



**HAL**  
open science

# Sensory-motor relationships in speech production in post-lingually deaf cochlear-implanted adults and normal-hearing seniors: Evidence from phonetic convergence and speech imitation

Lucie Scarbel, Denis Beaudemp, Jean-Luc Schwartz, Marc Sato

## ► To cite this version:

Lucie Scarbel, Denis Beaudemp, Jean-Luc Schwartz, Marc Sato. Sensory-motor relationships in speech production in post-lingually deaf cochlear-implanted adults and normal-hearing seniors: Evidence from phonetic convergence and speech imitation. *Neuropsychologia*, 2017, *Neuropsychologia*, 101, pp.39 - 46. 10.1016/j.neuropsychologia.2017.05.005 . hal-01546755

**HAL Id: hal-01546755**

**<https://hal.science/hal-01546755>**

Submitted on 26 Jun 2017

**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.

**Sensory-motor relationships in speech production in post-lingually deaf cochlear-implanted adults and normal-hearing seniors: evidence from phonetic convergence and speech imitation**

Lucie Scarbel<sup>1</sup>, Denis Beautemps<sup>1</sup>, Jean-Luc Schwartz<sup>1</sup>, & Marc Sato<sup>2</sup>

<sup>1</sup> GIPSA-LAB, Département Parole & Cognition, CNRS & Grenoble Université, Grenoble, France

<sup>2</sup> Laboratoire Parole & Langage, CNRS & Aix-Marseille Université, Aix-en-Provence, France

Authors' note:

Correspondence can be addressed to Lucie Scarbel, GIPSA-LAB, UMR CNRS 5216, Grenoble Université, Domaine Universitaire, BP 46, 38402 Saint Martin d'Hères, France. Tel: (+33) (0)476 824 128. Fax: (+33) (0)476 574 710. Email: lucie.scarbel@gipsa-lab.grenoble-inp.fr.

The authors declare no competing financial interests.



**Abstract:**

Speech communication can be viewed as an interactive process involving a functional coupling between sensory and motor systems. One striking example comes from phonetic convergence, when speakers automatically tend to mimic their interlocutor's speech during communicative interaction. The goal of this study was to investigate sensory-motor linkage in speech production in postlingually deaf cochlear implanted participants and normal hearing elderly adults through phonetic convergence and imitation. To this aim, two vowel production tasks, with or without instruction to imitate an acoustic vowel, were proposed to three groups of young adults with normal hearing, elderly adults with normal hearing and post-lingually deaf cochlear-implanted patients. Measure of the deviation of each participant's  $f_0$  from their own mean  $f_0$  was measured to evaluate the ability to converge to each acoustic target.

Results showed that cochlear-implanted participants have the ability to converge to an acoustic target, both intentionally and unintentionally, albeit with a lower degree than young and elderly participants with normal hearing. By providing evidence for phonetic convergence and speech imitation, these results suggest that, as in young adults, perceptuo-motor relationships are efficient in elderly adults with normal hearing and that cochlear-implanted adults recovered significant perceptuo-motor abilities following cochlear implantation.

*Key-words:* Phonetic convergence, imitation, cochlear implant, elderly, sensory-motor interactions, speech production

## 1 **1. Introduction**

### 2 **1.1.Sensory-motor interactions in speech perception and speech production**

3 In the speech communication domain, an old and fundamental debate concerns the nature of processes  
4 and representations involved in speech perception and production. Concerning speech perception,  
5 auditory theories assume that speech perceptual processing and categorization are based on acoustic  
6 features and auditory representations (Stevens and Blumstein 1978, 1979; Lindblom et al. 1988, 1990).  
7 Conversely, the motor theory of speech perception (Liberman et al., 1985) or the direct realist theory  
8 (Fowler, 2005) respectively claim that speech perception involves the recovery of the speaker's motor  
9 intentions or articulatory gestures. More recently, various perceptuo-motor theories introduced  
10 syntheses of arguments by tenants of both auditory and motor theories, and proposed that implicit  
11 motor knowledge and motor representations are used in relationship with auditory representations and  
12 processes to elaborate phonetic decisions (Skipper et al., 2007; Schwartz et al., 2012). These theories  
13 capitalize on the increasing amount of evidence for the role of motor representations and processes in  
14 speech perception (see a recent review in Skipper et al., 2017).

15 Concerning speech production, various theories have also been introduced to characterize the  
16 speaker's goals. Motor or articulatory theories, like Task dynamics (Saltzman, 1986) and the  
17 associated Articulatory Phonology (Browman & Goldstein, 1992), consider that speech targets are  
18 defined in the articulatory space. On the contrary, auditory theories like Stevens' Quantal Theory  
19 (Stevens 1972, 1988) and Perkell's speech production control model (Perkell et al. 1995, 2000),  
20 suggest that targets are specified in auditory terms. Finally, sensory-motor models claim that speech  
21 production control is multimodal and combines auditory and somatosensory information (see Perrier,  
22 2005; and the DIVA model, Guenther et al. 1998, Guenther & Vladusich, 2012). A number of data  
23 about the effect of auditory (e.g. Lametti et al., 2014; Shiller & Rochin, 2014) or somatosensory  
24 (Tremblay et al., 2003) perturbations applied to speech production in adults or children are in  
25 agreement with this sensory-motor framework.

26 From a number of these theories, sensory-motor relationships seem to play an important role in  
27 both speech perception and speech production in normal conditions. The present study deals with  
28 subjects for which some degradation of the sensory-motor link could be expected, because of sensory  
29 deficits. The first and primary situation is the case of post-lingually deaf subjects equipped with  
30 cochlear implants. Indeed, post-lingually deaf subjects are expected to have acquired efficient sensory-  
31 motor relationships before deafness. Their ability to achieve intelligible speech production is  
32 interpreted as evidence that they have maintained a stable and rather accurate internal model of these  
33 sensory-motor relationships all along their deafness life (Perkell et al., 2000). Cochlear implantation  
34 then results in providing them with a new kind of auditory input, quite different from the one they had  
35 acquired before deafness considering the very special nature of cochlear-implant coding of acoustic  
36 information. The question is to know how the internal relationships between these “new” auditory  
37 inputs, the “old” ones acquired at the first stages of development, and the motor representations are  
38 established and organized, and how efficient they are.

39 A number of data on speech production after cochlear implantation display an increase in acoustic  
40 dispersion of vowels after a few months of implantation (e.g. Langereis et al., 1998; Lane et al., 2005;  
41 Ménard et al., 2007), suggesting that the internal model has indeed benefited from the new auditory  
42 input provided by cochlear implantation. Interestingly however, some studies explored the response to  
43 articulatory perturbation using a bite-block task (Lane et al., 2005), a robotic device displacing the jaw  
44 during speech (Nasir & Ostry, 2008) or a lip-tube task (Turgeon, Prémont, Trudeau-Fisette, & Ménard,  
45 2015), with or without auditory feedback. Results showed that post-lingual cochlear-implanted  
46 participants (CI) are able to adapt their articulatory trajectory when it is perturbed in order to reach  
47 their auditory goals and make their production intelligible, even when the implant is turned off. Such  
48 compensatory strategies show that perceptuo-motor abilities acquired during speech acquisition in CI  
49 subjects are still at work after deafness, though the addition of auditory information provided by the  
50 cochlear-implant does result in enhanced precision and efficiency of the sensory-motor internal model.  
51 Finally, a PET-scanning study on visual speech perception in post-lingually deaf CI patients (Rouger  
52 et al., 2012) showed that after a short adaptation period with the implant, there was a decrease of the

53 initially abnormal activity in the superior temporal sulcus, a cross-modal brain area, accompanied by a  
54 progressive reactivation of frontal premotor speech areas. This suggests that sensorimotor  
55 neuroplasticity after cochlear implantation provides a progressive reactivation of the audio-visuo-  
56 motor linkage in CI subjects.

57 The second situation considered in the present study is the case of elderly subjects. They are of  
58 interest for us for two reasons. The first one is the interest of assessing the consequences of the  
59 potential decline of cognitive and language abilities in relation with decline in sensory and motor  
60 accuracy on the efficiency of sensory-motor relationships. The second reason is that a number of the  
61 cochlear-implanted post-lingually deaf participants are rather old, and we considered important to  
62 include senior subjects with no severe auditory deficit as a control population. As a matter of fact, very  
63 few studies investigated perceptuo-motor relationships in elderly population. One of them shows that  
64 elderly people adapt their production in case of degraded auditory feedback (Liu et al., 2010, 2011).  
65 Sensorimotor neuroplasticity was also observed in elderly people, linked with age-dependent  
66 intelligibility effects mainly found in auditory and motor cortical areas (Tremblay, Dick, & Small,  
67 2013; Bilodeau-Mercure, Ouellet, & Tremblay, 2015). Taken together, these behavioral and neuro-  
68 imaging studies suggest that sensory-motor relationships are well preserved in aging.

### 69 **1.2. Phonetic convergence and imitation, a paradigm for studying sensory-motor** 70 **relationships in speech**

71 Imitation is a quite widespread phenomenon in speech communication and can be viewed as a key  
72 mechanism in the acquisition of human language. In interactive situations, adult speakers tend to  
73 continuously adapt their productions to those of their interlocutor, in order to facilitate communicative  
74 exchanges. These adaptive changes in speech production can be voluntary, that is when the speaker  
75 consciously imitates his interlocutor, but also unintentional. Indeed, unintentional imitation, or  
76 phonetic convergence, that is the tendency to automatically imitate a number of acoustic-phonetic  
77 characteristics in the productions of an interacting speaker, has been displayed in various studies (see a  
78 recent collection of papers on this topic in Nguyen et al., 2013, and for recent reviews see Babel, 2009;  
79 Aubanel, 2011; Lelong, 2012). In these studies, convergence effects have been observed both in

80 paralinguistic features, such as gestures, and in speech acoustic parameters, such as intensity,  
81 fundamental frequency  $f_0$  or formants. These phonetic convergence effects may be related to the social  
82 component of communication, assuming that they would contribute to setting a common ground  
83 between speakers (Giles, Coupland, & Coupland, 1991) and that they could be associated to the  
84 human desire of affiliation to a social group.

85 Interestingly however, while most reports on phonetic convergence are based on conversational  
86 exchanges in natural conditions, a few studies showed that phonetic convergence can also be observed  
87 using laboratory settings involving non-interactive situations of communication (Goldinger & Azuma,  
88 2004; Gentilucci & Cattaneo, 2005; Delvaux & Soquet, 2007; Garnier, Lamalle, & Sato, 2013; Sato et  
89 al., 2013). For example, Delvaux & Soquet (2007) obtained phonetic convergence effects during the  
90 production of auditorily presented words without interaction, and they also reported offline adaptation  
91 to the auditory targets in post-tests following stimulus exposure. These studies hence suggest that  
92 phonetic convergence is not only a matter of social attunement, but could also involve a more basic  
93 stage of continuous automatic adaptation of the speech production system to the external speech  
94 sounds environment (for a recent review, see Sato et al., 2013).

95 Such an automatic adaptation mechanism requires the existence of a sensory-motor coupling  
96 process in which variations in the external environment provide sensory targets that drive motor  
97 control procedures to adapt and produce stimuli closer to these external targets. More in detail, the  
98 speaker would program motor commands to achieve an articulatory and ultimately an auditory (or  
99 more generally sensory) goal according to the linguistic message to be conveyed to the listener. Then,  
100 in the course of speech production, the speaker would compare sensory feedback to the initial sensory  
101 objective. Variations in the external environment would shift the sensory targets and accordingly result  
102 in modifications of motor commands to converge towards speech sounds in the environment. Hence,  
103 the convergence/imitation paradigm is a natural tool to study sensory-motor coupling in speech  
104 communication.

105 Phonetic convergence and imitation in laboratory settings can involve various kinds of acoustic  
106 features, though fundamental frequency effects seem to be larger and easier to obtain than variations in



107 e.g. formant values (Sato et al., 2013). The present study therefore focus on convergence and imitation  
108 on  $f_0$  variations.

### 109 **1.3.f0 perception and motor control in cochlear-implanted adults and normal-hearing** 110 **seniors**

111 Past studies on pitch control in speech production by CI subjects reported higher  $f_0$  values for CI  
112 adults than for adults with normal hearing and, more importantly, larger variations in  $f_0$  for CI than for  
113 adults with normal hearing (e.g. Lane & Webster, 1991). Langereis (1998) and Kishon-Rabin et al.  
114 (1999) also reported that while pitch production is generally altered in CI patients soon after  
115 implantation, almost normal  $f_0$  values were observed one year post-implantation for two participants in  
116 Kishon-Rabin et al. (1999) study. Notice that technological evolutions led to important progress in the  
117 coding of pitch in cochlear implants since these studies.

118 Concerning speech perception, several factors appear to influence performance in CI patients. In  
119 Blamey et al. (2012), 2251 CI patients participated to a battery of auditory tests, where they had to  
120 recognize phonemes, words and sentences. The experimenters reported that both duration of implant  
121 experience, age at onset of severe to profound hearing loss, age at cochlear implantation and duration  
122 of deafness influence speech perception to a certain extent – though inter-subject variability was quite  
123 large in this kind of study. Regarding  $f_0$ , for cochlear-implanted patients, pitch is difficult to estimate  
124 because of different reasons. First, as the implant is composed of 12 to 22 electrodes, the merging of  
125 frequencies in a given electrode likely impacts frequency discrimination by patients. Furthermore,  
126 because of their proximity, electrodes can stimulate nerve fibers, which do not correspond to the  
127 intended frequency, which leads to an imperfect tonotopy restitution.

128 Regarding elderly people, it is well known that auditory capacities decline with aging. More  
129 specifically, even with close-to-normal auditory thresholds, elderly people with normal hearing  
130 present difficulties in adverse listening conditions, e.g. in a noisy environment (Gelfand, Piper, &  
131 Silman, 1985; Ohde & Abou-Khalil, 2001; Fullgrabe, 2013). The role of cognitive functions in this  
132 decline of speech perception is still under debate (Cienkovski & Camey, 2002; Cienkovski & Vasil-  
133 Dilaj, 2010; Moore & Fullgrabe, 2013; see a review in Fullgrabe & Rosen, 2016). On the motor side,  
134 elderly people also show deficits in orofacial movements, which might impact speech production.

135 Indeed, a majority of studies demonstrate a degradation of speech production in elderly people, with  
136 more variable and less stable speech utterances than observed in younger adults. An alteration of voice  
137 control is also observed in elderly people, notably with higher pitch variability and less stable pitch  
138 production at the level of individual subjects (Morgan & Rastatter, 1986; Russell, Penny, &  
139 Pemberton, 1995; Lortie, Thibeault, Guitton, & Tremblay, 2015).

#### 140 **1.4. Goal of the study**

141 In the present study, voluntary and unintentional imitative changes were tested during speech  
142 production in both post-lingually deaf or hearing impaired adults equipped with a cochlear implant and  
143 in seniors with normal hearing. Our goal was to explore in these populations the underlying perceptuo-  
144 motor mechanisms at work in phonetic convergence and speech imitation. To this aim, we capitalized  
145 on two recent studies by Sato et al. (2013) and Garnier et al. (2013) displaying both unintentional and  
146 voluntary imitative changes on fundamental frequency of auditorily presented vowel targets during  
147 speech production in a non-interactive situation of communication. In these studies, participants were  
148 asked to produce different vowels according to acoustic targets with their pitch varying around the  
149 averaged pitch of their own voice, with or without instruction to imitate the target. Results showed that  
150 participants strongly converged to the vowel target not only in the imitative task, but also, at a lower  
151 degree, in the production task even if no instruction to imitate the vowel target was given.

152 To our knowledge, convergence and imitation abilities have never been studied either in CI  
153 subjects or in elderly people with normal hearing. We here report online imitative changes on the  
154 fundamental frequency in relation to acoustic vowel targets in a non-interactive situation of  
155 communication during both unintentional and voluntary imitative production tasks in the two studied  
156 populations. Our goal was to determine to which extent CI and elderly participants display  
157 convergence and imitation abilities, and to compare these abilities with those observed in young adults  
158 with normal hearing (NHY), with the larger aim to evaluate sensory-motor linkage during speech  
159 production in CI and elderly population. Three populations have been contrasted: post-lingually deaf  
160 or hearing-impaired cochlear-implanted adults (generally not very young), senior adults with no  
161 specific auditory deficit apart from aging, and normal-hearing young adults as a control. The same

162 paradigm as the one developed by Garnier et al. (2013) and Sato et al. (2013) was exploited in this  
163 study.

164

165

166

167 **2. Methods**168 **2.1. Participants**

169 Three groups of participants performed the experiment. The first group consisted of fifteen young  
170 participants with normal hearing (NHY) (10 females and 5 males, mean age: 30 years old, range: 20-  
171 40) who reported no history of speaking, hearing or motor disorders. The second group consisted of  
172 ten elderly participants with normal hearing (NHE) (4 females and 6 males, mean age: 69 years old,  
173 range: 63-78). The third group consisted of ten post-lingually deaf cochlear-implanted (CI)  
174 participants (7 males and 3 females, mean age: 58.9 years old, range: 27-76). As indicated in Table 1,  
175 the CI group was heterogeneous in several aspects: age at onset of deafness varied from 7 to 65 years,  
176 duration of deafness varied from 1 month to 58 years and duration of cochlear implant experience  
177 varied from 1 month to 9 years. In addition, seven of the ten participants wore classical hearing aid in  
178 the non-implanted ear, and one participant was bilaterally implanted. The experiment was performed  
179 in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki.

180

---

Insert Table 1 about here

---

181

182 **2.2. Stimuli**

183 A vowel database was created from /e/, /oe/, /o/ French vowels produced by one male and one female  
184 speaker. For each vowel, one clearly articulated occurrence was selected and digitized at 44.1 kHz.  
185 From these stimuli,  $f_0$  was artificially shifted by steps of  $\pm 5$ Hz (from 80Hz to 180Hz for the male  
186 vowels, and from 150 to 350Hz for the female vowels) using the PSOLA module integrated in the  
187 Praat software (Boersma and Weenink, 2013). These stimuli allowed us to present to each participant  
188 nine distinct stimuli per vowel ranging from -20% and +20% from the mean participant's  $f_0$  by steps of  
189 5% (-20%, -15%, -10%, -5%, 0%, +5%, +10%, +15%, +20%).

190 **2.3. Experimental procedure**

191 The experiment was carried out in a sound-proof room. Participants sat in front of a computer monitor  
192 at a distance of approximately 50 cm. The acoustic stimuli were presented at a comfortable sound level

193 through a loudspeaker, with the same sound level set for all participants. The Presentation software  
194 (Neurobehavioral Systems, Albany, CA) was used to control the stimulus presentation during all  
195 experiments. All participants' productions were recorded for off-line analyses.

196 The experiment consisted of three vowel production tasks. First, participants had to individually  
197 produce /e/, /æ/ and /o/ vowels, according to a visual orthographic target. This allowed the  
198 experimenter to measure the participant's  $f_0$ . In the subsequent task (which will be referred by  
199 "convergence task"), participants were asked to produce the three vowels according to an acoustic  
200 target. Importantly, no instruction to "repeat" or to "imitate" the acoustic targets was given to the  
201 participants. Finally, the third task (which will be referred by "imitative task") was the same as the  
202 second task except that participants were explicitly asked to imitate the acoustic targets. The only  
203 indication given to participants was to imitate the voice characteristics of the perceived speaker.  
204 Importantly, we hence kept a fixed order of the "convergence" and "imitative" tasks so as to avoid  
205 cases in which prior explicit imitation instructions could modify further convergence effects and add  
206 conscious imitation strategies to automatic convergence processes.

207 Acoustic targets in the "convergence" and "imitative" tasks for each participant consisted in the 27  
208 stimuli selected from the vowel database (9 per vowel category), repeated three times each so as to  
209 obtain 81 trials altogether. Each participant was presented with one model talker of the same sex, with  
210 the 9 quantified  $f_0$  frequencies varying from -20% to +20% by steps of 5% around his/her own pitch,  
211 as measured in the first task.

212 At the end of the experiment, participants were asked to perform a frequency discrimination test to  
213 estimate their pitch just noticeable difference (JND), based on an experimental paradigm proposed by  
214 Vinay & Moore (2010). Two groups of four sounds were presented to participants in a random order,  
215 composed either of identical sounds (AAAA) or of two sounds A and B differing in pitch (ABAB),  
216 and participants had to determine in which group sounds were different among each other. We used  
217 synthesized stimuli, obtained by addition of harmonics of a given  $f_0$  value, with 0 phase and amplitude  
218 based on the spectrum of a vowel /e/ produced by either a male or a female, depending of the  
219 participant's gender. For sound A, the  $f_0$  value was set at the participant's mean production value. For  
220 sound B, the selected  $f_0$  value was systematically higher than  $f_0$  for sound A, beginning with a 5%

221 deviation further varying along the test. Moreover, small random intensity variations were applied to  
222 the two groups of sounds from one trial to the other, in order to force the participant to focus on pitch  
223 rather than timbre or intensity perception (Vinay & Moore, 2010). Variations of the  $f_0$  values for sound  
224 B were driven by an adaptive two-alternative forced choice procedure: after two consecutive correct  
225 answers, the frequency difference between sounds A and B decreased, whereas one mistake drove an  
226 increase of this difference. The test result corresponded to the mean of the difference between  $f_0$  values  
227 for sound A and sound B for the last eight trials.

#### 228 **2.4.Data analysis**

229 All acoustic analyses of participants' productions were performed using the Praat software (Boersma  
230 & Weenink, 2013). In the second and third tasks which both involved an acoustic target (/e/, /æ/ or /o/)  
231 that could possibly be misunderstood by the participant, we annotated the participant productions in  
232 order to estimate the percentage of errors in production in relation to the target.

233 Analysis began by a semi-automatic segmentation procedure aiming at segmenting each individual  
234 vowel produced by a given participant in a given task. Using intensity and duration criteria, the  
235 algorithm automatically identified pauses between each vowel and segmented individual productions  
236 accordingly. Segmentation was hand-corrected when appropriate, checking waveform and  
237 spectrogram. Only correctly produced vowels (that is, checked by the experimenter as corresponding  
238 to the phonemic target) were further analyzed. For each correctly produced vowel,  $f_0$  was determined  
239 within a time frame around  $\pm 25$ ms of the maximum intensity of the sound file.

#### 240 **2.5.Statistical analysis**

241 For each subject and task, there were altogether 9 values of  $f_0$  production per target  $f_0$  value between -  
242 20% and 20% (with 3 vowel types and 3 utterances per type). The mean of the 9 produced  $f_0$  values  
243 was computed, and linear regression between mean produced  $f_0$  values and  $f_0$  targets was computed for  
244 each participant separately in the convergence and imitation tasks,. The slope and correlation  
245 coefficients were then determined for each task and each participant. Individual t-tests were done on  
246 these correlation coefficients to determine if they were significant.

247 For each group, one-tailed t-tests were then performed on individual slope and correlation  
248 coefficients compared to zero in order to determine significant imitative changes in each task. To test

249 for possible differences in imitative changes from one task to the other, additional t-tests were  
250 performed on slope and correlation coefficients between the two tasks. However, it is important to  
251 remind that the order between the two tasks being not counterbalanced, any difference between them  
252 can be interpreted both as a consequence of difference in tasks and order effect.

253 In addition, in order to test whether imitative changes in the convergence and imitation tasks may  
254 correlate among speakers, a Pearson's correlation analysis was performed between slope coefficients.  
255 Finally, Pearson's correlation analyses were performed between slope coefficients and JND values  
256 within each group of participants and for each task, as well as between slope coefficients and age,  
257 duration of deafness, age at implantation, age at the beginning of deafness and duration of implant  
258 experience for CI participants in each task.

259 To compare the results of the three groups, a repeated-measures ANOVA was finally performed on  
260 slope coefficients with the group (NHY vs. NHE vs. CI) as between-subject variable and the task  
261 (convergence vs. imitation) as within-subject variable. Additionally, a one-factor ANOVA was  
262 performed on JND values, with group as categorical factor. The sphericity assumption was tested  
263 using a Mauchly test and, when necessary, Greenhouse-Geiser corrections were applied (Greenhouse  
264 and Geiser, 1959).

265

266

267

268

269

### 270 3. Results

#### 271 3.1. Young participants with normal hearing (see Figure 1)

272 For NHY participants, imitative changes were observed in both tasks, though stronger during  
273 voluntary imitation. Slope coefficients differed significantly from zero in both the convergence  
274 ( $t(14)=5.98$ ;  $p<0.001$ ) and imitation ( $t(14)=35.78$ ;  $p<0.001$ ) tasks. In addition, slope coefficients were  
275 higher in the imitation compared to the convergence tasks (on average: 0.87 vs. 0.45;  $t(14)=6.02$ ;  
276  $p<.001$ ). No significant correlation was observed between the slope coefficients in the convergence  
277 and imitation tasks ( $r^2=0.12$ ).

278 Similarly, correlation coefficients differed significantly from zero in both the convergence ( $t(14)=8.3$ ;  
279  $p<0.001$ ) and imitation ( $t(14)=93.34$ ;  $p<0.001$ ) tasks, and were higher in the imitation compared to the  
280 convergence task (on average: 0.94 vs. 0.64;  $t(14)=4.3$ ;  $p<0.001$ ). Individual analyses showed that  
281 fourteen of the sixteen participants converged to the target in the convergence task ( $r^2$  from 0.05 to  
282 0.99), and in the imitative task all of the participants converged to the target ( $r^2$  from 0.91 to 1).

283 JND values for NHY participants were at 1.06 Hz on average (range: 0.14 to 3.88 Hz), with no  
284 significant correlation between the slope coefficients in the convergence or the imitation tasks and  
285 JND values (convergence:  $r^2=0.00$   $p=.82$ , imitation:  $r^2=0.00$   $p=.9$ ).

---

286 Insert Figure 1 about here

---

287

#### 288 3.2. Elderly participants with normal hearing (see Figure 2)

289 For NHE participants, imitative changes were observed in both tasks, though stronger in voluntary  
290 imitation. Slope coefficients differed significantly from zero in both the convergence ( $t(9)=3.52$ ;  
291  $p<0.01$ ) and imitation ( $t(9)=8.5$ ;  $p<0.001$ ) tasks. In addition, slope coefficients were higher in the  
292 imitation compared to the convergence task (on average: 0.75 vs. 0.33;  $t(9)=4.97$ ;  $p<.001$ ). There was  
293 no significant correlation between convergence and imitation tasks ( $r^2=0.59$ ).

294 Similarly, correlation coefficients differed significantly from zero in both the convergence ( $t(9)=5.96$ ;  
295  $p<0.001$ ) and imitation ( $t(9)=11.78$ ;  $p<0.001$ ) tasks, and were higher in the imitation compared to the



296 convergence tasks (on average: 0.84 vs. 0.54;  $t(9)=4.12$ ;  $p<0.001$ ). Individual analyses showed that  
297 eight of the ten participants converged to the target in the convergence task ( $r^2$  from 0.35 to 0.99), and  
298 in the imitative task all of the participants except one converged to the target ( $r^2$  from 0.28 to 0.99).

299 JND values for NHE participants were at 0.95 Hz on average (range: 0.43 to 1.8 Hz) with no  
300 correlation with either the convergence or the imitation slopes (convergence:  $r^2=0.03$   $p=.63$  imitation:  
301  $r^2=0.00$   $p=.79$ ).

---

302 Insert Figure 2 about here

---

303

### 304 3.3. Cochlear-implanted participants (see Figure 3)

305 As for NHY and NHE participants, imitative changes were also observed in both tasks for CI  
306 participants. Slope coefficients differed significantly from zero in both the convergence ( $t(9)=3.24$  ;  
307  $p<0.02$ ) and imitation ( $t(9)=4.84$ );  $p<0.001$ ) tasks. In addition, slope coefficients were higher in the  
308 imitation compared to the convergence tasks (on average: 0.39 vs. 0.14;  $t(9)=3.53$ ;  $p<0.001$ ). As for  
309 NH participants, there was no significant correlation between convergence and imitation tasks  
310 ( $r^2=0.28$ ).

311 Similarly, correlation coefficients differed significantly from zero in both the convergence ( $t(9)=4.08$ ;  
312  $p<0.01$ ) and imitation ( $t(9)=5.53$ ;  $p<0.001$ ) tasks, and were higher in the imitation compared to the  
313 convergence tasks (on average: 0.49 vs. 0.24;  $t(9)=2.95$ ;  $p<0.02$ ).

314 JND values for CI participants were at 14.06 Hz on average (range: 1.71 to 45.35 Hz). There was also  
315 no correlation between the convergence or imitation tasks and the others factors, that are JND values,  
316 age, deafness duration, age at implantation, age at the beginning of deafness and duration of implant  
317 experiment (all  $r^2<0.3$ ).

318 While young and elderly participants with normal hearing made no errors in both tasks, cochlear-  
319 implanted participants made a number of errors in both tasks, with similar error percentage in the  
320 convergence (15%) and imitation tasks (13%). In order to verify if these errors influence convergence  
321 effects, correlations between production and target were measured in both tasks on error trials only, for  
322 participants presenting at least 15% of errors (four participants in each task). The correlation values  
323 are displayed in Table 2. For those participants, correlations between production and target for error

324 trials appear similar to those obtained for correct trials in both convergence and imitation tasks. As a  
325 matter of fact, the correlation values for correct trials and for error trials (when they are sufficient to do  
326 computations) are highly correlated among the 8 cases where both can be computed ( $r^2=0,71$ ,  
327  $t(6)=3,81$ ,  $p<0,005$ ).

328 To relate convergence effects with etiology variability and JND values, we report determination  
329 coefficients in both tasks for each participant (see Table 2). In the convergence task, we obtained  
330 significant correlation between target and production for five participants, with determination  
331 coefficients from 0.05 and 0.75. In the imitative task, correlation is significant for seven participants,  
332 with determination coefficients from 0.09 and 0.94. As shown in Table 2, the participants with no  
333 significant convergence in one or the other tasks do not display obvious similarity in etiology in e.g.  
334 age at deafness onset or deafness duration. Importantly, clear convergence in both tasks is obtained for  
335 some subjects after a quite reduced duration of implantation: see the cases of subjects CI1, CI2, CI5 or  
336 CI9 with significant convergence or imitation on  $f_0$  after 1 to 5 months of implantation.

337 Concerning JND values, the participants with no significant convergence in one or the other tasks  
338 do not present the worst JND values, and the participant with the highest JND value obtains significant  
339 convergence effect in both tasks. More strikingly, the corresponding JND value for this subject, at 45  
340 Hz, is actually larger than the largest variation in F0 applied to the stimuli in both convergence tasks.  
341 As a matter of fact, F0 for this subject equals 125 Hz, and the maximal variation of F0 at 20% results  
342 in a 25 Hz modification. Altogether, these facts converge to suggest that the task used for JND  
343 estimation, well adapted for normal-hearing subjects (Vinay & Moore, 2010), is probably too complex  
344 for cochlear-implanted subjects.

---

345 Insert Table 2 about here

---

346

---

347 Insert Figure 3 about here

---

348

349 **3.4. Comparison between groups (see Figure 4)**

350 The main effect of group was significant ( $F_{(2,32)}=31.3$ ;  $p<.001$ ) with lower mean slope coefficient for  
351 CI participants than for NHY and NHE participants (CI: 0.26 ; NHY: 0.65 ; NHE: 0.55). No difference  
352 was obtained between NHY and NHE groups. The main effect of task was also significant  
353 ( $F_{(1,32)}=67.29$  ;  $p<.001$ ), with lower slope coefficients for the convergence task than for the imitation  
354 task (convergence: 0.33 ; imitation: 0.69). No significant interaction was found between the group and  
355 the task.

356 For the JND test, the effect of group was significant ( $F_{(2,32)}=11.03$ ;  $p<.001$ ), with larger values for  
357 CI participants than for NHY and NHE participants (CI: 14.06; NHY: 1.06; NHE: 0.95) without  
358 significant difference between NHY and NHE participants. This shows that CI participants stay  
359 largely impaired in  $f_0$  discrimination.

---

360 Insert Figure 4 about here

---

361

362

363 **4. Discussion**

364 The goal of this study was to investigate sensory-motor linkage in speech production on post-  
365 lingually deaf cochlear-implanted patients, in comparison with young and elderly adults with normal  
366 hearing, through the abilities of  $f_0$  convergence and imitation. To this aim, we used a paradigm of  
367 intentional and non-intentional imitation of vowels with modified fundamental frequency.

368 **4.1. Young adults**

369 We firstly replicated the previous findings by Sato et al. (2013) and Garnier et al. (2013) for young  
370 adults with normal hearing. Imitative changes towards the acoustic target were indeed observed in  
371 both tasks, with stronger convergence in the task with direct instruction than in the convergence task.  
372 It is likely that this is mainly due to the effect of the explicit instruction, though we cannot discard the  
373 possibility that order intervened here, with a trend to increase convergence in the second task just  
374 becomes it comes after the first one. As in the two previous studies, no correlation was found between  
375 the two tasks, that is individual  $f_0$  changes in the convergence task were not related to those observed  
376 in the imitative task among subjects. This lack of correlation could be due to ceiling effects provided  
377 by the very high degree of convergence in the imitative task and the very low variability of these  
378 imitative changes for almost all participants. These results confirm that convergence can occur even in  
379 a non-interactive situation of communication, in line with previous studies (Goldinger & Azuma,  
380 2004; Gentilucci & Cattaneo, 2005; Delvaux & Soquet, 2007; Garnier et al. 2013; Sato et al. 2013).  
381 Such non-interactive phonetic convergence effects likely rely on low-level sensory-motor mechanisms  
382 described in the Introduction section, according to which participants would first analyze the target  
383 stimulus to elaborate their motor and sensory goals, resulting in an adaptation of their production to  
384 the acoustic target.

385 **4.2. Elderly population**

386 The results for elderly participants with normal hearing were similar to those of young adults. Indeed,  
387 NHE imitated the auditory target in the imitative task and converged towards the acoustic target in a  
388 probably subconscious way in the convergence task. In addition, as for young participants with normal

389 hearing, no correlation was observed between the two tasks. The observed convergence effects in both  
390 voluntary and unintentional imitation in this elderly population, as high as in young subjects, suggests  
391 that sensory-motor relationships during speech production are efficient in these participants, which is  
392 in line with previous studies suggesting that sensory-motor relationships are still active and efficient in  
393 seniors (Liu et al., 2010, 2011, Tremblay et al. 2013).

#### 394 **4.3. Post-lingually deaf cochlear-implanted participants**

395 In spite of their strongly impaired ability to discriminate frequencies, as displayed by JND values 10 to  
396 15 times larger than those of subjects with normal hearing, it is striking that CI participants were also  
397 found to be able to imitate and to converge towards the acoustic targets. This suggests that they are  
398 able to estimate the pitch of a vowel target to a certain extent and to monitor their own vocal source to  
399 attempt to get closer to this target. Even more strikingly, they do it even in the convergence task where  
400 no explicit imitation instruction was provided – hence the importance to know that all participants  
401 performed this task before the “imitative task”, reducing the risk that they could have considered that  
402 imitation mattered in the task.

403 Although these results demonstrate that cochlear-implanted patients are able to perceive and imitate  
404 the fundamental frequency of the vowel targets, no correlation was however found between their  
405 imitative abilities and their auditory discrimination scores in the JND test. This suggests that their  
406 imitation abilities are not related to their auditory capacity of perceiving frequencies in a simple way.  
407 Nevertheless, cochlear-implanted participants showed less convergence and imitation than adults with  
408 normal hearing whatever their age, and higher pitch variability between participants. At this stage, it is  
409 unclear if pitch variability in production is due to inaccurate estimation of the target pitch or to  
410 degraded voice control, but interestingly, the lack of correlation between convergence and JND  
411 suggests that apart from sensory ability, the sensory-motor relationships could differ from one subject  
412 to the other. Considering the importance and efficiency of somatosensory control for deaf subjects  
413 (Nasir & Ostry, 2008), together with the range of inter-individual variations in the relative importance  
414 of auditory and somatosensory control (Lametti et al., 2012; Feng et al., 2011), it is likely that  
415 variations in somatosensory dependence could be at hand in these data.

416 Contrary to Blamey et al. (2012) who showed that deafness duration and age at deafness onset are  
417 two factors influencing speech perception performance, we did not find any correlation between these  
418 factors and CI abilities to imitate or to converge towards an acoustic target. However, it is important to  
419 note that in the Blamey et al. study (2012), residual variability not explained by CI- and deafness-  
420 related factors was considerable (representing 90% of variation in their study). It seems therefore  
421 likely that the limited sample of our CI participants impeded any clear correlation with CI- and  
422 deafness-related characteristics to appear in the present study.

423 Finally, contrary to NH participants, CI participant's responses were not always correct (with up to  
424 13-15% errors in each task), which suggests some auditory difficulties in the decoding of isolated  
425 vowels. Actually, it is known that CI listeners need formant transition to accurately categorize a  
426 vowel, as shown by Hanna (2011). The isolated vowels used in our study are hence probably difficult  
427 to categorize efficiently by the CI participants. Moreover, convergence effects tested in error trials  
428 appear to be similar to those of correct trials, which suggests that timbre and pitch estimations seem to  
429 be relatively independent for CI participants. Nevertheless, the major result of this study is that a  
430 number of CI participants have already recovered a good ability to associate auditory with motor  
431 parameters even a short time after implantation (with significant convergence after only one, two or  
432 three months of implantation for some subjects). Indeed, results displayed both automatic and  
433 conscious imitation abilities in post-lingually deaf cochlear-implanted patients, displaying their ability  
434 to adapt their production to the environment, which is uniquely due to a functional sensory-motor  
435 linkage. This should be crucial for the retuning of their speech production system (see Perkell, Lane,  
436 Svirsky, & Webster, 1992), enabling them to improve their internal model (Perkell et al., 2000).  
437 Interestingly, recovering perceptuo-motor abilities could also be of importance for their speech  
438 perception abilities, considering the proposals about the role of the motor system in speech perception  
439 (e.g. Skipper, vanWassenhove, Nusbaum, & Small, 2007; Schwartz, Basirat, Ménard, & Sato, 2012).

440

## 441 **5. Conclusion**

442 The present study reports convergence and imitation abilities in adults with normal hearing, both  
443 young and elderly, and in cochlear-implanted subjects. Indeed, in the three groups, we obtain

444 significant and generally strong imitation of the acoustic target during a task with explicit instruction,  
445 and also convergence effects in a task without direct instruction. These results first confirm that  
446 sensory-motor relationships during speech production are still efficient in elderly subjects, with  
447 basically no difference with younger ones. Moreover, crucially, we also show for the first time that  
448 cochlear-implanted participants have the ability to converge to an acoustic target, intentionally and  
449 unintentionally. Therefore, they have recovered significant sensory-motor abilities, which could be  
450 crucial for improving both their speech production and speech perception abilities.

451

452

453 **6. Acknowledgments**

454 This work was supported by the French National Research Agency (ANR) through funding from  
455 the PlasMody project (Plasticity and Multimodality in Oral Communication for the Deaf).

456



457 **References**

- 458 Aubanel V. (2011). Variation phonologique régionale en interaction conversationnelle. *Doctoral*  
459 *Dissertation, Aix-Marseille University.*
- 460 Babel M. (2009). Phonetic and social selectivity in speech accommodation. *Doctoral Dissertation,*  
461 *University of California, Berkeley.*
- 462 Bilodeau-Mercure, M., Ouellet, C., & Tremblay, P. (2015). Vieillesse oro-faciale et production de  
463 la parole. *Affiche présentée au Congrès québécois de recherche en adaptation-réadaptation.*  
464 *Boucherville, Qc, Canada.*
- 465 Blamey, P., Artieres, F., Baskent, D., Bergeron, F., Beynon, A., Burke, E., Dillier, N., Dowell, R.,  
466 Fraysse, B., Gallégo, S., Govaerts, P.J., Green, K., Huber, A.M., Kleine-Punte, A., Maat, B., Marx,  
467 M., Mawman, D., Mosnier, I., Fitzgerald O'Connor, A., O'Leary, S., Rousset, A., Schauwers, K.,  
468 Skarzynski, H., Skarzynski, P.H., Sterkers, O., Terranti, A., Truy, E., Van de Heyning, P., Venail,  
469 F., Vincent, C., & Lazard, D.S. (2012). Factors affecting auditory performance of postlinguistically  
470 deaf adults using cochlear implant: an update with 2251 patients. *Audiology & Neurology*, 18, 36-  
471 47.
- 472 Blumstein, S. E. & Stevens, K. N. (1979). Acoustic invariance in speech production: Evidence from  
473 measurements of the spectral characteristics of stop consonants. *Journal Of the Acoustical Society*  
474 *of America*, 66, 1001-1017
- 475 Boersma, P. (2001). Praat, a system for doing phonetics by computer, *Glott International*, 5(9/10),  
476 341-345.
- 477 Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: an overview. *Phonetica*, 49(3-4),  
478 155-180.
- 479 Cienkowski, K. M. & Carney, A. E. (2002). Auditory-visual speech perception and aging. *Ear and*  
480 *Hearing*, 23(5), 439-449.
- 481 Cienkowski, K. M., & Vasil-Dilaj, K. (2010). The Role of Central Auditory and Cognitive Processing  
482 Ability in Speech Perception for Elderly Listeners. *Perspectives on Hearing and Hearing*  
483 *Disorders Research and Diagnostics*, 14(2), 30.

- 484 Delvaux, V. & Soquet, A. (2007). The influence of ambient speech on adult speech productions  
485 through unintentional imitation. *Phonetica*, 64, 145–173.
- 486 Feng, Y., Gracco, V. L., & Max, L. (2011). Integration of auditory and somatosensory error signals in  
487 the neural control of speech movements. *Journal of neurophysiology*, 106(2), 667-679.
- 488 Fowler, C. A. (1986). An event approach to the study of speech perception for a direct-realist  
489 perspective. *Journal of Phonetics*, 14.
- 490 Füllgrabe, C. (2013). Age-Dependent Changes in Temporal-Fine-Structure Processing in the Absence  
491 of Peripheral Hearing Loss. *American Journal of Audiology*, 22(2), 313.
- 492 Füllgrabe C, Rosen S (2016) On the (un)importance of working memory in speech-in-noise processing  
493 for listeners with normal hearing thresholds. *Frontiers in Psychology*
- 494 Garnier, M., Lamalle, L., & Sato, M. (2013). Neural correlates of phonetic convergence and speech  
495 imitation. *Frontiers in psychology*, 4, 600.
- 496 Gentilucci, M. & Cattaneo, L. (2005). Automatic audiovisual integration in speech perception.  
497 *Experimental Brain Research*, 167:66-75.
- 498 Goldinger, S.D. & Azuma, T. (2004). Episodic memory reflected in printed word naming.  
499 *Psychonomic Bulletin and Review*, 11(4): 716-722.
- 500 Gelfand, S. A., Piper, N., & Silman, S. (1985). Consonant recognition in quiet as a function of aging  
501 among normal hearing subjects. *The Journal of the Acoustical Society of America*, 78(4),  
502 1198-1206.
- 503 Giles H., Coupland N., & Coupland J. (1991). Accommodation theory: Communication, context, and  
504 consequence. In Giles H, Coupland N, Coupland J (Editors) *Contexts of Accommodation:*  
505 *Developments in Applied Sociolinguistic*, Cambridge University Press pp. 1-68.
- 506 Guenther, F. H., Hampson, M., & Johnson, D. (1998). A theoretical investigation of reference frames  
507 for the planning of speech movements. *Psychological Review*, 105(4), 611-633.
- 508 Guenther, F.H. and Vladusich, T. (2012). A neural theory of speech acquisition and  
509 production. *Journal of Neurolinguistics*. 25, pp. 408-422.
- 510 Hanna, N. (2011). Vowel identification by cochlear implant users: comparison on vowel edges and  
511 vowel centers. Honors College Thesis, University of South Florida.

- 512 Kishon-Rabin, L., Taitelbaum, R., Tobin, Y., & Hildesheimer, M. (1999). The effect of partially  
513 restored hearing on speech production of postlingually deafened adults with multichannel cochlear  
514 implants. *The Journal of the Acoustical Society of America*, 106(5), 2843-2857.
- 515 Lametti, D.R., Sazzad, Nasir, R.M., & Ostry, D.J. (2012). Sensory preference in speech production  
516 revealed by simultaneous alteration of auditory and somatosensory feedback. *Journal of*  
517 *Neuroscience*, 32(27), 9351-9358.
- 518 Lametti, D.R., Rochet-Capellan, A., Neufeld, E., Shiller, D.M., & Ostry, D.J. (2014) Plasticity in the  
519 human speech motor system drives changes in speech perception. *Journal of Neuroscience*,  
520 30;34(31), 10339-46.
- 521 Lane, H. & Webster, J.W. (1991). Speech deterioration in postlingually deafened adults. *The Journal*  
522 *of the Acoustical Society of America*, 89, 859-866.
- 523 Lane, H., Denny, M., Guenther, F. H., Matthies, M. L., Menard, L., Perkell, J. S., Stockmann, E.,  
524 Tiede, M., Vic, J. & Zandipour, M. (2005). Effects of bite blocks and hearing status on vowel  
525 production. *The Journal of the Acoustical Society of America*, 118(3 ; 1), 1636-1646.
- 526 Langereis, M.C., Bosman, A.J., Van Olphen, A.F., & Smoorenburd, G.F. (1998). Effect of cochlear  
527 implantation on voice fundamental frequency in post-lingually deafened adults. *Audiology*, 37,  
528 219-230.
- 529 Lelong, A. (2012). *Convergence Phonétique en Interaction*. Doctoral Dissertation, Grenoble  
530 University.
- 531 Liberman, A. M., & Mattingly, I. G. (1985). The Motor Theory of Speech Perception Revised.  
532 *Cognition*, 21(1), 1-36.
- 533 Lindblom, B., & Maddieson, I. (1988). Phonetic Universals in Consonant Systems. In *Language,*  
534 *Speech and Mind* (p. 62-78). London: Routledge.
- 535 Lindblom, B. (1990). Explaining Phonetic Variation: A Sketch of the H&H Theory. In W. J.  
536 Hardcastle & A. Marchal (éd.), *Speech Production and Speech Modelling* (p. 403-439). Springer  
537 Netherlands.
- 538 Linville, S. E., & Fisher, H. B. (1985). Acoustic characteristics of women's voices with advancing  
539 age. *Journal of Gerontology*, 40(3), 324-330.

- 540 Linville, S. E., & Korabic, E. W. (1987). Fundamental frequency stability characteristics of elderly  
541 women's voices. *The Journal of the Acoustical Society of America*, 81(4), 1196-1199.
- 542 Liu, H., Wang, E. Q., Chen, Z., Liu, P., Larson, C. R., & Huang, D. (2010). Effect of tonal native  
543 language on voice fundamental frequency responses to pitch feedback perturbations during  
544 sustained vocalizations. *The Journal of the Acoustical Society of America*, 128(6), 3739-3746.
- 545 Liu, H., Meshman, M., Behroozmand, R., & Larson, C. R. (2011). Differential effects of perturbation  
546 direction and magnitude on the neural processing of voice pitch feedback. *Clinical  
547 neurophysiology: official journal of the International Federation of Clinical Neurophysiology*,  
548 122(5), 951-957.
- 549 Liu, P., Chen, Z., Jones, J. A., Huang, D., & Liu, H. (2011). Auditory Feedback Control of Vocal Pitch  
550 during Sustained Vocalization: A Cross-Sectional Study of Adult Aging. *PLoS ONE*, 6(7), e22791.
- 551 Lortie, C. L., Thibeault, M., Guitton, M. J., & Tremblay, P. (2015). Effects of age on the amplitude,  
552 frequency and perceived quality of voice. *Age (Dordrecht, Netherlands)*, 37(6), 117.
- 553 Nasir, S. M., & Ostry, D. J. (2008). Speech motor learning in profoundly deaf adults. *Nature  
554 Neuroscience* 11, 1217 - 1222 (2008).
- 555 Nguyen, N., Sato, M., Postma, M., Pardo, J., Babel, M. & Postma, E. (2013). Speech imitation: the  
556 cognitive underpinnings of adaptive vocal behavior. *Frontiers in Psychology*.
- 557 Max, L., Daniels, J., Curet, K., & Cronin, K. (2008) Modulation of auditory and somatosensory  
558 processing during the planning of speech movements *Proceedings of the 8th International Seminar  
559 on Speech Production*, 41-44, Strasbourg, France.
- 560 Ménard, L., Polak, M., Denny, M., Burton, E., Lane, H., Matthies, M. L., ... Vick, J. (2007).  
561 Interactions of speaking condition and auditory feedback on vowel production in postlingually deaf  
562 adults with cochlear implants. *The Journal of the Acoustical Society of America*, 121(6),  
563 3790-3801.
- 564 Moore, D. R., & Fullgräbe, C. (2013). Cognitive contributions to hearing in older people. *Journal of  
565 Hearing Science*, 2(4), 58-60.
- 566 Morgan, E. E., & Rastatter, M. (1986). Variability of voice fundamental frequency in elderly female  
567 speakers. *Perceptual and Motor Skills*, 63(1), 215-218.

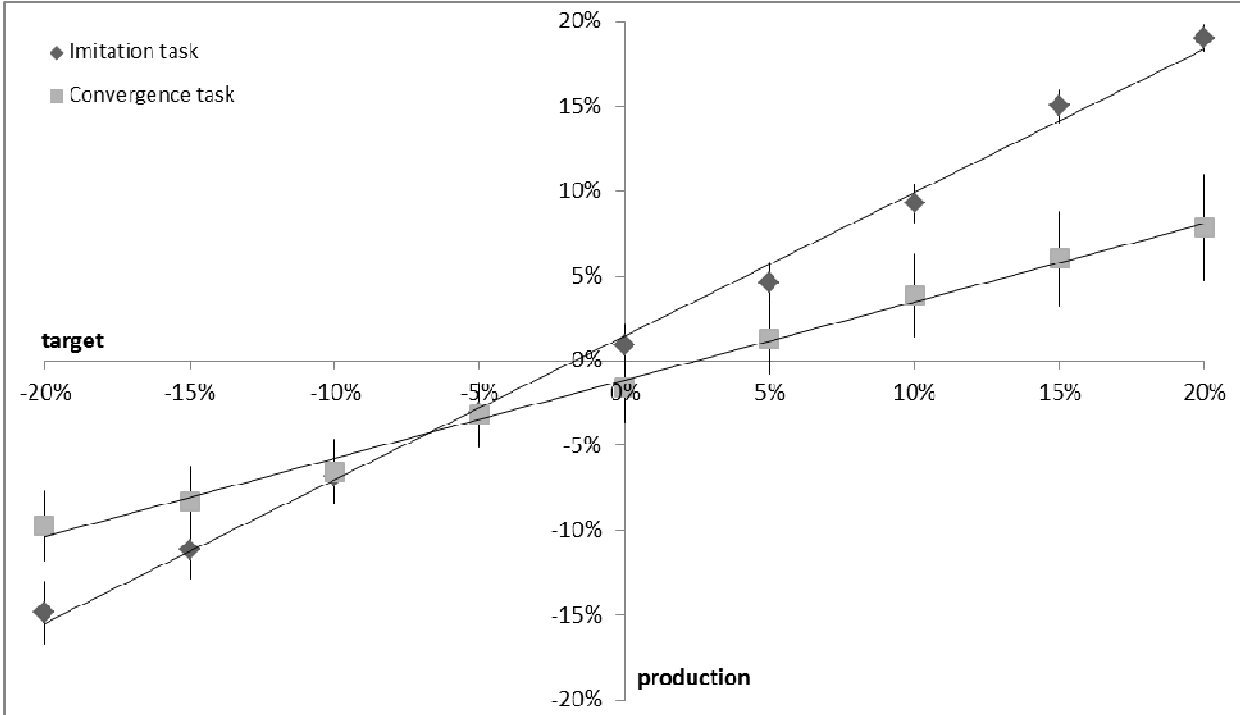
- 568 Ohde, R. N., & Abou-Khalil, R. (2001). Age differences for stop-consonant and vowel perception in  
569 adults. *The Journal of the Acoustical Society of America*, 110(4), 2156-2166.
- 570 Perkell, J.S., Lane, H., Svirsky, M., & Webster, J. (1992). Speech of cochlear implant patients: A  
571 longitudinal study of vowel production. *The Journal of the Acoustical Society of America*, 91,  
572 2961-2979.
- 573 Perkell, J. S., Matthies, M. L., Svirsky, M. A., & Jordan, M. I. (1995). Goal-based speech motor  
574 control: A theoretical framework and some preliminary data. *Journal of Phonetics*, 23(1), 23-35.
- 575 Perkell, J.S., Guenther, F.H., Lane, H., Matthies, L.M., Perrier, P., Vick, J., Wilhelms-Tricarico, R., &  
576 Sandipour, M. (2000). A theory of speech motor control and supporting data from speakers with  
577 normal hearing and with profound hearing loss. *Journal of phonetics*, 28, 233-272.
- 578 Perrier, P. (2005). Control and representations in speech production. *ZAS Papers in Linguistics*, 40,  
579 109-132.
- 580 Rouger, J., Lagleyre, S., Démonet, J.-F., Fraysse, B., Deguine, O., & Barone, P. (2012). Evolution of  
581 crossmodal reorganization of the voice area in cochlear-implanted deaf patients. *Human Brain*  
582 *Mapping*, 33(8), 1929-1940.
- 583 Russell, A., Penny, L., & Pemberton, C. (1995). Speaking fundamental frequency changes over time  
584 in women: a longitudinal study. *Journal of Speech and Hearing Research*, 38(1), 101-109.
- 585 Saltzman, E. (1986). Task Dynamic Coordination of the Speech Articulators: A preliminary model.  
586 *Experimental Brain Research Series*, 15, 130 - 144.
- 587 Sato, M., Schwartz, J.-L. & Perrier, P. (2014). Phonemic auditory and somatosensory goals in speech  
588 production. *Language, Cognition and Neuroscience*, 29(1): 41-43.
- 589 Sato, M., Grabski, K., Garnier, M., Granjon, L., Schwartz, J.L., & Nguyen, N. (2013). Converging to a  
590 common speech code: imitative and perceptuo-motor recalibration processes in speech production.  
591 *Frontiers in Psychology*, 4, 422.
- 592 Savariaux, C., Perrier, P. & Orliaguet, J. P. (1995) Compensation strategies for the perturbation of the  
593 rounded vowel [u] using a lip tube: A study of the control space in speech production. *Journal of*  
594 *the Acoustical Society of America*, 98 (5), 2428-2442.

- 595 Schwartz, J.L., Basirat, A., Ménard, L. & Sato, M. (2012). The Perception-for-Action-Control Theory  
596 (PACT): A perceptuo-motor theory of speech perception. *Journal of Neurolinguistics*, 25, 336-354.
- 597 Shiller, D.M., & Rochon, A.M. (2014) Auditory-perceptual learning improves speech motor  
598 adaptation in children. *Journal of Experimental Psychology. Human Perception and Performance*,  
599 40(4):1308-15.
- 600 Skipper, J. I., vanWassenhove, V., Nusbaum, H. C., & Small, S. L. (2007). Hearing lips and seeing  
601 voices: How cortical areas supporting speech production mediate audiovisual speech perception.  
602 *Cerebral Cortex*, 17, 2387-2399.
- 603 Skipper, J. I., Devlin, J. T., & Lametti, D. R. (2016). The hearing ear is always found close to the  
604 speaking tongue: Review of the role of the motor system in speech perception. *Brain Lang*, 164,  
605 77-105.
- 606 Stevens, K. N. (1972). The quantal nature of speech: Evidence from articulatory-acoustic data. In  
607 *Human communication: A unified view* (p. 51-66). EE David et PB Denes editors.
- 608 Stevens, K. N., & Blumstein, S. E. (1978). Invariant cues for place of articulation in stop consonants.  
609 *The Journal of the Acoustical Society of America*, 64(5), 1358-1368.
- 610 Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics*, 17, 3-46.
- 611 Tremblay, S., Shiller, D.M., & Ostry, D.J. (2003) Somatosensory basis of speech production. *Nature*,  
612 19;423(6942):866-9.
- 613 Tremblay, P., Dick, A. S., & Small, S. L. (2013). Functional and structural aging of the speech  
614 sensorimotor neural system: functional magnetic resonance imaging evidence. *Neurobiology of*  
615 *Aging*, 34(8), 1935-1951.
- 616 Turgeon, C., Prémont, A., Trudeau-Fisette, P., & Ménard, L. (2015). Exploring production-perception  
617 relationships in normal hearing and cochlear implant adults: a lip-tube perturbation study. *Clinical*  
618 *Linguistics and Phonetics*, 29(5), 378-400.
- 619 Vinay, S.N. & Moore, B.C. (2010). Effects of the use of personal music players on amplitude  
620 modulation detection and frequency discrimination. *The Journal of the Acoustical Society of*  
621 *America*, 128(6), 3634-41.
- 622

623

624 **Figures**

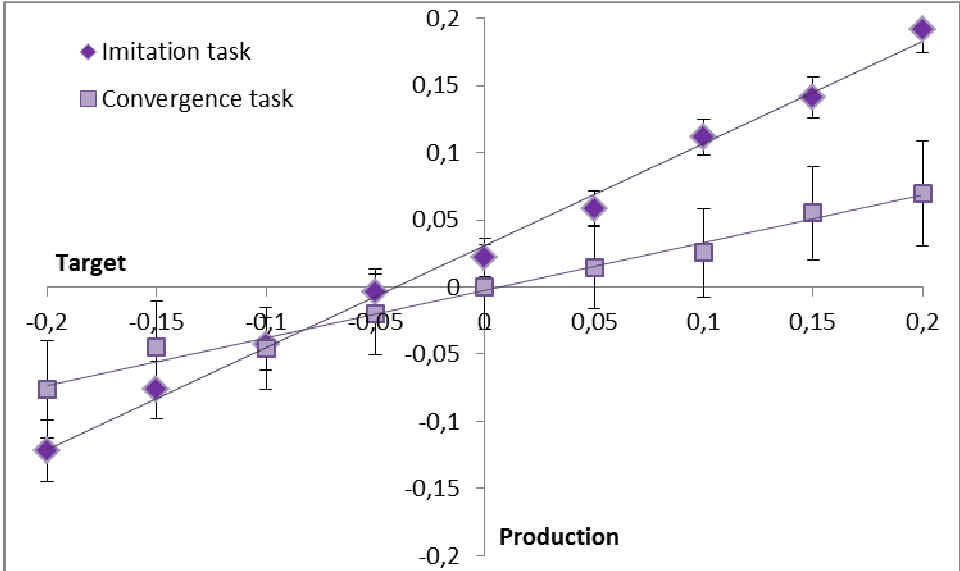
625 Figure 1: Phonetic convergence and voluntary imitative changes on fundamental frequency observed  
626 in young participants with normal hearing (X axis: deviation percentage with respect to mean  
627 participant'  $f_0$ , Y axis: mean production for all participants, error bars corresponding to standard  
628 deviation



629

630

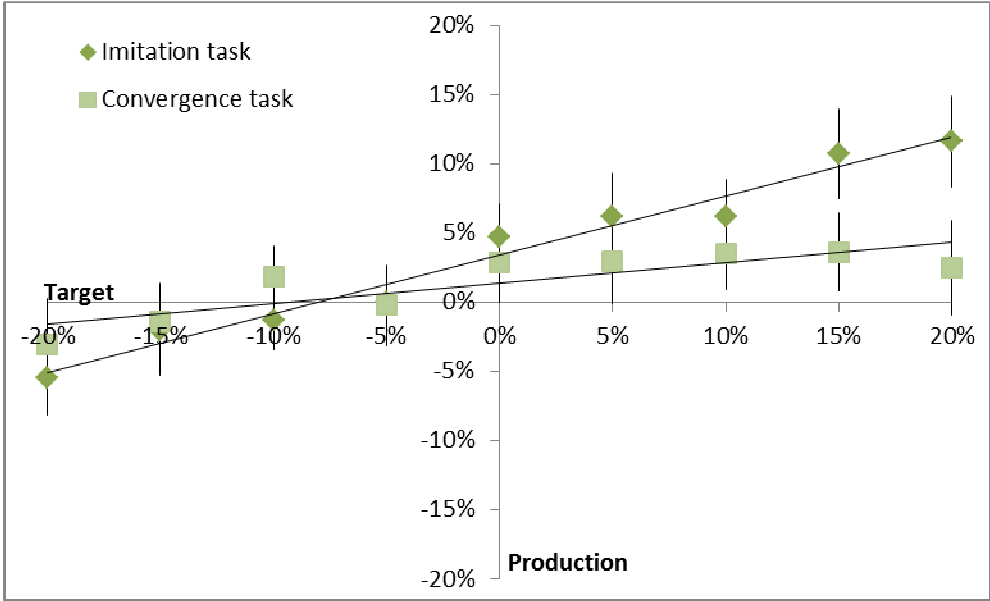
631  
632 Figure 2: Phonetic convergence and voluntary imitative changes observed on fundamental frequency  
633 in elderly participants with normal hearing (X axis: deviation percentage with respect to mean  
634 participant'  $f_0$ , Y axis: mean production for all participants, error bars corresponding to standard  
635 deviation)



636  
637

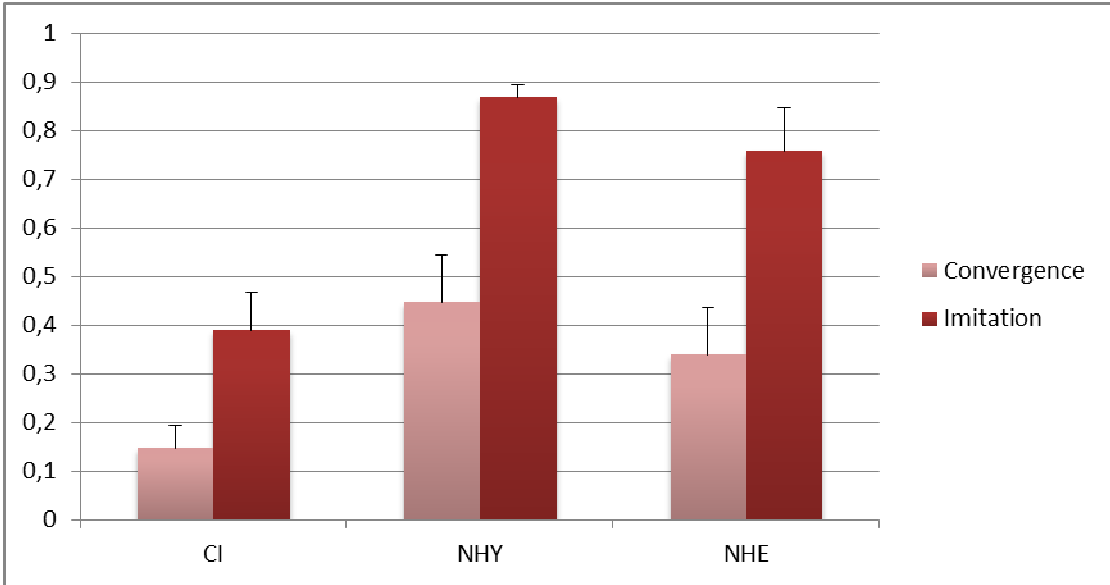


638  
639 Figure 3: Phonetic convergence and voluntary imitative changes observed on fundamental frequency  
640 in cochlear-implanted participants (X axis: deviation percentage with respect to mean participant'  $f_0$ , Y  
641 axis: mean production for all participants, error bars corresponding to standard deviation)  
642



643

644  
645 Figure 4: Phonetic convergence and voluntary imitative slope changes on fundamental frequency in  
646 young (NHY) and elderly (NHE) participants with normal hearing and cochlear-implanted participants  
647 (CI) (error bars corresponding to standard deviation)



648  
649  
650  
651  
652

653

654 **Tables**

655 Table 1: Characteristics of participants with cochlear implants

	<b>Gender</b>	<b>Age (y.)</b>	<b>Age at onset of deafness (y.)</b>	<b>Hearing aid</b>	<b>Duration of deafness</b>	<b>Duration of CI experiment</b>
<b>CI1</b>	M	65	7	No	58 years	1 month
<b>CI2</b>	M	56	20	Yes	35 years	3 months
<b>CI3</b>	F	66	32	Yes	25 years	9 years
<b>CI4</b>	M	60	59	Yes	1 month	1 years 4 months
<b>CI5</b>	F	43	20	Yes	13 years	2 months
<b>CI7</b>	M	27	25	Yes	2 months	2 years 6 months
<b>CI8</b>	F	67	65	Yes	2 years	7 months
<b>CI9</b>	M	72	40	Yes	30 years	5 months
<b>CI10</b>	M	76	48	No	27 years	3 years 4 months
<b>CI11</b>	M	57	48	Yes	8 years	10 months

656

657

658

659

660 Table 2: Determination coefficients in both tasks for each participant of the cochlear-implanted group  
 661 (CI), normal-hearing elderly adults group (NHE) and normal-hearing young adults (NHY). For the  
 662 cochlear-implanted group, determination coefficients measured on errors trials only are indicated in  
 663 parentheses for participants with at least 15% errors (enabling to obtain a sufficient number of errors  
 664 for computing correlations).

CI			NHE		NHY	
convergence	imitation	JND	convergence	imitation	convergence	imitation
0,25 (0,33)	0,51(0,32)	15,22	0,72	0,95	0,96	0,95
0,51	0,70	45,35	0,36	0,28	0,99	1,00
0,78 (0,51)	0,77 (0,73)	1,71	0,98	0,99	0,69	0,99
0,04 (0,17)	0,39	9,76	0,49	0,93	0,99	0,97
0,27	0,94	1,79	0,99	0,96	0,87	0,99
0,12	0,77 (0,41)	4,55	0,88	0,97	0,71	0,91
0,49	0,62	18,52	0,50	0,98	0,05	0,97
0,75	0,92	10,55	0,92	0,98	0,96	0,99
0,11 (0,11)	0,22 (0,03)	16,21	0,99	0,99	0,99	0,99
0,60	0,09	28,54	0,35	0,91	0,10	0,98
					0,95	0,94
					0,98	0,99
					0,84	0,95
					0,63	0,97
					0,97	0,99

665

666

667