Effects of Production Training and Perception Training on Lexical Tone Perception

Shuang Lu, Ratree Wayland & Edith Kaan

Department of Linguistics
University of Florida
Introduction

• The relationship between perception and production has been debated for a long time

  o Liberman and Mattingly (1985): The Motor Theory of speech perception
    - Speech perception and speech production share the same neural network
    - Speech is perceived by recognizing the intended vocal articulatory gestures
    - Speech perception and production are linked through an innate mechanism
    - Several behavioral and neuropsychological studies have supported these two assertions (e.g. Ojemann, 1983; Adank et al., 2010; see Galantucci et al., 2006 for a review)

  o Speech production relies on speech perception (e.g. Guenther et al., 2006; Levelt et al., 1998; Flege, 1995)
Introduction

- Perception training vs. Production training (novel phonological categories)
  - Perception training
    - Short-term laboratory perception training improves participants’ discrimination/identification of novel phonemic contrasts (e.g. Jamieson & Morosan, 1986; Logan et al., 1991)
    - Improvement gained through the perception training transferred to production learning (e.g. Bradlow et al., 1997)
Introduction

- Perception training vs. Production training (novel phonological categories)
  - Production training (mixed results)
    - Production training improves participants’ *production and perception* of novel phonological contrasts (e.g. Hirata, 2004; Leach & Samuel, 2007)
    - Production training only improves participants’ production but not perception (e.g. Hattori, 2009)
    - Production training hinders perception learning; perception training transferred more effectively (e.g. Herd, 2011; Baese-Berk, 2010)
Gaps in the literature

- Very limited studies have compared the effectiveness of production training and perception on the perception of suprasegmental aspect of speech, e.g. *lexical tones*

- No ERP study has specifically investigated the effect of production training on lexical tone perception
Background

- **Tone in language**
  - In English (a non-tonal language), tone is only used to indicate sentence type
    - e.g. declarative vs. interrogative
  - Lexical tones: form of prosody used to create semantic contrast between otherwise phonologically identical words
    - [ma] with 4 Mandarin tones
      - Tone1: ‘mother’
      - Tone2: ‘hemp’
      - Tone3: ‘horse’
      - Tone4: ‘scorn’

Xu (1997)
Background

- Non-native listeners of tone languages have difficulties in both comprehending and producing lexical tones
  - Differences in perceptual dimensions (onset vs. pitch contour)
  - Interference from L1 (English) suprasegmental features

- Both perception and production training improve the comprehension and production of lexical tones by non-native speakers (e.g. Wang et al., 1999 & 2003; Leather, 1990)
  - No comparative study between two types of training
Research Questions

- Is production training more effective than perception training in facilitating lexical tone perception?
  - Discrimination paradigm
  - Identification paradigm
Research Questions

- Does production training affect the unintentional perception of lexical tones differently from perception training?
  - ERP data – passive oddball paradigm
    - Mismatch Negativity (MMN): negative wave elicited by auditory stimuli that are infrequently presented and deviate from a frequently presented standard stimuli: $X X X X Y X X X X X$

Sams et al. (1985)
Overview

- **Experiment 1**: Behavioral and ERP study - effects of laboratory production and perception training on lexical tone perception
  - discrimination training paradigm
  - Behavioral & ERP: no difference

- **Experiment 2**: Behavioral only
  - identification training paradigm
  - Some hint that Perception-plus-production training outperformed Perception-only training
Overview

• Experiment 1: Behavioral and ERP study - effects of laboratory production and perception training on lexical tone perception

  ○ Day1: baseline
    - ERP & behavioral

  ○ Day 2: Training session:
    - perception-only vs. perception-plus-production training

  ○ Day3: post-training tasks
Methods

Participants

- 22 native speakers of English
  - 11 Perception-only group
  - 11 Perception-plus-production group

- None of the English participants were familiar with any tone language or have received any formal music training

- Right-handed and healthy participants with no history of speech/neurological impairment

  - Hearing test
  - Music aptitude test (AMMA) (p > .50)
  - Memory test (p > .18)
Methods

- Stimuli
  - Syllables [pʰa], [kʰo], [ʈʰa], [ʈʰi], [ʈʰɛ] and [ʈʰo] in pre- & post-training
  - Syllables [pʰa], [pʰi], [kʰɛ] and [kʰo] in training
  - ERP: [tʰu]
    - All of the syllables exist in American English
    - The syllables were produced by two female native speakers of American English
    - Associated with three linear tones that resemble Mandarin Tone1 (high-level), Tone2 (high-rising) and Tone4 (high falling)
Methods

Stimuli

- Three linear pitch contours were superimposed starting from the voiced part of each syllable
  - ERP: pitch contour started at 108ms after the stimulus onset; mean duration of all tokens was 553 ms
- Judged by native Mandarin speakers as perceptually natural
Methods

- Procedures
  - Day 1
    - ERP recording (baseline)
      - passive oddball paradigm: watch a silent movie while the stimuli are presented
      - 39 Ag/AgCl electrodes will be recorded
      - 6 blocks
        - Deviants T2 & T4 in standard T1
        - Deviants T1 & T4 in standard T2
        - Deviants T1 & T2 in standard T4
Methods

Procedures

Day 1

- Behavioral task: discrimination task
  - All possible pairings of tone within each syllable (288 trials) were presented through headphones
    - Same speaker within each pair
    - Inter-stimulus-interval: 500ms
  - Participants had to decide whether the tones in each pair were the same or different
Methods

Same or Different?
Methods

Same or Different?
Methods

- Procedures
  - Day 2: 1 hour training

  *Perception-only training:*

  (1) 500 ms ISI  (2) (3) (4) (5) (6) (7) (8)
  Stimulus1 → Stimulus2 → Same or Different → feedback → Stimulus1 again → Say ‘Next’ → Stimulus2 again → Say ‘Next’
  with the tone and the tone graph e.g. Tone1

  *Perception-plus-production training:*

  (1) 500 ms ISI  (2) (3) (4) (5) (6) (7) (8)
  Stimulus1 → Stimulus2 → Same or Different → feedback → Stimulus1 again → Imitate Stimulus1 → Stimulus2 again → Imitate Stimulus2
  with the tone and the tone graph e.g. Tone2

  - Day 3: same as Day 1
Results

- Behavioral

- Both groups improved in the post-training tasks (ps < .001)
- No difference between the two groups
Results

- Behavioral

  - No improvement in terms of RT
Results

- ERP (Grand Average)

  - MMN: smaller in the post-training task than in the pre-training task ($p = .018$).
Results

- ERP
- Difference between deviant and standard
Results

- ERP (difference between deviant and standard)
  - The MMN was significant only for T21, T41 & T42

![Graph showing ERP waveforms and MMN](image)
Results

- **ERP**
  - for the *T41 condition*, the two groups differed in the lateralization of the MMN component
    - perception-only group showing a larger MMN at the left-frontal electrodes
    - the perception-plus-production group displaying a larger MMN at the right-frontal electrodes.
To sum up

- **Discrimination training paradigm** - Effects of Laboratory Perception training and Production Training on Tone Perception

  - The additional production component in the perception-plus-production training did not result in more improvement in the tone discrimination at the intentional & unintentional level.

  - Both groups showed a smaller MMN in the post-training task than in the pre-training task.
    - Differences in pitch onset (Kaan et al., 2008)
    - More accurate if first token had a lower F0 than when the first token had a higher F0 (Francis & Ciocca, 2003)
To sum up

- **Future study**
  - Longer training
  - Include feedback in production training
  - Further analysis for individual differences (fast learners vs. slow learners)
Overview

- **Experiment 1: Behavioral and ERP study - effects of laboratory production and perception training on lexical tone perception**
  - discrimination training paradigm
  - Behavioral & ERP no difference

- **Experiment 2: Behavioral only**
  - identification training paradigm
  - Some hint that Perception-plus-production training outperformed Perception-only training
Overview

- **Experiment 1**: Behavioral and ERP study - effects of laboratory production and perception training on lexical tone perception
  - **Day 1**
    - baseline
    - Training session: identification-only vs. identification-plus-imitation training
  - **Day 2**
    - post-training task
Methods

- **Participants**
  - 20 native speakers of English
    - 10 identification-only group
    - 10 identification-plus-imitation group

- **Stimuli**
  - \([k\text{\varepsilon}], [k\text{ho}], [p\text{ha}], [p\text{hi}], [t\text{ha}], [t\text{i}], [t\text{\varepsilon}]\) and \([t\text{ho}]\) in pre- & post-training
  - \([k\text{ha}], [k\text{\varepsilon}], [k\text{i}], [k\text{ho}], [p\text{ha}], [p\text{\varepsilon}], [p\text{hi}]\) and \([p\text{ho}]\) in training
Methods

- Procedures
  - Day 1
    - Identification task (baseline)
Methods

- Procedures
  - **Day 1**
    - Identification task (baseline)
    - One-hour training
Methods

- Procedures
  - Day 2
    - Post-training identification task
Results

- Identification

- Both groups improved in the post-training tasks ($ps < .001$)
- No difference between the two groups
The pre-training difference between the two groups was not significant ($p > .25$).

The identification-plus-imitation group identified the tones more quickly in the post-training task than in the pre-training task.

The identification-only group did not show any improvement.
To sum up

- **Identification training paradigm** - Effects of Laboratory Perception training and Production Training on Tone Perception
  - The identification-plus-imitation training was more effective to improve the tone identification than the identification-only training
Discussion

- **Effects of Laboratory Perception training and Production Training on Tone Perception**
  - **Discrimination training paradigm.**
    - The additional production component in the perception-plus-production training did not result in more improvement in the tone discrimination at the intentional & unintentional level.
  - **Identification training paradigm**
    - The identification-plus-imitation group showed more improvement
      - Even though the experimental design was biased toward the identification-only group
        - The identification-only training reinforced the tone types of the stimuli during the production process while the identification-plus-imitation training directed the participants to pay more attention to the sounds rather than the tone types.
Why did different training paradigms present different results?

- The categorization ability might be crucial for production learning to be able to transfer to the perceptual modality

  - Discrimination training paradigm.

    - Encourages participants to pay more attention to within-category acoustic differences that are irrelevant to the key identify of the categories
    - The participants did not gain enough categorization skills through the discrimination training
    - Therefore, the transfer from the production learning to the perceptual learning did not take place
Discussion

- Why did different training paradigms present different results?
  - The categorization ability might be crucial for production learning to be able to transfer to the perceptual modality
    - Identification training paradigm
      - Encourages participants to form distinct categories and to focus more on the differences between categories
      - Participants gained enough categorization skills, which enable the production learning to take place and in turn to facilitate the perceptual learning.
Discussion

- Relationship between speech perception and production?
  - The categorization skill might be the prerequisite for the occurrence of the production learning, which in turn may facilitate the perceptual learning.
  
  So...speech production may rely on speech perception since the categorical learning precedes the production learning, which was contrary to the Motor theory.
Future research

- Compare the effect of perceptual learning on speech production & the effect of production learning on perceptual learning in order to uncover the relationship between speech perception and production.
Both music and language use time-varying pitch to convey information

- Tone in music
  - Used to differentiate notes
To what extent does processing in one cognitive domain affect processing in another?

- Music training positively affects the perception of pitch variations in music & language
  - Perception
    - Lee and Hung (2008): Musicians identified Mandarin tones more accurately than non-musicians
    - Wong et al. (2007): Pitch tracking in the brainstem was more faithful and robust in musicians than in non-musicians
    - Chandrasekaran et al. (2009): Musicians were more sensitive to lexical tone contrasts than non-musicians
  - Production
    - Gottfried (2007): Musicians’ spoken imitation of Mandarin tones were rated as significantly more native-like than those of non-musicians
Comprehensive studies on trained instrumental musicians, relatively little is known about formally trained vocal musicians and the differences between vocalists and instrumentalists.

- Conflicting research on non-linguistic pitch processing
  - Zwissler (1971): accurate singers performed significantly better on auditory pitch discrimination task than the inaccurate singers (support motor theory) – non-linguistic pitch stimuli
Research Questions

- Are the effects of production training and perception training domain general or domain specific?
- Is production training more effective than perception training in facilitating lexical tone perception?
  - Specifically, investigate whether there is a difference between different types of musicians (instrumentalists and vocalists) with regard to lexical tone perception/production ability
    - Vocalists expected to outperform instrumentalists on production, but not necessarily perception, as per the Motor Theory
Overview

- Experiment 3: behavioral study – effects of music perception and production training on lexical tone perception and production
Methods

- **Participants**
  - 15 English vocalists
  - 15 English instrumentalists
  - 15 English non-musicians
  - 15 Chinese non-musicians

- **Criteria for musicians:**
  - All musicians had at least 5 years of continuous music training
  - were still receiving music training at the time of participation
  - No pure percussionists
  - Vocalists had vocal training in addition to instrumental training, predominately piano; instrumentalists had diverse instrumental backgrounds

- All non-musicians had less than 2 years of music experience and had not received any music training within the past 4 years

- Native English speakers had no prior experience with tonal languages
Methods

- Stimuli consisted of six syllables with each of the four mandarin tones
  - [mao], [mi], [lao], [li], [tʰao], [tʰi]
- Stimuli were produced by two female native speakers
  - Stimuli were excised from the phrase: *Qing shuo TARGET zhege zi*
    - ‘please say TARGET this word’
- Five native Chinese speakers judged the stimuli to be 100% natural Mandarin stimuli
Methods

- Procedure
  - Discrimination task
  - Imitation task
**Methods**

- **Perceptual discrimination task**
  - 296 pairs of natural Mandarin single-syllable words were played through headphones
  - All possible pairings of tone within each syllable presented
    - Same speaker within each pair
  - Participants had to decide whether the tones in each pair were the same or different
Methods

- Perceptual discrimination task
  - 296 pairs of natural Mandarin single-syllable words were played through headphones
  - All possible pairings of tone within each syllable presented
    - Same speaker within each pair
    - 500ms ISI
  - Participants had to decide whether the tones in each pair were the same or different by means of a button-press
Methods

- **Self-paced imitation task**
  - Heard one sound, and then imitated it
    - Utterances were digitally recorded for later analysis
    - Participants were encouraged to repeat or correct whenever they feel necessary.

- **Data analysis**
  - Judged by 36 native Mandarin Chinese speakers
  - The production was considered accurate when two of the judges had the same identification
Results

- the English vocalists outperformed the English non-musicians in tone discrimination, but the English instrumentalists did not show such advantage
- the difference between the English vocalists and the English instrumentalists did not reach significance

Figure 1: Mean d’ score with standard error for the four groups in the perception task.
Results

- No difference between the three English groups
- Pair22: the vocalists marginally outperformed the instrumentalists

Figure 3: Mean accuracy percentages with standard error for the four groups in the perception task.
Results

- The vocalists and the instrumentalists performed almost as accurately as the Chinese non-musicians
- both the musician groups outperformed the English non-musicians
Results

- Production

- No difference between the three English groups
Conclusion

- Music training improves the pitch perception in the language domain

- To some extent the vocal training in addition to instrumental training was an advantage for lexical tone perception.

- However, lexical tone production was not significantly facilitated by either type of music training.
Thank you!