Effects of perceptual phonetic training on the perception and production of second language syllable structure

Amanda Huesch & Annie Tramblay (2015)

Presented by Dong-Jin Shin

Overview

1. Introduction
2. Relationship bw L2 perception and L2 production
   a. Speech Learning Model (SLM)
   b. Perceptual Assimilation Model (PAM) –L2
3. L2 perception and production of syllable structure
4. Methods
5. Discussion and Conclusions

1. Introduction

- Investigate to which degree the L2 speech perception and production systems are linked;
  - L2 learners' speech production errors may stem from **perceptual errors** (e.g., Brannen, 2002; Broselow, 2009; Flege, 1995; Flege, MacKay, & Meador, 1999; Peperkamp & Dupoux, 2003; Rochet, 1995)
  - L2 speech production errors are driven by the **non-native-like timing of gestures** (e.g., Colantoni & Steele, 2008; Davidson, 2010; Davidson & Stone, 2003; Oh, 2008; Zsiga, 2003)

1. Introduction

- Two L2 speech perception theories
  1. **Speech Learning Model (SLM)** (Flege, 1991, 1995, 2003)
  - L1 native categories are not fixed and can change over time
  - Accurate perception is the prerequisite of the accurate production
  - But, accurate perception does not guarantee accurate production
  - No one-to-one relationship bw production and perception
  2. **Perceptual Assimilation Model (PAM-L2)** (Best&Taylor 2007)
  - Takes articulatory gestures as the direct primitives of speech perception
  - Posits the direct link bw the production and perception

2. Relationship between L2 perception and L2 production

- Non-native speakers may learn to produce L2 segments accurately without being able to perceive them accurately.
  - Japanese L2 learners of English could produce the English /ɹ/-/l/ contrast accurately despite having poor perception of it (Goto, 1971, Sheldon & Strange, 1982)

2. Relationship between L2 perception and L2 production

- Accurate perception may be a necessary precursor of accurate production.
  - Rochet (1995):
    - Portuguese L2 learners of French produced /y/ as an/i/-like vowel,
    - English speakers more often produced /y/as an/u/-like vowel;
    - These results paralleled findings from his perception task,
    - /y/ vowels were categorized more often as /i/ by Portuguese speakers and /u/ by English speakers.

2. Relationship between L2 perception and L2 production

- Some evidence bw production and perception (Evan & Iverson, 2007)
  - Examined 11 English vowels of speakers from Northern England attending university in South England.
  - No significant interaction between production and perception, BUT;
    - Those who produced more southern perceived southern vowels better
    - Those who produced more northern vowels produced northern vowels better

2. Relationship between L2 perception and L2 production

- A more robust test of the relationship between perception and production would instead be to investigate how learning in one skill (e.g., perception) affects learning in the other (e.g., production).
  - **High-Variability Phonetic Training (HVPT)**
  - HVPT can modify L2 learners' perceptual representations

High-Variability Phonetic Training (HVPT)

- Bradlow et al. (1997)
  - Investigated the effects of perceptual training on the production of /ɹ/ and/l/in Japanese L2 learners of English.
  - The trained learners‘ perception accuracy and production scores were significantly higher than untrained learners.
  - Substantial variability in production among the trained learners
  - No correlation between perception and production
- Different participants may learn the required motor commands at different rates
- Timing of the articulatory commands is somewhat independent of the creation of perceptual representations.

3. L2 perception and L2 production of syllable structure

- L1 syllable structure influence on L2 speech perception
  - L1 Japanese learners of English (Dupoux et al., 1999)
    - Native speakers of Japanese were much more likely to hear an “illusory” vowel between two word-medial voiced obstruents (e.g., ebzo) than native speakers of French.
  - L1 Korean learners of English (Kabak & Idsardi, 2007)
    - Syllable structure violation causes an illusory vowel effect in Korean L2 learners of English

3. L2 perception and L2 production of syllable structure

- L1 phonetic cue influences on the L2 perception (Dupoux et al., 2011)
  - Japanese and Brazilian Portuguese listeners were more likely to hear an illusory vowel between word-medial consecutive stops than were European Portuguese listeners.
  - Most of the stimuli contained sequences of voiced stops, which can occur at the surface level in European Portuguese (via vowel deletion) but not in Japanese or Brazilian Portuguese.

3. L2 perception and L2 production of syllable structure

- Studies suggest that L1 production errors stem from
  - Sonority Hierarchy (e.g., Broselow & Finer, 1991; Eckman & Iverson, 1993; Berent, Steriade, Lennertz, & Vaknin, 2007).
  - Inaccurate timing of overlapping articulatory gestures (Davidson, 2006 & 2010)

The current study

1. What the relationship is between L2 perception and production of vowel epenthesis:
   - If epenthesis errors are due to the mistiming of articulatory gestures
     - Then, perception and production would be **correlated**
   - if L2 learners' speech perception is not rooted in such gestures
     - Then, **no direct relationship** bw perception and production
2. Examined the effect of HVPT on Korean speakers' L2 syllable structure (e.g., English coda consonants)

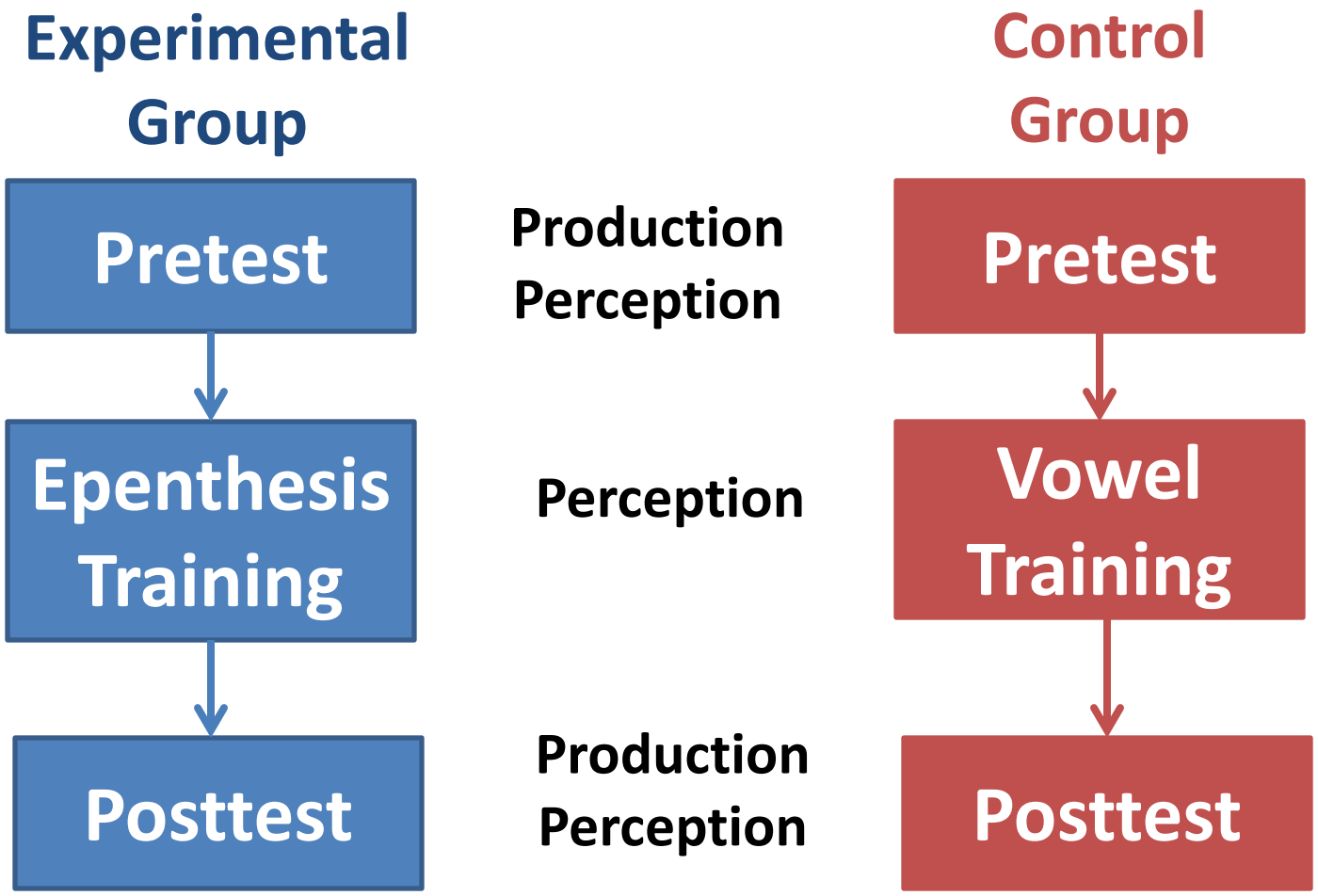
4. Methods : Participants

- 24 adult Korean learners of English
  - Randomly assigned to two groups.
  - **The experimental group** (n = 12, 5 women)
    - perceptual training on palatal codas,
  - **The control group** (n = 12, 5women)
    - perceptual training on vowel pairs (no words in the control training contained palatal codas).

4. Methods: Materials

- 48 minimal pairs of natural tokens (half real, half nonce word)
  - CVC vs CVCy (e.g. push/pushy vs. mish/mishy)
  - 24 pairs were included in the training
  - Other 24 pairs were only used in pre and post tests
- Minimal pairs were presented in isolation as well as within the carrier sentences
  - (He said ___ angrily vs. He said ___ frequently)
- All of the stimuli were recorded by six native speakers of English (3 women)

4. Methods: Procedures



4. Methods: Procedures

- Pretest phase
  - The pretest phase consisted of both perception and production experiments
  - The perception tests included a forced-choice word-identification task of words in isolation and words in carrier phrases.
  - 96 experimental words(from the 48 minimal pairs) and 56 filler words(from the 28 minimal pairs)
  - For production test, participants received a visual word or sentence prompt and read the word or sentence

4. Methods: Procedures

- Experimental training phase
  - eight, 20-min daily sessions of online training on palatal coda
  - Two blocks (isolation, carrier sentence)

    (a) 48 words in isolation from one talker along with16 fillers

    (b) 96 words in each of the carrier contexts from one talker along with 32 fillers

4. Methods: Procedures

- Control training phase
  - Training on the three vowel pairs(/æ/-/ɛ/, /i/-/ɪ/, and/oʊ/-/ʊ/),
  - The stimuli contained ten consonants(/d t n b p m k g h s/)in on set and/or coda position.
  - Talkers were eight native speakers (4 women)
    - Speech rate varied (slow/careful, normal/casual, fast)
    - participants heard the word pairs in a self-paced exposure phase and were tested on the word pairs in a subsequent testing phase.
    - 20 min training

4. Methods: Procedures

- Post-test Phase
  - The post-test phase was identical to the pretest phase

4. Methods : Analysis

- Data analysis
  - Answers on the perception tests were scored as either accurate or inaccurate (Transformed into *d'*)
  - Production responses were rated by 67 English native listeners (NLs) in two listening tasks
    - A paired-comparison task & a forced choice word identification choice
    - 43 of them completed both the forced-choice word-identification and paired-comparison tasks.
    - The remaining 24 listeners completed only one of the two tasks.

4. Methods: Analysis

- Paired-comparison task
  - a group of English NLs performed a paired-comparison task with the learners' pretest and post- test productions.
    - ‘1’ = the first version was better than the second
    - ‘4’ = no noticeable differences between the two versions
    - ‘7’ = the second version was better than the first
    - NLs used all seven points on the rating scale
  - Scores were converted from a scale of 1–7 to a scale of −3 to 3
    - A negative score indicated a preference for the pre-test item
    - A positive score indicated a preference for a post-test item.

4. Methods: Analysis

- Forced-choice word-identification task
  - Learners' productions were also presented to NLs in a forced-choice word-identification task.
    - NLs saw the two words from each pair (e.g., push-pushy) presented on the left and right side of the screen.
    - Then, a version of the word was played and NLs were asked to choose the correct response.
    - NLs heard both the experimental (48 minimal pairs) and filler (28 minimal pairs) stimuli.

5. Results

- Perception results
  - Five native English listeners completed the perception pretest.
    - d′ score
      - real words = 4.19 (SD = 0.49); nonce words = 4.42 (SD = 0.33).
      - A paired-samples t-test revealed no significant difference between real and nonce-words $t(4)=-1.58$, $p<0.190$
      - The task was therefore adequate to test for the perception of palatal codas.

5. Results

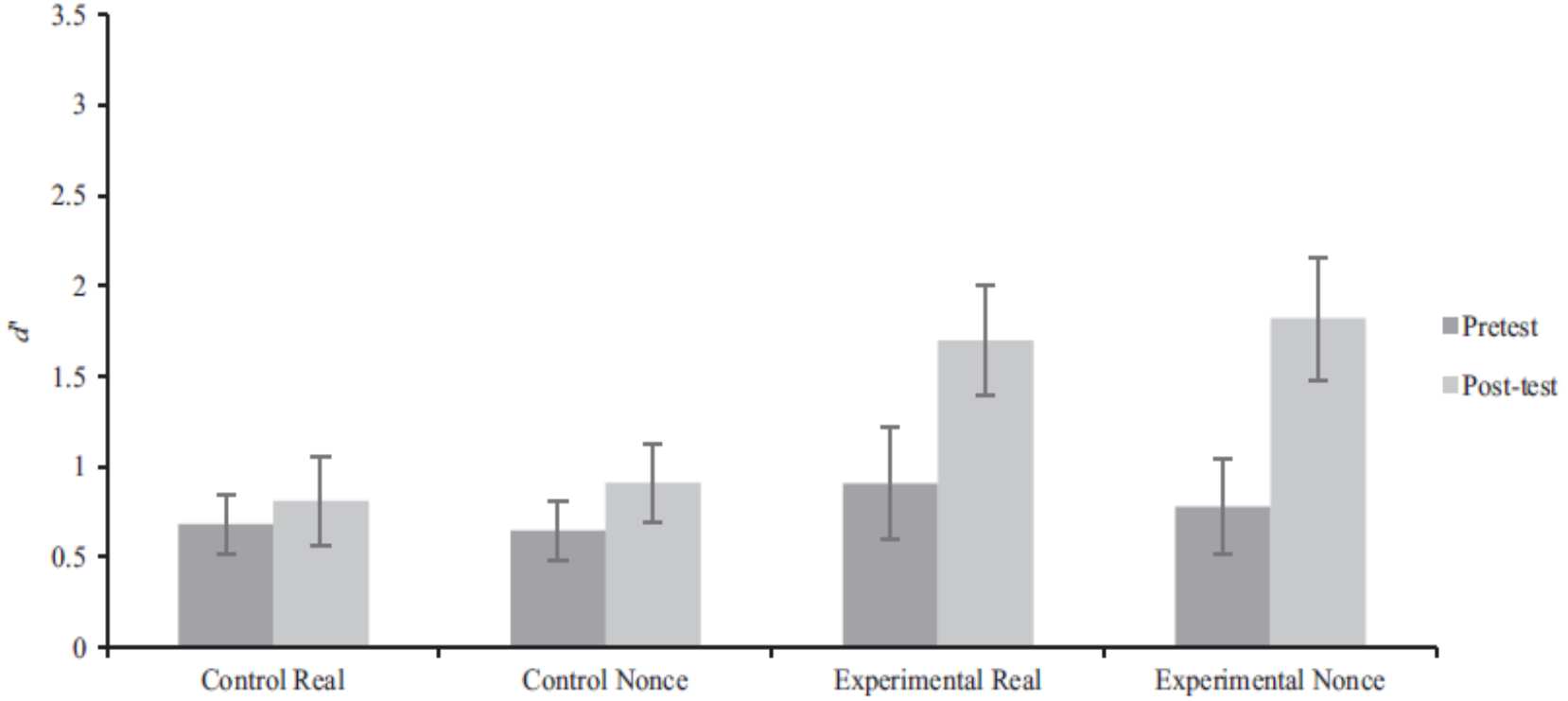


Fig. 1. Pretest and post-test mean *d'* perception scores (and standard error bars) by group and word type.

5. Results

- A mixed-design repeated-measures ANOVA was performed with test (pretest, post-test) and word type (real, nonce) as within- subject variables and with group (experimental, control) as between-subject variable.
- A main effect of test ($p<0.001$).
- An significant interaction between test and group ($p< 0.006$)
- Paired-samples t-tests on each group's pretest and post-test d’ scores a sig difference bw pre-post for the experimental group only ($p<0.001$).
- **Thus, the experimental group showed significant improvement between the pretest and post-test for palatal codas in sentences, but the control group did not.**

5. Results

- Investigated the relationship between the individual participants' perception improvement scores and the amount of time they spent on perceptual training.
  - **No significant correlation**
  - less time spent on perceptual training is not related to perception accuracy.

5. Results

- The experimental group's pretest and post-test d’ scores on four categories of words:
  - (1) new words spoken by new talkers
  - (2) new words spoken by talkers from the training
  - (3) words from the training spoken by new talkers
  - (4) words from the training spoken by talkers from the training.

5. Results

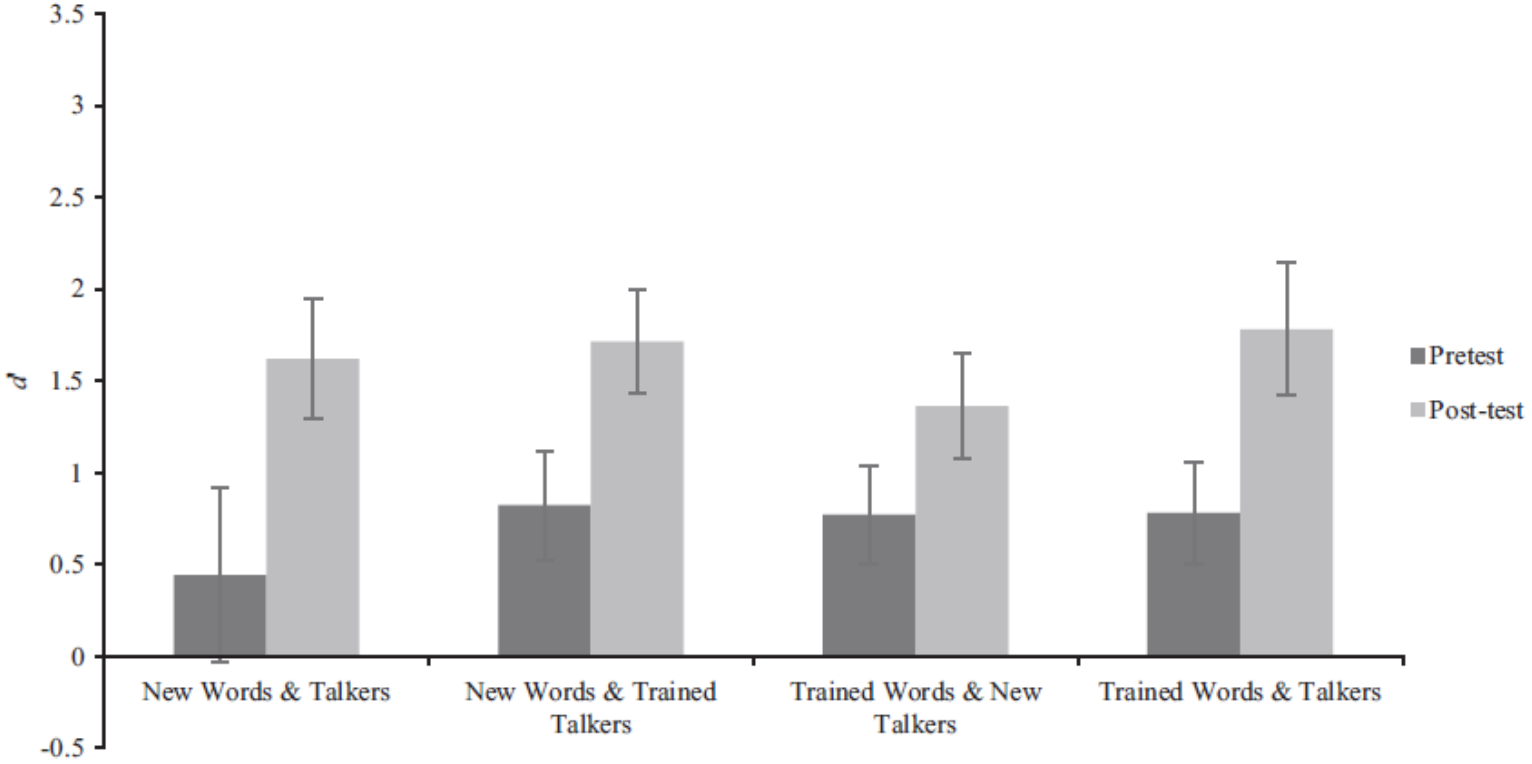


Fig. 2. Pretest and post-test mean *d'* perception scores (and standard error bars) for the experimental group with new and trained words and talkers.

5. Results

- **A repeated-measures ANOVA** performed with test(pretest, post-test), word(new, trained),and talker(new, trained) as within- subject variables revealed
  - a main effect of test ($p<0.001$) only
- learners showed similar improvements between pretests and post-tests on all four categories of words.
- Participants were able to generalize learning to new words and new talkers.
- **Summary : Training worked well in the sentence context and subjects generalized what they learned through the training sessions.**

5. Results

- Production results: paired-comparison task

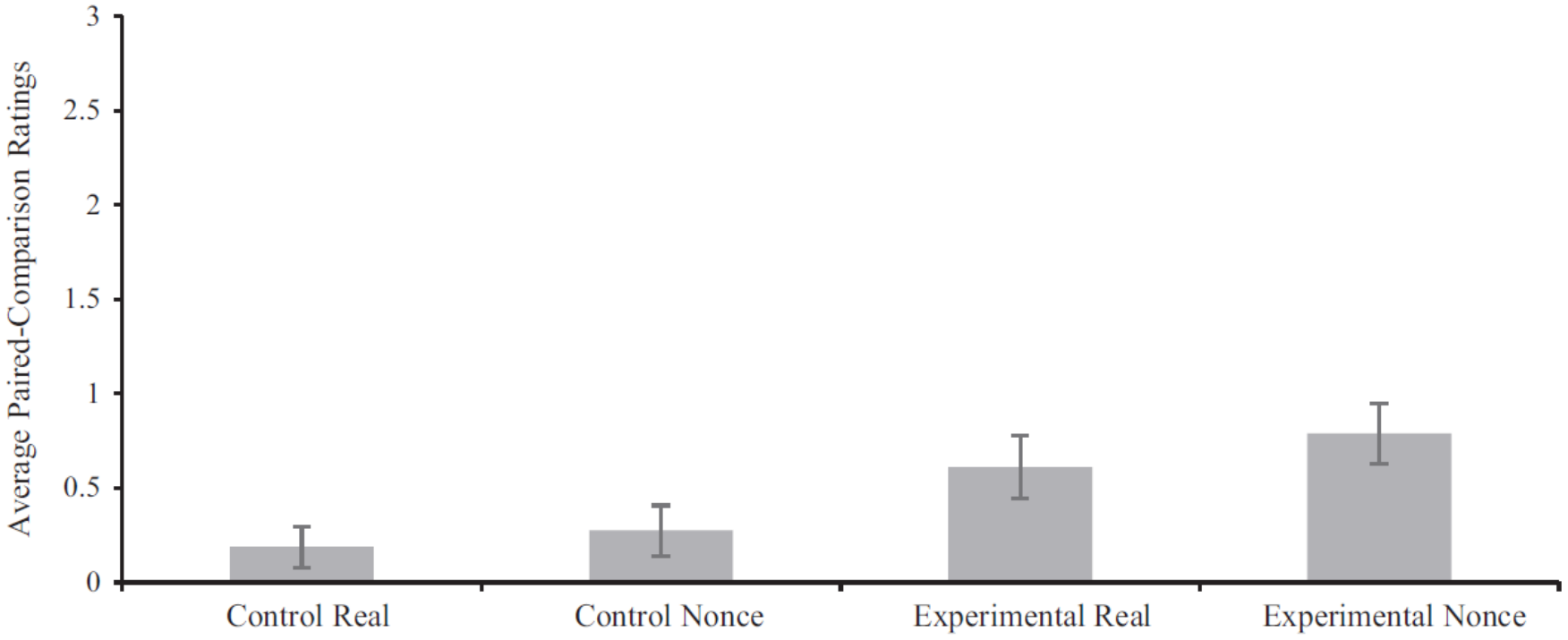


Fig. 3. Mean paired-comparison ratings (and standard error bars) by group and word type.

5. Results

- A mixed-design repeated-measures ANOVA with word type (real, nonce) as within-subject variable and group (experimental, control) as between-subject variable
  - Main effect of word type ($p<0.005$)
  - Main effect of group ($p<0.027$)
  - No interaction between type and group
- More of the experimental group's post-test productions of palatal codas in the sentence context were rated more native-like in comparison to those of the control group.

5. Results

- The perceptual phonetic training on palatal codas helped learners improve their production of palatals
- The improvements were minimal for the control group.
- These results also show that improvements were greater for nonce words than real words.
  - NLs may have been more lenient in their judgments of nonce words as opposed to real words because for the former there was no top down information to influence their decisions.
- Averages for the control group are above zero, indicating a slight preference for post-test productions.

5. Results

- Forced-choice word-identification task

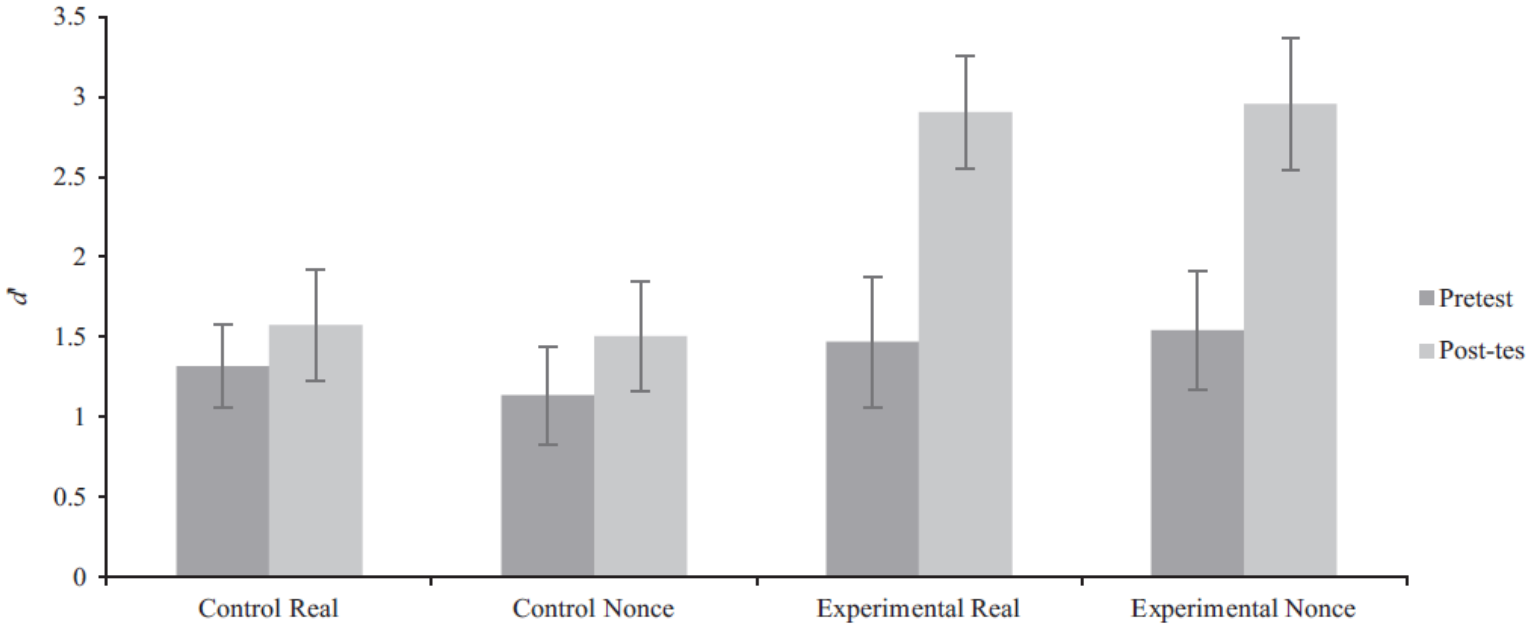


Fig. 4. Pretest and post-test mean *d'* production scores (and standard error bars) from the word-identification task by group and word type.

5. Results

- A mixed-design repeated-measures ANOVA was performed with test (pretest, post-test) and word type (real, nonce) as within- subject variables and with group (experimental, control) as between-subject variable.
  - A main effect of test ($p<0.001$)
  - An interaction between test and group ($p<0.014$)
  - A significant difference between the pretest and post-test scores for the experimental group ($p<0.002$); but not for the control group ($p<0.186$)
  - For the experimental group, NLs were able to accurately identify more post-test productions than pretest productions.

5. Results

- Forced-choice word-identification task summary: phonetic training on palatal codas enhanced the learners‘ production of codas, as reflected in NLs‘ identification of the produced words.

5. Results

- Individual variability and the relationship between perception and production

Table 2
Pretest, post-test, and improvement scores (as *d'* scores) by participant.

	Palatal Codas: Sentences					
	Perception			Production		
Participant	Pretest	Post-test	Improvement	Pretest	Post-test	Improvement
P1	0.21	0.84	0.63	0.60	3.53	2.93
P2	0.16	2.11	1.95	2.45	3.41	0.96
P3	0.24	1.53	1.29	0.15	3.37	3.22
P4	0.37	1.66	1.30	1.58	3.27	1.70
P5	0.48	1.19	0.71	0.66	4.06	3.41
P6	0.06	0.06	0.01	0.06	0.59	0.53
P7	0.72	2.13	1.41	1.87	2.79	0.92
P8	−0.22	0.00	0.22	−1.30[a]	0.07	1.37
P9	1.38	3.06	1.68	3.27	4.09	0.82
P10	2.72	2.66	−0.06	2.54	2.80	0.25
P11	1.29	2.12	0.83	2.06	3.78	1.72
P12	2.33	3.48	1.15	3.37	2.86	−0.51
Mean	**0.81**	**1.74**	**0.93**	**1.44**	**2.89**	**1.44**

[a] The large negative *d'* score of Participant 8's perception pretest can be explained by the fact that this participant *almost* always produced CVCy and CVC sequences as CVC. If it were the case that the participant *only* produced CVC sequences, the *d'* would be 0 (no discrimination). However, because this participant showed some variation (specifically, she produced one CVC sequence as CVCy and one CVCy sequence as CVCy), this resulted in the large negative value. While large negative *d'* values might be explained by a participant purposefully giving the wrong answer (cf. Park, 2013), we do not believe this is the case with this particular participant.

5. Results

- In terms of production and perception correlation;
  - No significant correlation bw the perception pretest scores and perception improvement scores ($p<0.788$)
  - No significant correlation between the production pretest scores and production improvement scores ($p<0.068$)
- **No significant correlation bw the improvement scores on the perception task and those on the production task (word-identification measure) ($p<0.929$).**

5. Results

- Thus, despite finding that perceptual training contributed to enhancing both the perception and production of palatal codas in Korean L2 learners of English…
  - No direct relationship between the production and perception.

6. Discussion: first finding

- The beneficial effects of perceptual training can extend to syllable structure.
  - The experimental group improved their perception of palatal codas in English
  - Generalized improvements to both new words and new talkers
  - Improvements in perception were found for both real and nonce words.
  - Perceptual phonetic training is beneficial for establishing new segmental categories.

6. Discussion: second finding

- Perceptual training on palatal codas enhanced learners' **production** of palatal codas.
  - Perceptual training on difficult segments can yield production improvements for the same segments
    - Production and perception is related somehow
    - **No correlation** is found in current study
- Perception and production do not bear one-to-one relationship.

6. Discussion

- The representations underlying learners' speech perception and speech production differ.
- The results are more easily explained by the model situated within **a psychoacoustic theory** of speech perception(e.g. **SLM**) rather than Direct Realistic theory (e.g., PAM-L2)

6. Discussion

- Need more time to learn the motor skills
- Differences in cue weighting
  - Some subjects focused on perceptual cues and others concentrated on cues highly related to production.
- The difference explained in Bradlow et al. (1997)
  - The contrast pairs differed more in the timing of the articulatory gestures than in their place of articulation.
  - In order to develop native-like perception of palatal codas, learners must attend to the cues that signal these timing differences.
  - These cues include the duration of the stem vowel and the duration of the syllable coda/onset consonant.

6. Discussion

- The individual L2 learners improved perception and production at different rates.
  - Some learners may thus have tuned more to the perceptual cues to the realization of palatal codas in English.
  - Others may have tuned more to the mapping between these cues and the timing of articulatory gestures in speech production.

6.Discussion

- Open questions
  - How the different perceptual cues to palatal codas in English influence Korean listeners‘ perception of these codas
  - The present study did not focus on these individual cues
  - Whether the cues that enhance learners' perception of syllable structure after perceptual training are also produced after such a training
  - whether similar results would be obtained with participants at lower proficiencies

7. Conclusion

- The current research contributes to a better understanding of the relationship between perception and production systems.
  - A perceptual training beyond isolated words, finding improvements in the context sentences
  - Perceptual phonetic training can be beneficial not only for acquiring new segment contrasts, but also for acquiring segments in restricted syllable structures.
  - Representations are not directly shared between perception and production systems.

Limitations

- Duration of training
- Number of subjects
- Control group training?

Thank you!