



HAL
open science

Plasticity of sensory-motor goals in speech production: behavioral evidence from phonetic convergence and speech imitation

Marc Sato, Krystyna Grabski, Maëva Garnier, Lionel Granjon, Jean-Luc Schwartz, Noël Nguyen

► To cite this version:

Marc Sato, Krystyna Grabski, Maëva Garnier, Lionel Granjon, Jean-Luc Schwartz, et al.. Plasticity of sensory-motor goals in speech production: behavioral evidence from phonetic convergence and speech imitation. ISICS 2012: International Symposium on Imitation and Convergence in Speech, Sep 2012, Aix-en-Provence, France. pp.47-48. hal-00762370

HAL Id: hal-00762370

<https://hal.science/hal-00762370>

Submitted on 7 Dec 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Plasticity of sensory-motor goals in speech production: Behavioral evidence from phonetic convergence and speech imitation

Marc Sato¹, Krystyna Grabski¹, Maëva Garnier¹, Lionel Granjon¹, Jean-Luc Schwartz¹, Noël Nguyen²

¹Gipsa-Lab, CNRS & Grenoble Universités, France

²Laboratoire Parole & Langage, CNRS & Aix-Marseille Université

marc.sato@gipsa-lab.inpg.fr

Imitation is one of the major processes by which humans develop social interactions. In speech communication, imitative processes are used from birth to adulthood, as highlighted by children's mimicking abilities and by adult's tendency to automatically "imitate" a number of acoustic-phonetic characteristics in another speaker's speech. These adaptive changes are thought to play a key role in speech development/acquisition and to facilitate conversational exchange by contributing to setting a common perceptuo-motor link between speakers. Based on acoustic analyses of speech production in various laboratory tasks, the present study aimed to better characterize sensory-to-motor adaptive processes involved in unintentional as well as voluntary speech imitation, and to test possible motor plastic changes due to auditory-motor recalibration mechanisms.

Methods

Three groups of participants involved in speech production or imitation tasks were exposed via loudspeakers to vowel utterances spoken by different speakers. The first task was designed to induce unintentional imitation of acoustically presented vowels and to measure the magnitude of imitative changes in speech production as well as possible motor after-effects. To this aim, participants were instructed to produce vowels according to either an orthographic or an acoustic cue, without any instructions to repeat or to imitate the acoustic cues. A block design was used where participants produced a vowel target according first to an orthographic cue (baseline), then to an acoustic cue (phonetic convergence) and finally to an orthographic cue (motor after-effect). To compare phonetic convergence and voluntary imitation of the acoustic vowels, we asked the second group of participants to imitate the acoustically presented vowels. In a third task, we tested whether motor after-effects can also occur without prior unintentional or voluntary vowel imitation but only after auditory exposure of the acoustic targets.

The three tasks were performed in a soundproof room using the same experimental setting and participants' productions were recorded for offline analyses. A semi-automatic procedure was first devised for segmenting participants' recorded vowels (around 10000 utterances). For each participant, the procedure involved the automatic segmentation of each vowel based on an intensity and duration algorithm detection. The algorithm automatically identified pauses (with minimal duration of 1000 ms and low intensity energy inferior to 55 dB) between each vowel by marking boundaries. If necessary, these boundaries were hand-corrected, based on waveform and spectrogram information, so as to correctly mark the onset and offset of vowels. After individual sound file extraction of each vowel, omissions, wrong productions and hesitations were manually identified and removed from the analyses. Finally, for each vowel, F0 and F1 values were calculated from a period defined as ± 25 ms of the maximum peak intensity of the sound file. For each participant, median F0 and F1 values were first computed for each vowel and expressed in bark. For each experiment, median F0 and F1 exceeding ± 2 standard deviations from the mean were removed from the analyses.

Results

Phonetic convergence and imitation (Experiments A and B): For each participant and vowel, median F0- and F1-responses observed during the presentation of the acoustic cue were subtracted from the preceding baseline (i.e., median F0- and F1-responses observed in the preceding sub-block during the presentation of the orthographic cues). These values were then correlated with F0 and F1 values of the respective acoustic cue subtracted from the preceding baseline. Single subject correlation coefficients were calculated for both F0 and F1 and entered

into analyses of variance (ANOVA) with the experiment (phonetic convergence, imitation) as a between-subject variable. In addition, individual one-tailed t-tests (with Bonferroni corrected p-value) were performed for each experiment in order to test significant correlation coefficients (compared to zero).

ANOVA on single subject correlation coefficients for F0 show a significant effect of the task. In addition, correlation coefficients differed significantly from zero in both Experiment A and Experiment B. For F1, there was no significant effect of the task. Correlation coefficients also differed significantly from zero in both Experiment A and Experiment B.

After-effects (Experiments A, B and C): For each participant and vowel, median F0- and F1-responses observed during the second presentation of the orthographic cue were subtracted from the preceding baseline. These values were then correlated with F0 and F1 values of the respective acoustic cue subtracted from the preceding baseline. As previously, single subject correlation coefficients were calculated for both F0 and F1 and entered into analyses of variance (ANOVA) with the experiment (phonetic convergence, imitation, perceptual categorization) as a between-subject variable. In addition, individual one-tailed t-tests (with Bonferroni corrected p-value) were performed for each experiment in order to test significant correlation coefficients (compared to zero).

ANOVA on single subject correlation coefficients for F0 showed no significant effect of the task. In addition, correlation coefficients differed significantly from zero in both Experiment B and Experiment C but not in Experiment A. For F1, there was no significant effect of the task. Correlation coefficients did not differ significantly from zero in Experiments A, B and C.

Conclusion

These results demonstrate automatic imitative processes during speech communication even at a fine-grained acoustic-phonetic level and highlight the online plasticity of phonemic sensory-motor goals during speech production, although only for F0 acoustic parameter. They will be discussed in relation with forward and inverse internal models of speech production in which feedback control mechanisms allow evaluating the sensory consequence of the speech-motor act with actual sensory input in order to further control production.