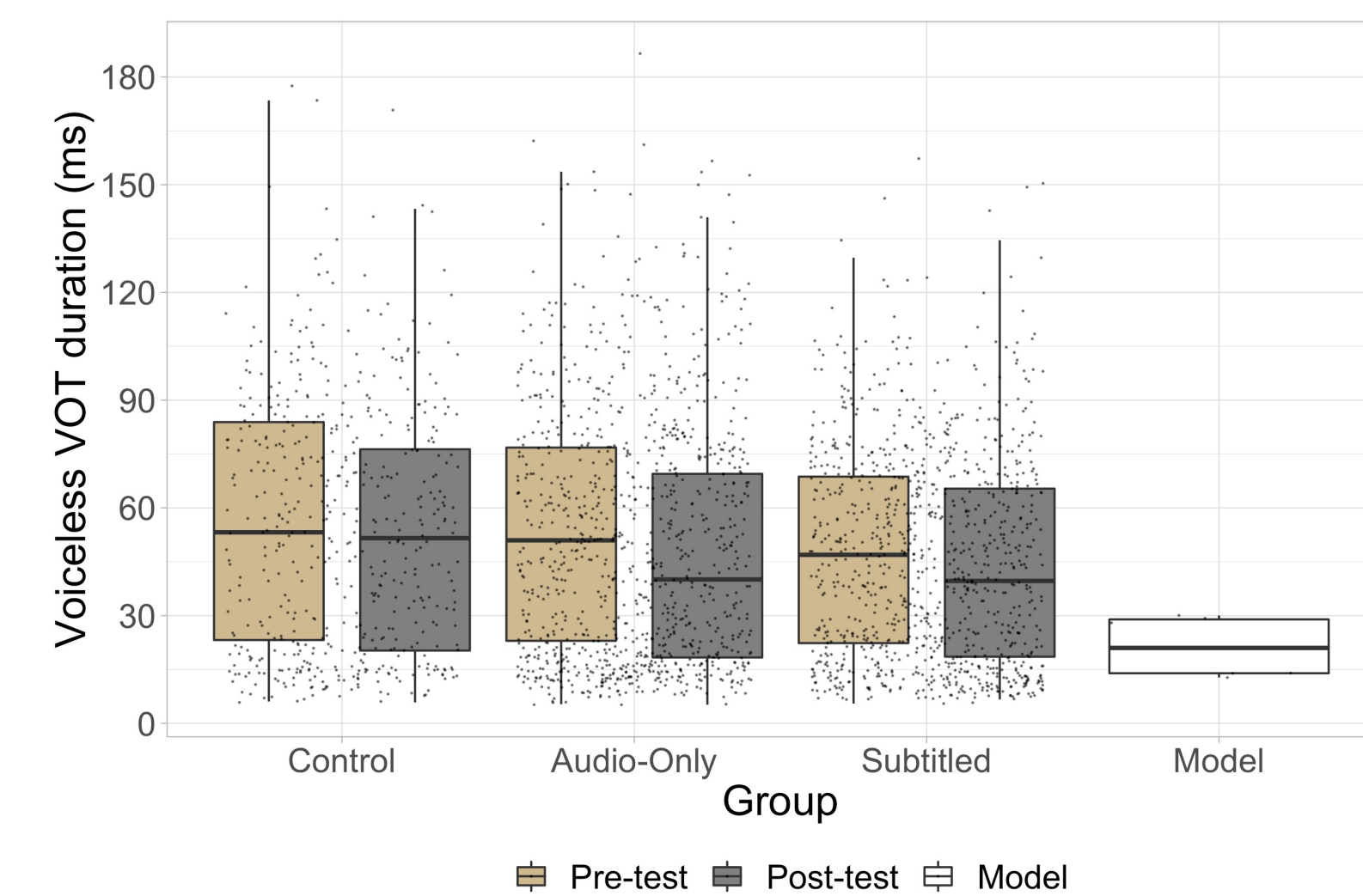


Fig 1. Voiceless VOT duration across groups and sessions with the model talker for reference.

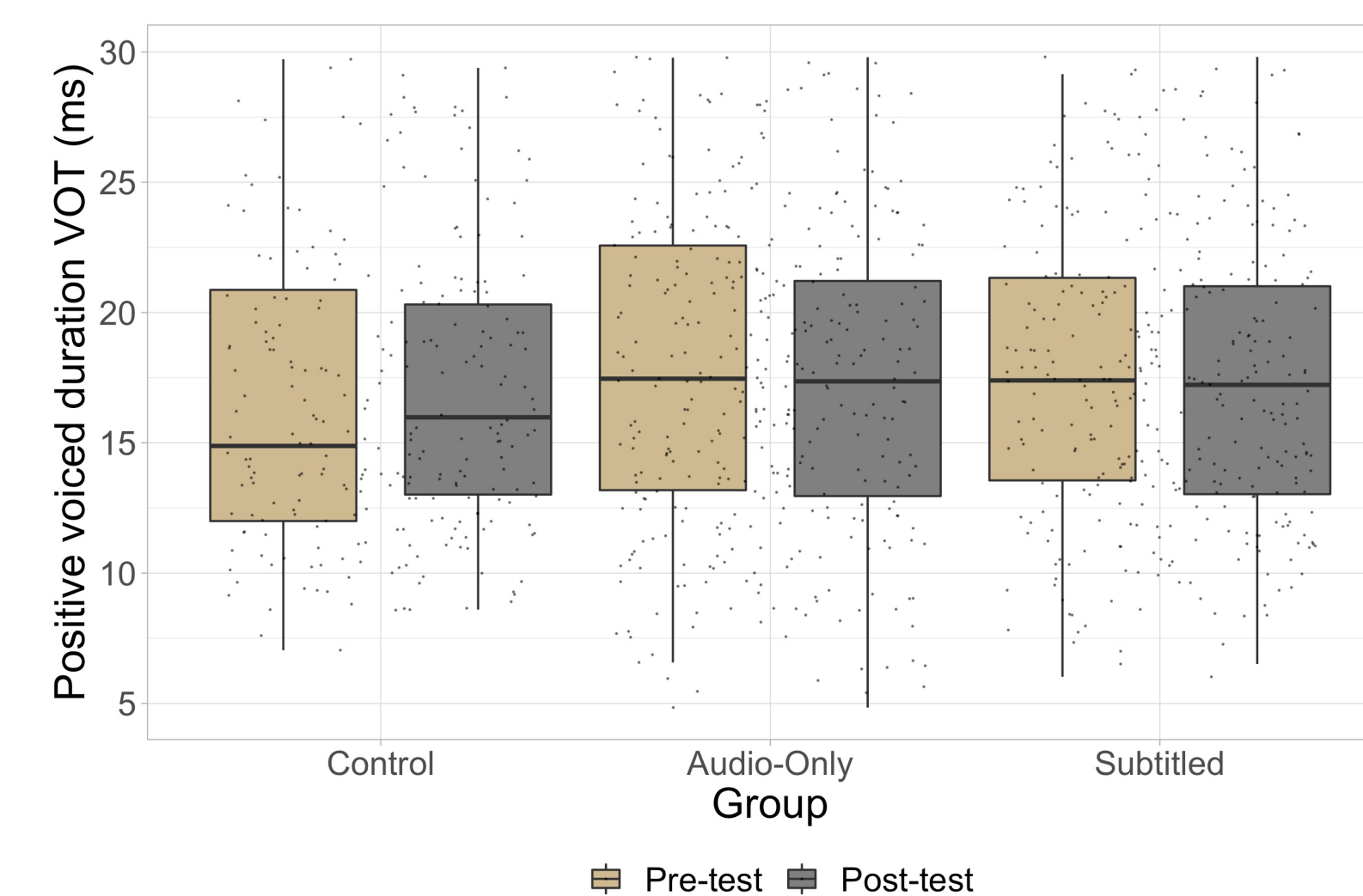


Fig 2. Positive voiced VOT duration across groups and sessions.

Voiceless Duration Results

- No statistical evidence to suggest that film affected voiceless VOT duration [Audio-Only vs. Control: $t = -0.65$; Subtitled vs. Control: $t = 0.07$]

Positive Voiced Duration Results

- Despite visible shortening of average positive voiced durations by experimental groups in the post-test, this was not significant [Audio-Only vs. Control: $t = -1.47$; Subtitled vs. Control: $t = -0.99$]

Negative Voiced Durations Results

- The Audio-Only group produced significantly different negative VOT durations in the post test [$t = 2.07$], but the Subtitled group did not [$t = 1.92$]

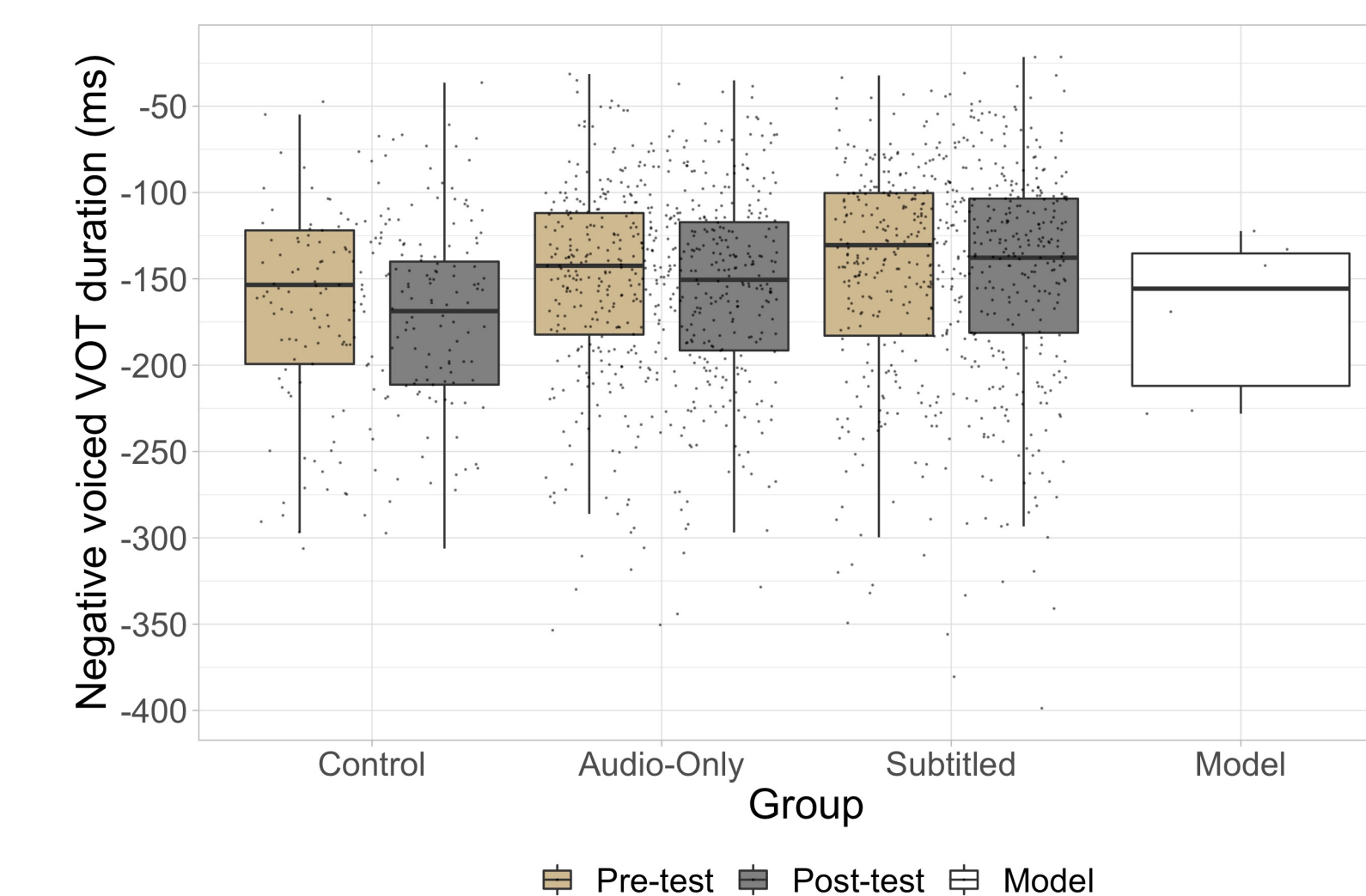


Fig 3. Negative voiced VOT duration across groups and sessions with model talker for reference.

VOT CATEGORY SHADOWING FREQUENCY RESULTS

Category Frequency Analysis

- Two separate mixed effects binomial logistic regression models: voiceless and voiced
- **Dependent variable:** Category (prevoiced, short lag, or long lag)
- **Fixed effects:** Group (Audio-Only, Subtitled, and Control) and Session (pre-test and post-test)
- **Random effects:** Subject and Item (Intercepts)

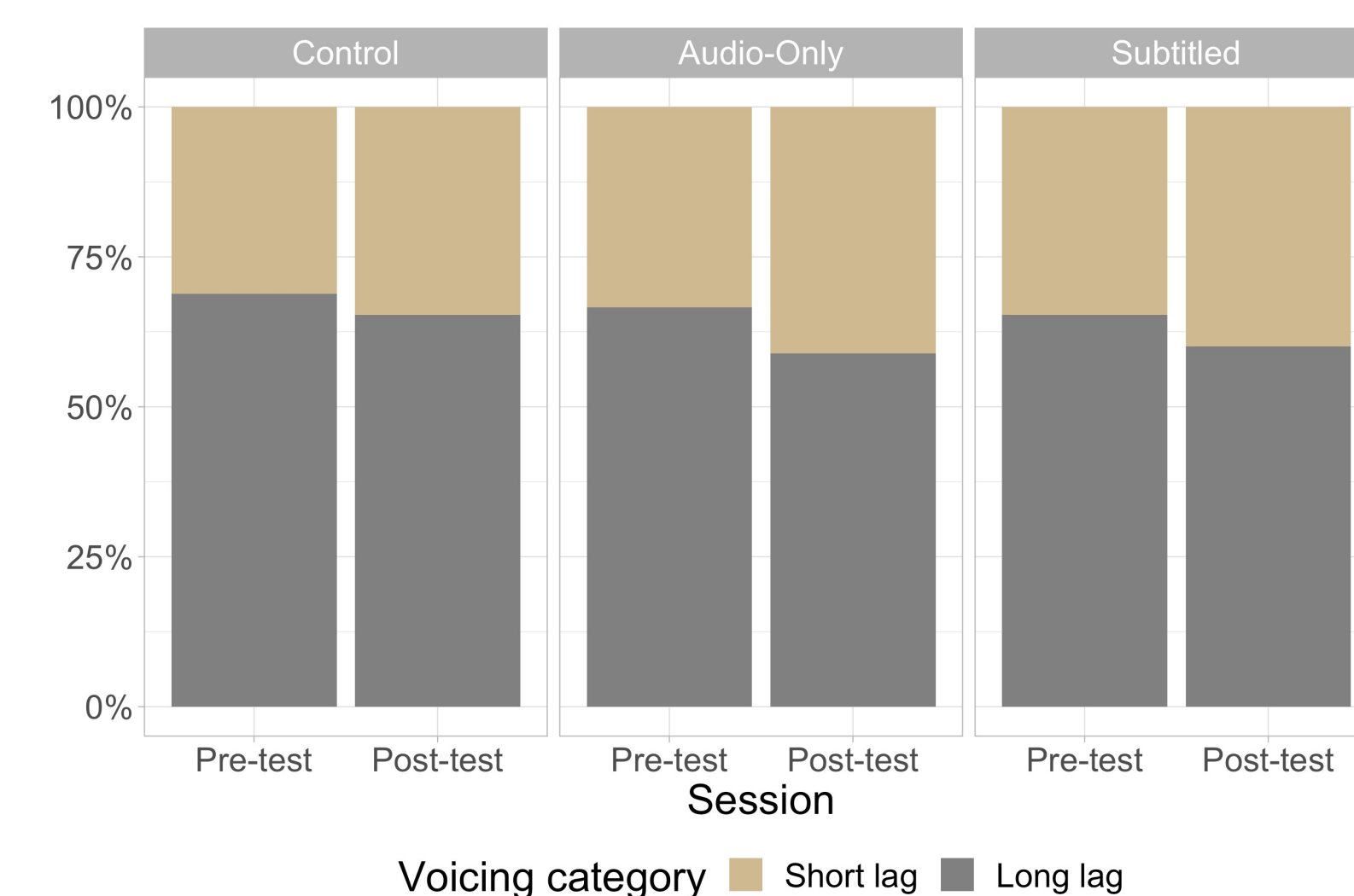


Fig 4. Frequency of voiceless categories (short and long lag) across groups and sessions.

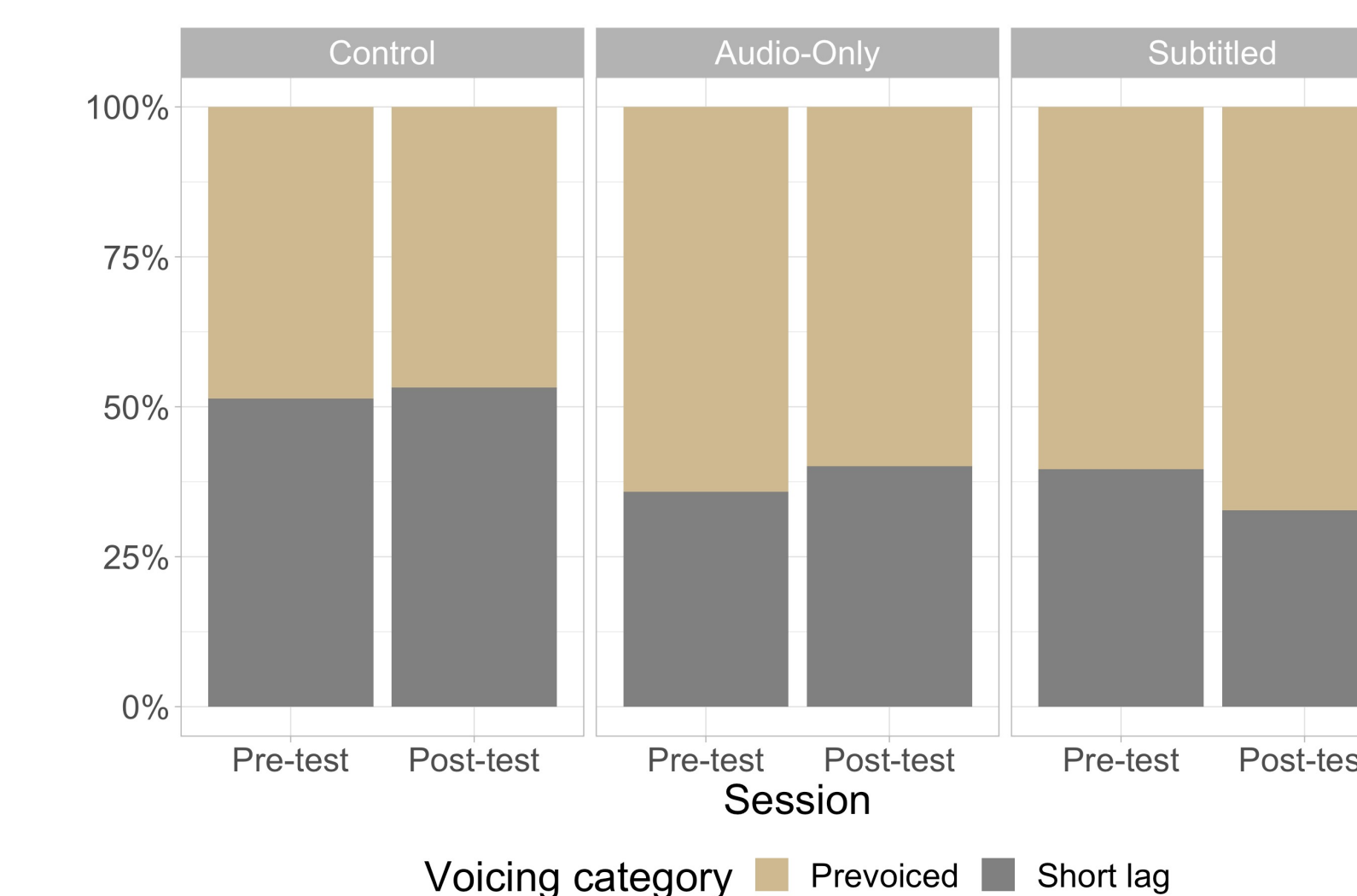


Fig 5. Frequency of voiced categories (prevoiced and short lag) across groups and sessions.

- Marginal attempts to realize more French voiceless stops as short lag and voiced stops as prevoiced (Subtitled group only) in the post-test were made, but film exposure did not significantly affect category distribution for voiceless [Audio-Only vs. Control: $p = 0.36$; Subtitled vs. Control: $p = 0.72$] or voiced stops [Audio-Only vs. Control: $p = 0.82$; Subtitled vs. Control: $p = 0.07$]

PERCEPTUAL JUDGMENT METHODS

Listeners

- 222 native French listeners (135 male, 84 female, and 3 non-binary; mean age 28.18 y.o., SD=9.40) were recruited using Prolific

Stimuli

- Items collected during the pre-and post-test shadowing sessions (second repetition)

Procedures

- AXB paradigm: Listeners were asked to determine whether A or B (corresponding to pre- or post-test shadowings) were a better representation of X
- Each item was presented four times total (twice in each order) and played to three listeners

PERCEPTUAL JUDGMENT RESULTS

Perceptual Judgement Analysis

- Two mixed effects binomial logistic regression models: items containing voiceless /p/ and voiced /b/
- **Dependent variable:** Listener response (item A or B; corresponding to pre- or post-test shadowings)
- **Fixed effects:** Talker Group (Audio-Only, Subtitled, and Control)
- **Random effects:** Talker, Item, and Listener (Intercepts)

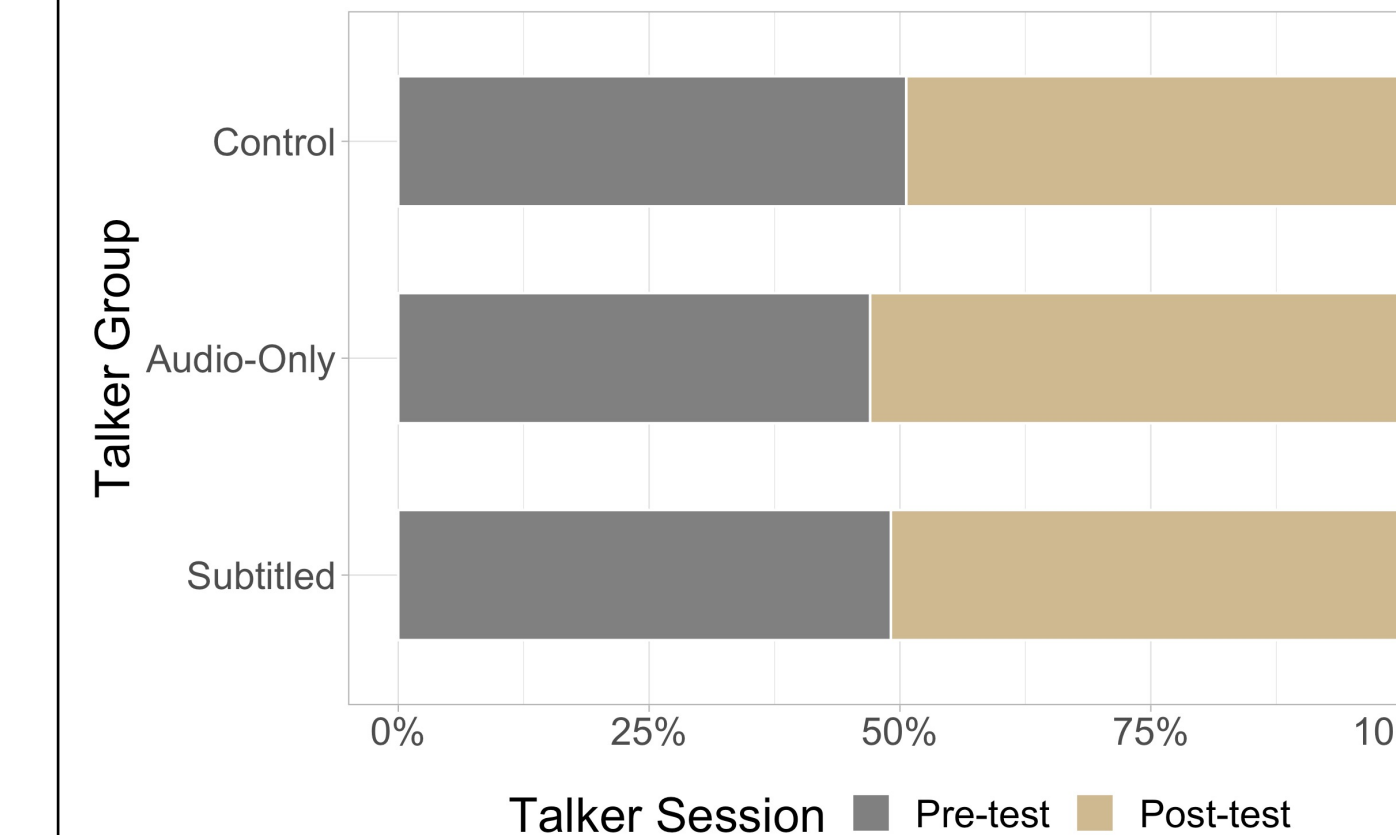
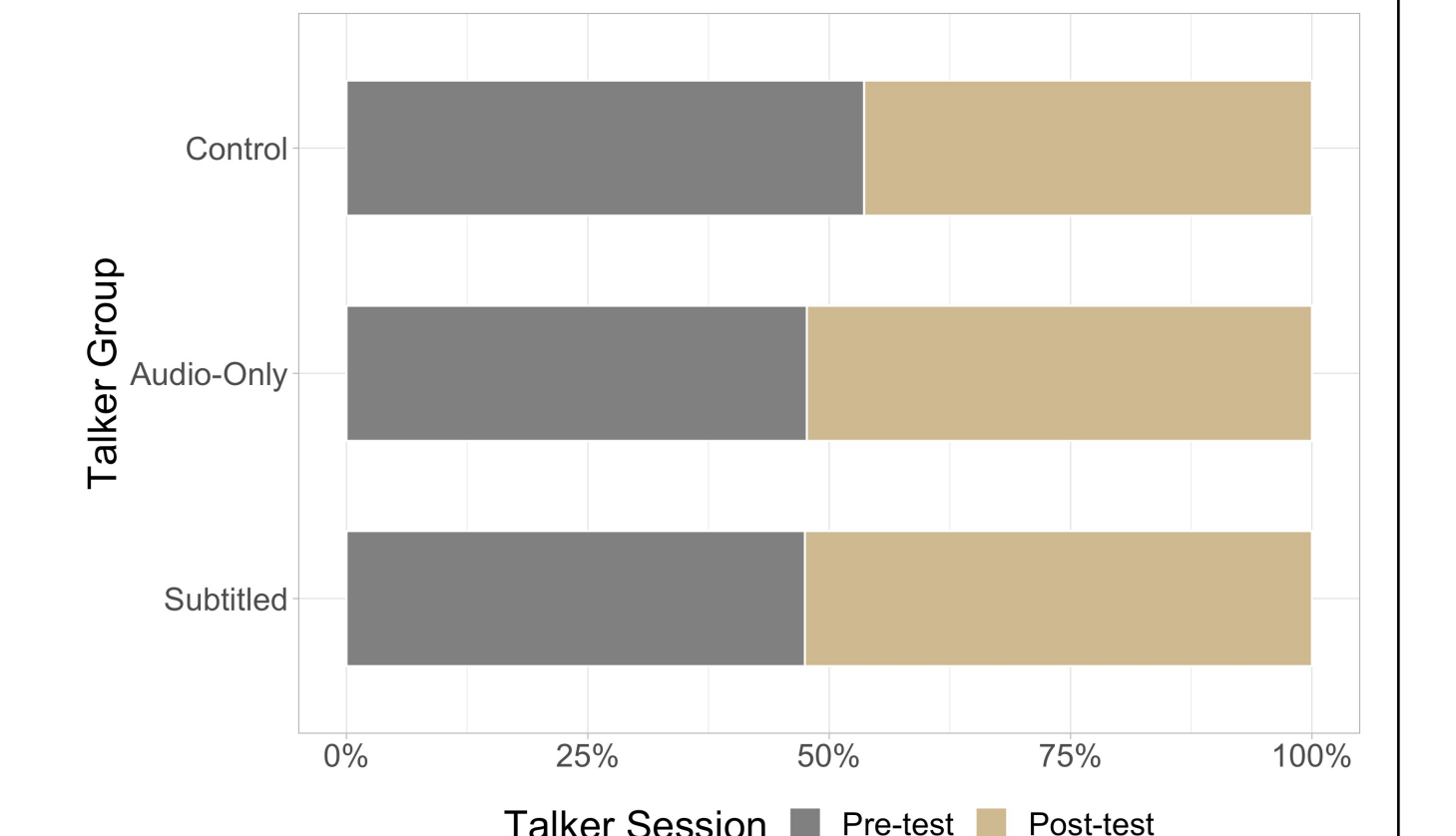


Fig 6. Listener selection of talker pre- or post-test /p/ productions as more similar to the model, by talker group

- **Voiceless /p/:** There was no effect of film exposure on the perceived similarity between model pronunciation and talkers’ shadowing [Audio-Only vs. Control: $p = 0.11$; Subtitled vs. Control: $p = 0.49$]

Fig 7. Listener selection of talker pre- or post-test /b/ productions as more similar to the model, by talker group



- **Voiced /b/:** Results revealed that listeners who heard experimental group participants were more likely to select post-test items than listeners who heard Control group talkers [Audio-Only vs. Control: $p = 0.02$; Subtitled vs. Control: $p = 0.02$]

DISCUSSION

- There is no evidence that exposure to foreign film had an affect on duration or category frequency of voiced or voiceless VOT
 - VOT is not a perceptually salient feature making learning through film challenging
- Perceptual judgements provided by native French listeners demonstrate that acoustic adjustments made by experimental group talkers when producing /b/ following film exposure were perceptually salient and selected at higher rates than pre-test productions
- Post-exposure, talkers are modifying acoustic correlates of voicing other than VOT, which is contributing to more frequent selection of post-test /b/ items