Neutralization avoidance and naturalness in the learning of palatalization

Heng Yin
James White
University College London
Learning bias

- Learners exhibit a variety of learning biases when learning phonological alternations. (see Moreton & Pater, 2012a/b for a summary)

- Most previous studies looking at learning biases focused on non-neutralizing alternations. (e.g., Wilson, 2006; Peperkamp & Dupoux, 2007; White, 2014)

- Neutralizing alternations?
Neutralization avoidance (contrast preservation)

- Previous studies have appealed to neutralization avoidance in analyses of phonological patterns:
  - **Diachronic perspective**: selection process against neutralizing patterns, especially those resulting in ambiguous speech; functional load hypothesis (e.g., Wedel, 2006; Silverman, 2010; Wedel et al., 2013)
  - **Synchronic perspective**:
    - MAXIMIZE CONTRASTS (Flemming, 1996, 2004)
    - *NEUT (Bolognesi, 1998)
    - *MERGE (Padgett, 2003, 2009)
    - PRESERVE CONTRAST (Lubowicz 2007)
Some neutralization avoidance effects

- Preservation of contrasts with phonetic shifts. (Flemming, 1996)
  - English: Difficult to maintain voicing in initial stops, so /b/ → [p], /p/ → [pʰ].

- Rules applying to an isolated sound rather than a natural class of sounds.
  - E.g. stop nasalisation in traditional Tokyo dialect of Japanese: (Labrune 2012)
    - [g] → [ŋ]  
    - *[d] → [n], *[b] → [m] 
    - /n/ and /m/ are phonemes in Japanese while [ŋ] is not.

- Exceptional allomorph selection just in case it would avoid neutralization.
  - Polish (Lubowicz, 2007)

- Optional rules less likely to apply if they might neutralize.
  - Japanese nasal contraction (Kaplan & Muratani, 2015)
A lot of work discussing the role of the speaker and listener in neutralization avoidance. (see Silverman 2012 for an overview)

But not much work looking at the role of the learner.

**Question:** Are learners biased against neutralizing alternations in comparison with non-neutralizing ones?

**Artificial language approach:**
- Present equal input for neutralizing and non-neutralizing alternations.
- Test how well learners acquire the two types of alternation.
Experiment 1: method

- Participants: native English speakers (n=30)

- 3 phases:
  - Exposure phase
  - Test phase 1: trained items
  - Test phase 2: novel items

- 4 novel alternations involving palatalisation [t, d, s, z] ~ [tʃ, dʒ, ʃ, ʒ]

- 2 counterbalancing groups: Language A vs. Language B
Experiment 1: design

Alternations

Language A

\[ / t / \rightarrow [\text{ʃ}] \]
\[ / d / \rightarrow [dʒ] \]
\[ / s / \rightarrow [ʃ] \]
\[ / z / \rightarrow [ʒ] \]

Language B

\[ / t / \rightarrow [\text{ʃ}] \]
\[ / d / \rightarrow [dʒ] \]
\[ / s / \rightarrow [ʃ] \]
\[ / z / \rightarrow [ʒ] \]

Critical non-alternating phonemes

Language A

\[ / \text{ʃ}, \text{dʒ} / \]

Language B

\[ / ʃ, ʒ / \]

Filler non-alternating phonemes

Language A

\[ / p, b, k, g, f, v/ \]

Language B

\[ / p, b, k, g, f, v/ \]
Experiment 1: stimuli

◆ Exposure stimuli: 48 CVCVC singular nonwords with CVCVC-i plural forms:
  ● 8 alternating [t ~ tʃ] and [d ~ dʒ] (Neutralizing in Language A)
      [tusut] → [tusutʃi]
  
  ● 8 alternating [s ~ ʃ] and [z ~ ʒ] (Neutralizing in Language B)
      [duvis] → [duviʃi]
  
  ● 8 critical non-alternating trials ending in [tʃ, dʒ] (Language A) or [ʃ, ʒ] (Language B)
      [buvatʃ] → [buvatʃi]
  
  ● 24 non-alternating filler trials ending in [p, b, k, g, f, v].
      [vatuk] → [vatuki]
Experiment 1: stimuli

◆ ‘Illegal’ sequences never presented.
  ◆ *[ti, di] in Language A.
  ◆ *[si, zi] in Language B.

◆ Otherwise, consonant and vowel distribution roughly balanced across positions.
Experiment 1: exposure phase
Experiment 1: test phases

- 2 test phases: 24 trained items, then 48 untrained items.

- Forced-choice task: choose the correct plural form between an alternating option and a non-alternating option.
Experiment 1: test phase options

Incorrect changing options for non-alternating phonemes:

/ tʃ/ → [ʃ]
/dʒ/ → [ʒ]
/s/ → [tʃ]
/ʒ/ → [dʒ]
/ʃ/ → [tʃ]
/ʒ/ → [dʒ]
/p/ → [tʃ]
/b/ → [dʒ]
/k/ → [tʃ]
/g/ → [dʒ]
/f/ → [ʃ]
/v/ → [ʒ]
Experiment 1: results (Neut. vs. Non-neut. overall)

Better learning in Non-neutralizing condition.
Experiment 1: results (trained and new items)

- Main effect of training: Trained > New
- Main effect of Trial Type: Non-neut. > neut.
- No interaction.
Summary

1. Neutralizing alternations dispreferred relative to Non-neutralizing alternations, despite equal evidence for both.
   • Independent of which alternations were being learned.
A second interesting result: velar vs. labial fillers

More errors on velar fillers than labial fillers!
Summary

1. Neutralizing alternations dispreferred relative to Non-neutralizing alternations, despite equal evidence for both.
   • Independent of which alternations were being learned.

2. Participants spontaneously palatalized velar stops more often than labial stops.
   • Even though neither was palatalized in exposure.
Why velars more than labials?

◆ Substantive naturalness bias?
  ◆ Typologically, palatalization of velars more common than labials. (Kochetov 2011)
  ◆ [ki] and [ʧi] are more phonetically similar than [pi] and [ʧi], and thus considered a more likely alternation by learners. (e.g. Steriade, 2001)
  ◆ Consistent with previous studies (Wilson, 2006; Skoruppa et al., 2011; White, 2014; White & Sundara, 2014)

◆ L1 influence?
  ◆ English has alternations involving velar palatalization (‘velar softening’).
    ◆ *opaque* [ˈou̯ˈpeɪk], *opacity* [ˈou̯ˈpæsɪri]; *esophagus* [əˈsɒfəɡəs], *esophageal* [əˈsɒfeˈdʒiːl] (Halle 2005)
  ◆ Cross-linguistic study would be useful.
Accounting for the anti-neutralization effect

Why are neutralizing rules more difficult to learn?

We consider two possibilities:
  • pure statistical learning
  • learning bias
Just distributional learning?

Non-neutralizing alternation

Neutralizing alternation

Non-alternating fillers

Fully complementary distribution

Partially overlapping distribution

Fully overlapping distribution
Discussion: distributional learning

◆ A distributional learning model must be able to recognise both fully complementary distributions and partially overlapping distributions as cues for alternation. (Calamaro & Jarosz, 2015)

◆ But perhaps identifying partially overlapping distribution requires more input or computation, resulting in reduced learnability.

◆ Potential limitation: Is there more than just phonological distributions involved? Lexical, semantic, and/or pragmatic influences?
  ◆ If so, this is unlikely to be the full story.
Anti-neutralization learning bias

◆ Learners have a neutralization avoidance bias when learning alternations.

◆ Could be formalized in a learning model with weighted constraints (e.g. MaxEnt).
  ◆ A prior could initially assign the ANTI-NEUTRALIZATION constraint a high weight.
    ◆ Would block alternations resulting in neutralization.
  ◆ Exposure to neutralizing alternations would gradually lower the constraint’s weight.
    ◆ Would make neutralizing alternations learnable, but at a slower rate.
What type of constraint/bias?

• Phonological
  – Does it consider all possible forms (e.g., Lubowicz, 2007; Padgett, 2009) or only existing forms in the lexicon?
  – Is it restricted to paradigms, or can it consider any pair of forms? (see Kaplan & Muratani, 2015 for discussion)

• General/functional
  – How do we integrate the neutralization avoidance pressure with models of the grammar (assuming we want to do that)? (e.g. Flemming 1996, 2004)
  – What roles does homophony (actual or potential) or frequency play? (Wedel et al. 2013)

• Upshot: If we see neutralization avoidance effects in learning, learning experiments could provide us with a new way of looking at these questions.
Is neutralization avoidance driven by homophony avoidance?

- Phonologically neutralizing alternations tolerated when they result in little lexical neutralization. (e.g. Silverman 2010 on Korean)

- Diachronically, mergers less likely when they would result in high amounts of homophony. (Wedel et al. 2013)

- Synchronically, stochastic processes may occur less frequently when they result in potential homophones. (e.g. Kaplan & Muratani, 2015 on Japanese nasal contraction)
Does homophony affect the learnability of neutralizing alternations?

• Currently running Experiment 2 to address this question.

Exp. 1: Half lexical neutralization

Exp. 2: Homophony Condition

Exp. 2: No Homophony Condition
Product-oriented learning?

• Product-oriented learning (e.g. Bybee & Slobin, 1982) and neutralization avoidance predict different results in our experiment:
  – **Product-oriented learning:** Adding /tʃ/ → [tʃ] should boost /t/ → [tʃ] due to more cases of plural -tʃi.
  – **Neutralization avoidance:** Adding /tʃ/ → [tʃ] should reduce /t/ → [tʃ].

• We found neutralization avoidance, not product-oriented learning.
  – Fillers further suggest that product-oriented generalization was not a major factor in our study.

• Results diverge from Kapatsinski 2013.
  – He found that adding /tʃ/ → [tʃ] had no impact on (trained) /k/ → [tʃ]. (no neutralization avoidance)
  – But it did increase (untrained) /t/ → [tʃ]. (product-oriented learning)
Conclusions

◆ Learners are biased against neutralizing alternations.
  ◆ Use of neutralization avoidance in synchronic accounts is potentially warranted.
  ◆ Learner could play a role in diachronic neutralization avoidance effects.

◆ Learners favor palatalization of velar stops over that of labial stops.
  ◆ Suggests a substantive bias, though a L1 bias cannot be ruled out at in this experiment.
Future directions

◆ Running the study with different L1 backgrounds.

◆ Would we see the same effect with infants?
  ◆ Perhaps related to the Mutual Exclusivity Bias in pragmatic development? (Merriman, Bowman, & MacWhinney, 1989)
    ● Infants disfavour the learning of homophones.
    ● Neutralization avoidance could be a realization of mutual exclusivity at a phonological level.
    ● Infants have ‘proto-lexical’ knowledge (Ngon et al., 2013) and phonotactic generalizations (Mattys et al. 1999) at an early age.

◆ L2 acquisition of natural languages.
  ● L2 learners care more about semantic information than ALL participants.
  ● More explicit training on phonological system.
Acknowledgements:

◆ Help with experiment: Andrew Clark.

◆ Helpful discussion: Wing Yee Chow, Richard Breheny.

◆ London Phonology Seminar audience.
References:


References:


References:


