Effects of interlocutor feedback on speaker phonetic production in a simulated-communication task

Esteban Buz & T. Florian Jaeger

Department of Brain and Cognitive Sciences University of Rochester

Introduction

Is language production organized for robust communication?

• Speakers hyper-articulate contextually confusable words [1, 2].
• Some accounts argue this is partly due to production being organized to take into account perceived success of previous articulation [3, 4].
• Alternative accounts argue articulation differences are due to production ease or comprehension processes [1, 5 – 8].
• Determining whether speakers’ articulations change based on interlocutor feedback is crucial for distinguishing between accounts.

Can we use the web to study phonetic change?

• Usage of web-based paradigms allows for faster data collection but opens up questions if phonetic data is feasible.

Our Questions

1. Do we find contextually confusable hyper-articulation?
2. Does explicit listener feedback result in more hyper-articulation?

Study Design

Web-based speech collection

Participants

• 60 self reported native speakers of American English

Materials

• 36 critical target words with /k, p, t/ stop onsets with voiced minimal pairs (e.g. pig, stimuli from Kirov & Wilson 2012, study 2)
• Critical target words presented with or without minimal pair (context manipulation, Figures 1a)
• Between participants trials ended with three types of feedback (feedback manipulation, Figures 1c): None, positive only or Mixed

Believability of simulated-partner design

• 4 participants, unprompted, did not believe partner existed
• When told they might have interacted with a computer or human, partner ‘human-ness’ ratings were middling
• Experiment cover story plausibility was rated as good

Design results

• Longer VOTs for targets with co-present minimal pair (p < 0.01)
• Nominal but non-significant VOT difference across feedback types, non-significant interaction (p’s > 0.6)

Study Results

Believability of simulated-partner design

• 27 participants hand annotated for voice onset timing

• Linear mixed effects modeling predicting VOT (msecs)

- Significant linear term, marginal negative quadratic term (Fw = 74.17, -53.21, t = 3.4, -1.9)
- Context effect still significant (p < 0.01), feedback still non-significant (p = 0.68)

Acoustic analysis

- 8 participants, 29 percent preferred human

- Linear mixed effects modeling predicting VOT (msecs)

- No evidence that production ease (latency) affects VOT

- Significant linear term, marginal negative quadratic term (Fw = 74.17, -53.21, t = 3.4, -1.9)
- Context effect still significant (p < 0.01), feedback still non-significant (p = 0.68)

Results (cont’d)

• Follow-up analysis modeling VOT change by log latency

• No significant effect of log latency on VOT (p > 0.32)

Discussion

• We replicate findings of phonetic change with web-based design

- Encouraging for future large scale web-based phonetic experiments with simulated interactions

• We find nominal but non-significant change based on feedback

- Suggests possible role of feedback on phonetic change (caveat: contingent on reminding of the data)

- Suggests support for the robust communication view

- No evidence that production ease (latency) affects VOT

- Post-hoc we find significant trial effects

- Suggests phonetic changes may be partly contingent on experience with context

- Production ease accounts do not predict this effect

- Quadratic effect may be due to perceived partner behavior (always correct toward end of study)

Participants

• 60 self reported native speakers of American English

Materials

• 36 critical target words with /k, p, t/ stop onsets with voiced minimal pairs (e.g. pig, stimuli from Kirov & Wilson 2012, study 2)
• Critical target words presented with or without minimal pair (context manipulation, Figures 1a)
• Between participants trials ended with three types of feedback (feedback manipulation, Figures 1c): None, positive only or Mixed

- No feedback – trial ended and next began

- Positive only – partner always picked correct target

- Mixed – partner picked correct on subset of trials (5 critical, 1 filler)

Context manipulation (within participant)

heart pig daft

big pig daft

Figure 1a, preview for target pig with or without competitor.

Figure 1b, target cue in competitor present condition.

Figure 1c: No, positive and mixed feedback (from top to bottom).

Feedback manipulation (between participants)

big pig daft

big pig daft

Figure 2. Post test survey measures.

Figure 3. VOT by condition by feedback. Error bars are ±1 SE after by-participant aggregation. Red lines connect means across contexts.

Figure 4. Effect of speech onset latencies on VOTs. Solid line is linear fit, dashed is non-parametric fit.

Figure 5. VOT across trials by feedback. Solid line is quadratic fit, dashed is non-parametric fit.

Figure 1. Experimental trial timeline (trial preview, target cue, then feedback).