IMPLICIT PHONETIC IMITATION IS CONSTRAINED BY PHONEMIC CONTRAST

Kuniko Y. Nielsen, Department of Linguistics, University of California, Los Angeles

Introduction

The imitation paradigm (Goldinger,1998) has shown that subjects shift their production in the direction of the target, indicating the use of *episodic traces* in speech perception and production. By extending the paradigm to *non-shadowing*, the current study aims to test: 1) if/how this implicit phonetic imitation interacts with linguistic representations when the change might impair linguistic contrast; 2) whether phonetic imitation can be *generalized*, and 3) whether word-level specificity can be obtained through physical measurements of a phonetic feature.

Questions

1. Would the imitation effect be observed when a shift in production would impair a linguistic contrast?

Very short VOT of /p/ introduces linguistic ambiguity (confusion with voiced /b/), while there is no such danger for very long VOT. An extreme episodic view predicts the same imitation whether or not it endangers a contrast

2. Can phonetic imitation effect generalize to sub-lexical and/or sub-phonemic representations?

For example, /p/ and / k/ share the same acoustic cue "Long VOT". Listeners' knowledge of linguistic structure allows (and predicts) sub-phonemic generalization from heard /p/ to unheard /p/ and /k/

3. Is there a word-specific advantage in phonetic imitation?

An episodic view predicts a stronger specificity for 1) more recently experienced words, and 2) low-frequency words (Low frequency = fewer exemplars → weight of one exemplar is relatively larger)

1. **Effect of linguistic contrast**
   (key stimuli: Long vs. Short VOT)
2. **Generalizability of the effect to *new linguistic units***
   (phoneme, feature)
3. ***Effect of word specificity in VOT imitation effect***
   (high vs. low freq, target vs. novel item)

Methods

Participants:

46 native speakers of American English (22M & 24F): 27 in Group1 (long VOT) and 19 in Group 2 (short VOT)

Stimuli

Listening list [120 words] (for listening phase)
- **80 target words with initial /p/ with manipulated VOT**
- 40 filler words with initial sonorants

Production list [150 words] (for baseline and test phase)
- **120 target words**
  1) the 80 modeled words (= **the targets in listening list**)
  2) 20 new words, also with initial **/p/**
  3) 20 new words with initial **/k/,** which like /p/ is [+spread glottis]
- 30 filler words with initial sonorants

Lexical frequency:
40 of the target words had high frequency, and 40 had low. (Kučera & Francis (1967) Hi>50, Low<5: CELEX2 (Baayen, Piepenbrock and Gulikers, 1996) Hi>1000, Low<300)
All new words had low frequency

Phonological neighborhood density & number of syllables: Controlled between frequency groups
(Neighborhood density: from Sommers 2004)

Familiarity: 6.0-7.0 on the 7-point Hoosier Mental Lexicon scale (Nusbaum et al., 1984)

All target words had initial stress, no onset clusters

- A phonetically trained male American English speaker recorded the 120 words in the listening list
- The speaker produced: 1) All the words normally, and 2) The target words *with extra aspiration*
- The VOT for the normally produced initial /p/ was:

Lengthened by 40ms (for **Group 1**)
Spliced with the initial part of hyper-aspirated tokens
➔ average VOT = **113.26 ms** (SD=10.82ms)

Shortened by 40ms (for **Group 2**)
The most stable part of aspiration was taken out:
➔ average VOT = **32.29ms** (SD=12.39ms)

Procedure

The experiment used a slightly modified version of the imitation paradigm from Goldinger(1998). The participants first read the list silently (warm-up phase), to help avoid possible hyper-articulation found in our pilot

1. **Warm-up**: Subjects read the production list *silently*
2. **Baseline**: Subjects produced (read) the production list aloud
3. **Listening**: Subjects heard the listening list (no other task)
4. **Test**: Same as the Baseline Phase

The subjects' tokens were digitally recorded and VOTs were measured using both waveforms and spectrograms.

Results & Discussion

Between group factor:
Listening Stimuli (Long vs. Short VOT)

Within group factors:
Type of Production (Baseline vs. Test)
Lexical Frequency (High vs. Low)
Word Specificity (Target vs. Novel Items)
Imitated Unit (/p/ vs. /k/)

➢**Significant interaction between Type of Production x Listening Stimuli** (*p* <.001*)
= Group Difference

Group1 (Lengthened VOT)

Significant Main Effects on:

➢ **Imitation** (Baseline vs. Test) Longer VOT in the Test phase for *all types of stimuli* (*p* <.005*)
= Generalization

Lexical Frequency (High vs. Low)
Longer VOTs for Low freq (*p* <.05*)

Word-Specificity (target vs. novel)
Longer VOTs for Target (*p* <.01*)

Imitated Unit (/p/ vs. /k/)
Longer VOTs for /k/ (*p* <.01*)

➢ **However, no interactions** of Imitation with other factors**:** Lexical Frequency, Word-Specificity, Imitated Unit

Group 2: (Shortened VOT)

➢ No significant difference between Baseline and Test production across all types of stimuli (*p* >.1)
= No imitation effect

➢Clear lexical frequency effects in Goldinger (1998): Why not in our data?

Post-hoc analysis:

➢Subjects whose baseline/test difference was greater than 5% (16 out of 27 in Group1)
(some subjects showed no imitation at all, confounding possible effects)

➢**Significant interaction between Imitation and Lexical Frequency** (*p* <.05*)
= Frequency Effect on Phonetic Imitation

➢Also, marginal word-specificity effect (Target vs. Novel Items) (*p* =.054)

➢No effect of Imitated Unit (/p/ vs. /k/) (*p* >.1)

Figure 1: Group1 Imitation effect plotted across four types of stimuli (listened to stimuli with LONGER VOT)

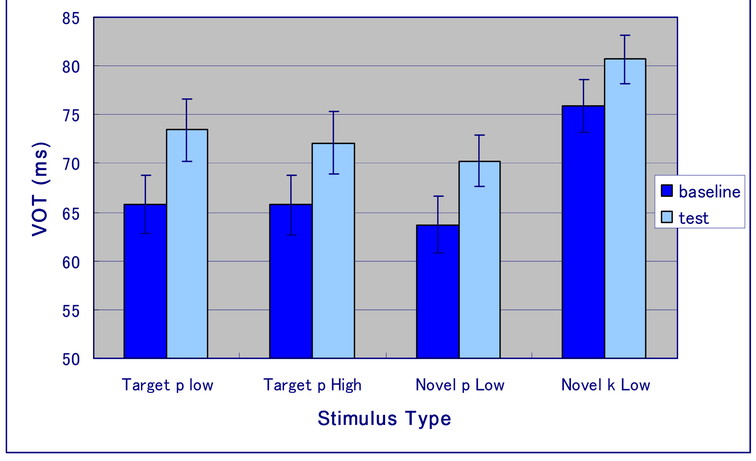
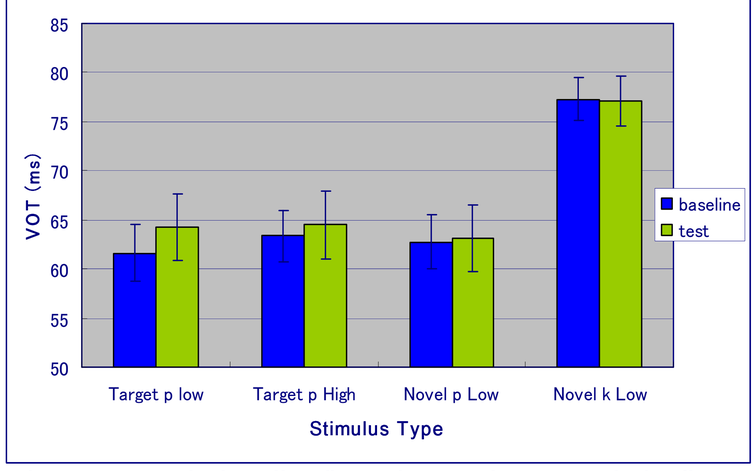


Figure2: Group2 Imitation effect plotted across four types of stimuli (listened to stimuli with SHORTER VOT)



Figures 1 & 2: Imitation effect (in VOT) across four types of stimuli
From left: (1) target words (initial /p/) with low frequency, (2) target words (initial /p/) with high frequency, (3) new words (initial /p/) with low frequency, and (4) new words (initial /k/) with low frequency

➢Subjects imitated long VOT, but did not imitate short VOT

➔**Knowledge of linguistic contrast modulates phonetic imitation**
Similar asymmetrical results in VOT Goodness Rating (Allen & Miller, 2001)

➢ The imitation effect was **generalized** to:
- **New words** which share the initial phoneme /p/
- **New phoneme /k/** which shares a feature [+spread glottis] (and: [-continuant, -sonorant, -voice]) = ***natural class***

➔**Support for sub-phonemic representation**

➢(For the subjects who showed a clear imitation effect,) words with low frequency showed stronger imitation effect than words with high frequency (**frequency effect**), as predicted by the exemplar view

➔**Support for word-level representation**

Conclusions

This study showed:

➢Speakers' sensitivities to linguistic structures:
Phonemic Contrast and
Sub-phonemic Representation

➢Also, some predictions of the extreme exemplar view were confirmed:
Phonetic Imitation and **Word-level Specificity (Lexical Frequency Effect)**

The results call for a linguistically informed exemplar model of speech perception and production, which incorporates both **sub-phonemic and word-level representations** as well as knowledge of **linguistic contrast**