Pronunciation in Foreign Language: How to train?

Effects of different kinds of training in perception and production of the Italian consonant / λ/ by adult German learners

Manuela Macedonia, Jonas Obleser, & Angela Friederici

Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig

macedonia@cbs.mpg.de

MAX-PLANCK-GESELLSCHAFT

MAX PLANCK INSTITUTE FOR HUMAN COGNITIVE AND BRAIN SCIENCES LEIPZIG

Introduction

This study investigates the role of perception and sensory motor learning on speech production. The experiment was based on the hypothesis that sensory motor cues together with imitation induce the development of articulatory motor programmes and accurate pronunciation.

Compared to natural language learning (Kuhl, 2000) acoustic input in adult formal instruction consists of listening and comprehension activities, is devoid of multiple sensory motor cues and lacks the imitation component. Transfer effects from perception training to production performance are documented (Bradlow et al.,1997; Wang et al., 2003; Hazan et al., 2005).

The Motor Theory of Speech Perception (Liberman & Mattingly, 1985) provides interesting theoretical underpinning for the interdependency between perception and production: it postulates that speech is perceived by reference to the articulator gestures necessary to produce it. Accordingly in foreign language learning the motor programmes have to be initiated and trained in order to perceive the sounds and to reproduce them accurately.

By running our study we aimed to assess if sensory motor training can induce the creation of motor programmes, native-like accuracy of performance in adults´ perception and articulation of the Italian phoneme / λ/.

Our first behavioural experiment focussed on the following questions:

- Do sensory motor cues combined with imitation enhance the accuracy of perception in adults (Navarra & Soto-Faraco, 2005)?
- If adults are provided with appropriate sensory motor cues and are trained to imitate, does this lead to more accurate pronunciation?

Predictions

Sensory motor cues and imitation procedures enhance the accurateness in phoneme perception and production;
Production training enhances perception.

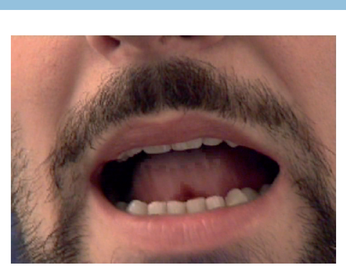
Methods

Participants

49 native Germans (25 males, 24 females) sessions 1, 2; session 3 only 34 subjects.
Age, 19 to 36 yrs (mean: 26.12). No experience in any romance language. No visual, auditory or neurological impairment.

Stimuli

a) Video clips files (length betw. 800 and 1600 ms), mouth region of 2 native Italians (male / female), frontal and side perspective.
Minimal pairs of syllables containing the Italian phonemes /λ/ and /l/ embedded in vocalic contexts with the structure CV, VCV like in / la / / λa / / ala / / aλa /.
Altogether: 102 stimuli: 51 with /l/ and 51 with /λ/.



b) Audio files extracted from the videos and identical in all respects to the video files.

Training conditions

"Audiovisual / Imitation" (AVI) condition (17 participants)
Task: watch the mouth, listen to the syllable and repeat what you hear.

"Audiovisual / Motor task" (AVM) condition (16 participants)
Task: watch the mouth + listen to the syllable and press your lips.

"Acoustic / Imitation" (AI) condition (16 participants)
Task: listen and repeat what you hear.

Experimental sessions:

1. Day 1 training + perception test;
2. Day 2 (after 24 hours) perception + production test;
3. Day 3 (after 60 days) refreshing, perception + production test.

Training procedures

Familiarization: 7 minutes
(Task: watch the video or listen). Stimuli ranged in minimal pairs like / la / / λa /, / le / / λe /…

2 training blocks: approx. 20 minutes each, total of 408 responses per participant (204 containing /l/ and /204/ containing / λ /).

Random presentation of stimuli: discrimination between "familiar" sound /l/ or the "new" sound /λ/ by button pressing. Feed back positive /negative smiley symbol.
Thereafter: repeated stimulus presentation and reproduction (AVM and AI) or lip pressing (AVM). Speech production instantly scored by an Italian phonetician and feed back sent to the participant´s screen.

Perception testing procedure

Sessions 2 and 3: without video and feedback. Perception scores: log files.

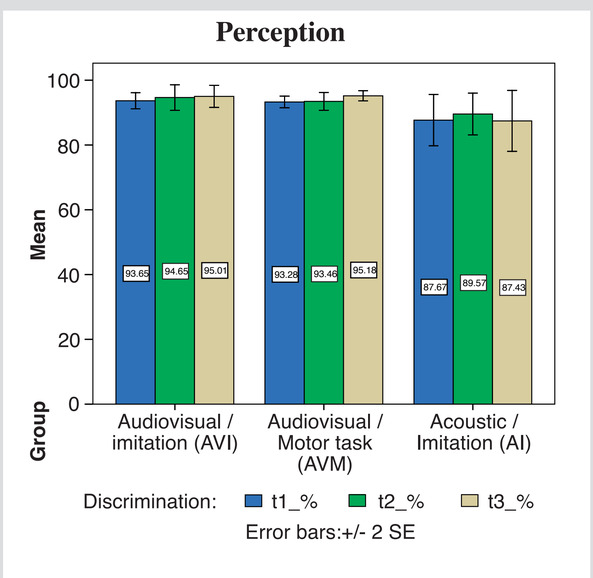
Production testing procedure

Sessions 2 and 3: production recorded and scored 3 times (2 blind scores) by phonetician.

Results

Perception t1 / t2 / t3

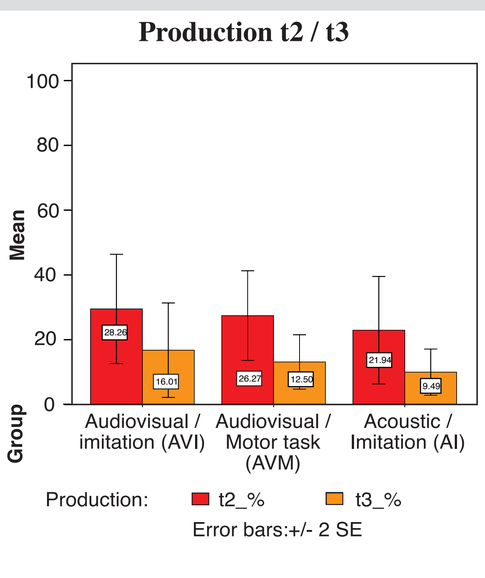
The graph shows ceiling effects in perception.



The mean error of the AI condition showed high variability among subjects. Therefore we divided the population in low and high proficiency subjects and we performed a repeated measures ANOVA with training and proficiency as between factors. Result: significant main effect of training .046 F (2,28) = 3,588. For the combined factor training /proficiency in perception the result approached the significance F (2,28) = 3,328 p = .051.

Production t2 / t3

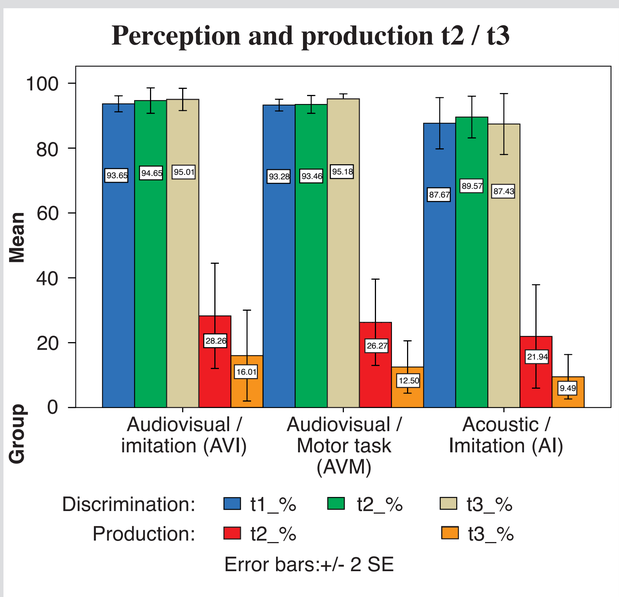
The graph shows slight mean differences among the training groups.



The repeated-measures-ANOVA showed that the factor time affected the performance significantly: $F(1, 33) = 11.2$ $p = .002$ whereas the combined factors time*group and the group factor alone lost significance respectively $F(2, 31) = 0.015$ $p = 0.985$ and $F(2, 31) = 0.015$ $p = .712$

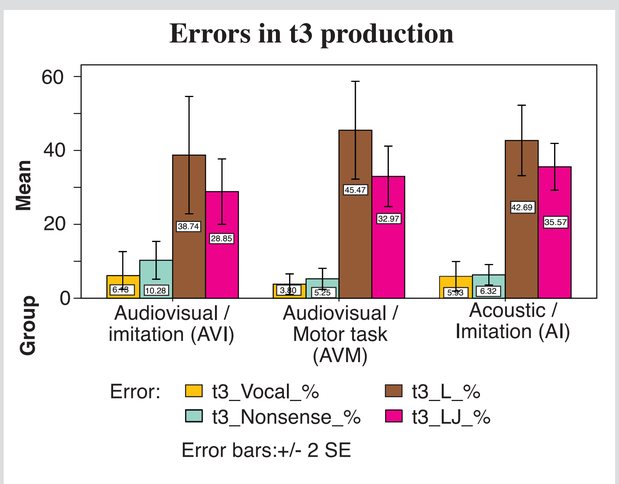
Relationship between perception and production t2 / t3

Accuracy in production was in comparison to perception low even after training and it decayed dramatically after the interval between both tests.

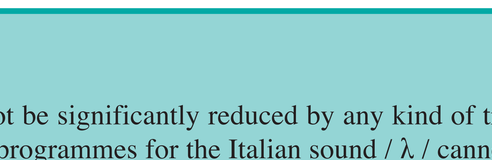


Analysis of errors t3

87,33% (mean) of the potential production was performed as "error".
The target sound /λ/ was substituted by: /l/ , the combination /l/ + /j/, new sound combinations ranging from aspirations to plosives like /d/ or combinations of both of them, the vocal /j:/.



Input [oλi] vs. production [o_llj]



Typical error: Substitution of [λ] by [l_j]
The audible difference to expert ears is often a short break between the two phonemes.

Discussion

- This first study comparing three different training procedures shows that the adult learner's pronunciation of a foreign phoneme is remarkably resistant to the often propagated influences of visual cues and imitation.
- The lack of visual cues affected low proficiency learners. They performed better if they were provided sensory motor training.
- The perception performance was enhanced by production training.
- The production performance, however, was very poor and no training procedure elicited significant improvement.

The gap between perception and production could not be significantly reduced by any kind of training at any time. This suggests that the creation of motor programmes for the Italian sound / λ / cannot only derive from perception transfer effects.

The error analysis shows that our participants hardly created the new motor programme and they rather "slipped" into mother tongue articulation. This could provide support for the motor theory of speech perception: thereafter the learners would perceive influenced by the motor programmes of their first language and reproduce what they have perceived by substituting the target phoneme with other sounds.

The production results might be related with "natural" reluctance in adults to learn pronunciation (Iverson, 2002) but also to the training procedure we used. They could be improved by higher frequency in training sessions, varied context of consonantal embedding, perception task variation by addition of the frequent error /l/ + /j/.

Learning accurate pronunciation in adult age might be successful but might require extensive focussed training and considerable time. Given the settings in formal language instruction (30 hours beginner classes) the question arises if it is possible at all or if new media could provide an efficient training environment.

References

[1] Bradlow, A.R.; Pisoni, D.B.; Akahane-Yamada, R.; Tohkura, Y.(1997): Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America, 101*, 2299-2310.

[2] Hazan, V.; Sennema, A.; Faulkner, A.; Ortega-Llebaria, M.; Iba, M.; Chunge, H. (2006): The use of visual cues in the perception of non-native consonant contrasts. *The Journal of the Acoustical Society of America, 119,* 1740-1751.

[3] Iverson, P.; Kuhl, P.K.; Akahane-Yamada, R.; Diesch, E.; Tohkura, Y.; Kettermann, A., Siebert, C. (2003): A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition, 87,* 47-57.

[4] Kuhl, P.K. (2000): A new view of language acquisition. *Proceedings of the National Academy of Science U S A., 97*, 11850-11857.

[5] Liberman, A.M.; Mattingly, I.G. (1985): The motor theory of speech perception revised. *Cognition, 21*, 1-36.

[6] Navarra, J.; Soto-Faraco, S.(2007): Hearing lips in a second language: visual articulatory information enables the perception of second language sounds. Psychological research, 71, 4-12.

[7] Wang, Y.; Jongman, A.; Sereno, J.A. (2003): Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *The Journal of the Acoustical Society of America, 113*, 1033-43.

Aknowledgments
This research was supported by a grant to M.Macedonia by the Cogito Foundation, Wollerau, Switzerland.