



**HAL**  
open science

# The role of orofacial somatosensory inputs in speech perception and production in individuals with hearing-impairment

Monica Ashokumar

► **To cite this version:**

Monica Ashokumar. The role of orofacial somatosensory inputs in speech perception and production in individuals with hearing-impairment. Psychology. Université Grenoble Alpes [2020-..], 2024. English. NNT: 2024GRALS008 . tel-04660602

**HAL Id: tel-04660602**

**<https://theses.hal.science/tel-04660602v1>**

Submitted on 24 Jul 2024

**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.

THÈSE

Pour obtenir le grade de

**DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES**

École doctorale : ISCE - Ingénierie pour la Santé la Cognition et l'Environnement

Spécialité : PCN - Sciences cognitives, psychologie et neurocognition

Unité de recherche : Grenoble Images Parole Signal Automatique

**Le rôle des entrées somatosensorielles orofaciales dans la perception et la production de la parole chez les personnes malentendantes**

**The role of orofacial somatosensory inputs in speech perception and production in individuals with hearing-impairment**

Présentée par :

**Monica ASHOKUMAR**

Direction de thèse :

**Jean-luc SCHWARTZ**

DIRECTEUR DE RECHERCHE EMERITE, CNRS DELEGATION ALPES

Directeur de thèse

**Takayuki ITO**

CHARGE DE RECHERCHE HDR, CNRS DELEGATION ALPES

Co-directeur de thèse

Rapporteurs :

**LUCIE MENARD**

FULL PROFESSOR, UNIVERSITE DU QUEBEC A MONTREAL

**DOUGLAS SHILLER**

FULL PROFESSOR, UNIVERSITE DE MONTREAL

Thèse soutenue publiquement le **30 avril 2024**, devant le jury composé de :

**PASCAL PERRIER,**

PROFESSEUR DES UNIVERSITES, GRENOBLE INP

Président

**JEAN-LUC SCHWARTZ,**

DIRECTEUR DE RECHERCHE EMERITE, CNRS DELEGATION ALPES

Directeur de thèse

**ITO TAKAYUKI,**

CHARGE DE RECHERCHE HDR, CNRS DELEGATION ALPES

Co-directeur de thèse

**LUCIE MENARD,**

FULL PROFESSOR, UNIVERSITE DU QUEBEC A MONTREAL

Rapporteuse

**DOUGLAS SHILLER,**

FULL PROFESSOR, UNIVERSITE DE MONTREAL

Rapporteur

**AXELLE CALCUS,**

ASSISTANT PROFESSOR, UNIVERSITE LIBRE DE BRUXELLES

Examinatrice



# **Funding Statement**

This work was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No 860755 (Comm4CHILD project) and by National Institute on Deafness and Other Communication Disorders Grant R01-DC017439.

# Abstract

Hearing-impaired (HI) individuals can perceive speech sounds with the help of hearing aids (HA) and/or cochlear implants (CI). Hence, factors concerning those devices, such as implantation age, type of device used and the usage duration of device, can affect the development of their speech perception ability. Recent findings suggest the involvement of orofacial somatosensory inputs in speech perception. The development of speech perception abilities in HI individuals might be related to the acquisition of this auditory-somatosensory integration. The aim of this thesis is to investigate the role of orofacial somatosensory inputs in speech perception and production in HI and normal-hearing (NH) individuals.

In a series of studies, orofacial somatosensory stimulation associated with facial skin deformation was used as a main tool. This somatosensory stimulation induces perceptual modulation, which is represented by the shift of categorical boundary in vowel perception, when the stimulation is applied with the speech sound. This thesis mainly focused on how this somatosensory effect in perception is related to the production and perception abilities.

The first study examined a relationship between the orofacial somatosensory effect in speech perception and production performance in NH individuals. Amplitude of somatosensory effect in speech perception in a given contrast between two target vowels was positively correlated with the acoustic distance between the target vowels in production. This suggests that acquiring a precise performance in speech production can be associated with greater role of the somatosensory modality in auditory-somatosensory integration.

The second study examined whether speech perceptual training with somatosensory stimulation modifies speech production. The perceptual training involved a vowel identification task with or without somatosensory stimulation synchronous with the sound. Following this perceptual training, production performance of the corresponding vowel sounds was reliably changed in a specific way when training involved the somatosensory stimulation. The results suggest that receiving specific associations of auditory and somatosensory inputs can play a role in speech motor learning.

The third study examined whether HI individuals also showed the somatosensory effect in speech perception and how this effect is related to their hearing ability. Hearing ability of HI individuals was evaluated as hearing threshold in noise, which was obtained using a speech-in-noise test, and also included factors concerning their hearing device. The results showed that the somatosensory effect was varied depending on their hearing ability. Long-time HA users

who also had poor hearing in noise, showed large somatosensory effect in speech perception. In contrast, CI users, who also had rather good hearing in noise, showed little or no changes in speech perception due to somatosensory stimulation. The development of hearing ability based on hearing experience can affect the acquisition of auditory-somatosensory integration.

The fourth study examined whether HI individuals showed sensorimotor adaptation using altered auditory feedback. The results showed that long-time CI users showed an adaptation but HA users did not. Since the long-time CI users were able to auditorily identify small differences in synthesized auditory stimuli, similarly to NH individuals, speech sensorimotor adaptation in HI individuals can be similar to the one in NH individuals when they can receive proper auditory feedback by using hearing devices.

In all, the development of auditory-somatosensory integration can be related to the development of perception and production abilities. Somatosensory inputs can play an important role in the development of speech perception in both NH and HI individuals.

# Résumé

Les personnes malentendantes (HI) peuvent percevoir les sons de parole à l'aide d'appareils auditifs (HA) et/ou d'implants cochléaires (CI). Différents facteurs concernant ces dispositifs, tels que l'âge d'implantation, le type de dispositif utilisé et la durée d'utilisation, peuvent affecter le développement des capacités de perception de la parole. Des découvertes récentes suggèrent l'implication des entrées somatosensorielles orofaciales dans la perception de parole. Le développement des capacités perceptives pourrait être lié à l'acquisition de cette intégration auditive-somatosensorielle. L'objectif de cette thèse est d'étudier le rôle des entrées somatosensorielles orofaciales dans la perception et la production de la parole chez les individus HI et ayant une audition normale (NH).

Dans une série d'études, nous explorons les effets d'une stimulation somatosensorielle orofaciale produite par une déformation de la peau du visage. Cette stimulation, lorsqu'elle est appliquée en même temps que le son de parole, induit une modulation perceptuelle, correspondant au déplacement de la frontière catégorielle dans la perception des voyelles. Cette thèse s'est principalement concentrée sur la façon dont cet effet somatosensoriel orofacial est lié aux capacités de production et de perception.

La première étude a examiné une relation entre cet effet et les performances de production chez les individus NH. L'amplitude de l'effet sur un contraste perceptif entre deux voyelles cibles apparaît positivement corrélé à la distance acoustique entre ces voyelles en production. Ceci suggère que l'acquisition de performances précises en production pourrait être associée à un plus grand rôle de la modalité somatosensorielle en perception.

La deuxième étude a examiné si l'entraînement à la perception de parole avec stimulation somatosensorielle modifiait la production. L'entraînement perceptuel portait sur une tâche d'identification de voyelles avec ou sans stimulation somatosensorielle. La production des voyelles a été modifiée de manière spécifique lorsque l'entraînement comportait la stimulation somatosensorielle. Ceci suggère que l'association d'entrées auditives et somatosensorielles peut jouer un rôle dans l'apprentissage moteur de la parole.

La troisième étude a examiné si les individus HI présentaient également un effet somatosensoriel sur la perception de parole et comment cet effet était lié à leur capacité auditive, évaluée par le seuil d'audition dans le bruit, obtenu à l'aide d'un test de parole dans le bruit, et par des facteurs décrivant leur aide auditive. Les résultats ont montré que l'effet somatosensoriel variait en fonction de leur capacité auditive. Les utilisateurs de longue date

d'HA, avec une faible audition dans le bruit, ont montré un effet somatosensoriel important. En revanche, les utilisateurs d'IC, avec une relativement bonne audition dans le bruit, n'ont montré que peu ou pas d'effet de la stimulation somatosensorielle. Le développement de la capacité auditive basé sur l'expérience auditive peut affecter l'acquisition de l'intégration auditive-somatosensorielle.

La quatrième étude a examiné si les individus HI présentaient une adaptation sensorimotrice en utilisant un feedback auditif altéré. Les utilisateurs de CI ont montré une adaptation, mais pas les utilisateurs de HA. Les utilisateurs de longue date de CI sont capables de percevoir de petites différences auditives entre stimuli synthétisés, comme les individus NH, ce qui montre que l'adaptation sensorimotrice de la parole chez les individus HI peut être similaire à celle des individus NH lorsqu'ils reçoivent un retour auditif approprié.

Ainsi, le développement de l'intégration auditive-somatosensorielle apparaît lié au développement des capacités de perception et de production. Les entrées somatosensorielles peuvent jouer un rôle important dans le développement de la perception de la parole chez les individus NH et HI.

# Acknowledgements

My first and foremost gratitude to the almighty God who has brought me so far and will lead me further, without whom I cannot imagine completing this PhD.

My journey throughout the PhD was possible due to the valuable support of the wonderful people who have assisted and inspired me along the way.

I am heavily indebted to both my supervisors- Takayuki Ito and Jean-Luc Schwartz for their patience and wisdom in guiding me through this journey. I'm grateful for their quick responses and expertise that enabled me to complete my PhD successfully.

I would like to thank my CSI members- Brigitte Charlier, H el ene L evenbruck and Paul Deltenre for their enthusiastic and useful inputs during all our CSI meetings.

My heartfelt thanks to the jury members- Axelle Calcus, Douglas Shiller, Pascal Perrier and Lucie M enard who agreed to be a part of my jury, to evaluate my work and wake up at odd times for the defense.

Thanks to Brigitte Charlier, C ecile Colin, Elaine Ng and Hamish Innes-Brown for their valuable insights and guidance during my secondments in Belgium and Denmark.

A big thanks to the colleagues in GIPSA and especially the PCMD team for being so supportive and kind throughout my time here.

I would like to thank the Comm4CHILD ESRs- Irem Adalilar, Yifan Wang, Niloofar Tavakoli, Alina Schulte, Cora Caron, Lyan Porto, Julia Chiossi, Kristina Burum, Elettra Casellato, Nathalie Czeke, Elodie Sabatier, Lucie Van Bogaert and Marie-Joe Kfoury who have always inspired me with their passion towards the field and for all the wonderful memories during the Comm4CHILD workshops.

I would also like to thank our interns- Anesie Martiniani, Clara Fanget, Cl ement Guichet and Guilhem Simeao for their invaluable help with participant recruitment.

A very special thanks to Sanjana Sankar and Tsiky Rakotomalala for being there through all the ups and downs in this journey and always lending an ear during my bouts of anxiety.

Finally, I am grateful to my family and friends for always being there and providing moral support from different parts of the world. I am thankful to my dad for always believing in me and from whom I learnt to never give up. To my mom and grandma, for keeping me in their prayers and encouraging me to do my best. And to my little sister for her constant cheering and belief in me that is a driving force for me to be a better person.



## Table of Contents

<b>List of Figures</b>	<b>10</b>
<b>List of Tables</b>	<b>13</b>
<b>List of Abbreviations</b>	<b>14</b>
<b>Introduction</b>	<b>15</b>
<b>Chapter 1: Hearing impairment and its impact on speech perception and production</b>	<b>16</b>
<b>1.1 Overview</b>	<b>16</b>
<b>1.2 The auditory system and hearing impairment</b>	<b>17</b>
<b>1.3 Diagnosis of hearing impairment</b>	<b>19</b>
<b>1.4 The sources of hearing impairment</b>	<b>21</b>
<b>1.5 Hearing interventions by hearing technologies</b>	<b>22</b>
1.5.1 Hearing aids	23
1.5.2 Cochlear implants	24
1.5.3 Other devices	25
<b>1.6 Impact of hearing impairment on speech perception and production</b>	<b>25</b>
1.6.1 Implantation age	25
1.6.2 Type of hearing devices	26
1.6.3 Usage duration of device	27
<b>Chapter 2: The role of orofacial inputs in speech perception and production</b>	<b>29</b>
<b>2.1 Overview</b>	<b>29</b>
<b>2.2 Links between speech perception and production</b>	<b>29</b>
2.2.1 The involvement of motor system in speech perception	29
2.2.2 Experimental evidence in favor of a functional link between speech perception and production	31
2.2.3 Development of speech perception-production link	33
<b>2.3 Multisensory integration in speech</b>	<b>34</b>
2.3.1 Audio-visual integration and its development	34
2.3.2 Audio-tactile integration	35
<b>2.4 Orofacial somatosensory interaction in speech perception</b>	<b>36</b>
2.4.1 Source of kinaesthetic information from the orofacial system	36
2.4.2 Movement related somatosensory inputs intervene in speech perception	40
<b>2.5 Research questions and Hypothesis</b>	<b>41</b>
<b>Chapter 3: Methodology</b>	<b>43</b>
<b>3.1 Overview</b>	<b>43</b>
<b>3.2 Ethics</b>	<b>43</b>
<b>3.3 Speech Tasks</b>	<b>44</b>
<b>3.4 Somatosensory stimulation associated with facial skin deformation</b>	<b>46</b>
<b>3.5 Modified version of French Digit triplet test</b>	<b>48</b>
<b>3.6 Instrumental configuration of experimental setup</b>	<b>49</b>
<b>Chapter 4: Relationship between speech perception and production performance in individuals with normal-hearing</b>	<b>50</b>
<b>4.1 Introduction</b>	<b>50</b>
<b>4.2 Methods</b>	<b>51</b>
4.2.1 Experimental Procedure	51
4.2.2 Statistical Analysis	51
<b>4.3 Results</b>	<b>51</b>

<b>4.4 Discussion</b>	<b>54</b>
<b>Chapter 5: Speech perceptual training with somatosensory inputs changes speech production in individuals with normal-hearing</b>	<b>56</b>
<b>5.1 Introduction</b>	<b>56</b>
<b>5.2 Methods</b>	<b>57</b>
5.2.1 Experimental Procedure	57
5.2.2 Statistical Analysis	58
<b>5.3 Results</b>	<b>61</b>
<b>5.4 Discussion</b>	<b>66</b>
<b>Chapter 6: Somatosensory function in speech perception in individuals with hearing impairment</b>	<b>70</b>
<b>6.1 Introduction</b>	<b>70</b>
<b>6.2 Methods</b>	<b>70</b>
6.2.1 Experimental procedure and data analysis	70
<b>6.3 Results</b>	<b>71</b>
<b>6.4 Discussion</b>	<b>74</b>
<b>Chapter 7: Speech sensorimotor adaptation in individuals with hearing impairment</b>	<b>78</b>
<b>7.1 Introduction</b>	<b>78</b>
<b>7.2 Methods</b>	<b>79</b>
7.2.1 Experimental procedure	79
7.2.2 Statistical Analysis	80
<b>7.3 Results</b>	<b>80</b>
<b>7.4 Discussion</b>	<b>82</b>
<b>Chapter 8: General Discussion</b>	<b>84</b>
<b>8.1 The somatosensory function in the perception-production link in NH and HI persons, in a developmental perspective</b>	<b>84</b>
<b>8.2 Future Works</b>	<b>86</b>
<b>8.3 Conclusions</b>	<b>88</b>
<b>Appendix 1</b>	<b>89</b>
<b>References</b>	<b>91</b>

# List of Figures

<b>Figure 1:</b> Auditory pathway. (Image from Graham et al., 2023)	18
<b>Figure 2:</b> Parts of a hearing aid and its types. (Image from Gorman, 2018)	23
<b>Figure 3:</b> Components of cochlear implant (Image from Gupta et al., 2012)	24
<b>Figure 4:</b> Perioral Musculature (Image from Howard, 2019)	38
<b>Figure 5:</b> Orofacial nerves (Image from ‘Trigeminal Nerve’, 2024)	39
<b>Figure 6:</b> Vowel space showing /e/-/ø/ contrast	44
<b>Figure 7:</b> Picture corresponding to “Dé” (right panel) and “Deux” (left panel) used in the picture naming task	46
<b>Figure 8:</b> (a) Experimental setup for vowel identification test with somatosensory stimulation. Reproduced with permission from (Ito & Ostry, 2010). (b) Temporal relationship between the somatosensory stimulation (onset, dotted line) and the auditory stimulation (onset, plain line).	47
<b>Figure 9:</b> Representative example of participant’s responses in the speech-in-noise test. Left panel represents digit-singlet phase (4 dB step change), and right panel represents digit-triplet phase (1 dB step change). Circles represent correct responses and crosses represent incorrect responses.	49
<b>Figure 10:</b> Experimental setup for the vowel identification test with somatosensory stimulation. This setup was also used in the speech-in-noise test by removing the tabs from the facial skin (no physical connection between the robotic devices and participants).	49
<b>Figure 11:</b> (a) Judgment probabilities and estimated psychometric function for a representative participant. The horizontal dotted line shows the 50 <sup>th</sup> percentile of the judgment probability. The point where it crosses the estimated psychometric function is the PSE which represents the category boundary between the two vowel targets. (b) Mean PSE values in control and skin-stretch conditions. Error bars show the standard error.	52
<b>Figure 12:</b> (a) Temporal pattern of target words ‘dé’ and ‘deux’ (top panel) and corresponding spectrograms (bottom panel). The trajectories of the three formants (F1, F2 and F3) extracted using PRAAT are displayed by thick solid lines. The 40ms time window used for formant measurement is temporally set at the middle of the vowel and marked by the two vertical thin solid lines. (b) Averaged formant values across participants. Error bars show standard error.	53
<b>Figure 13:</b> Scatterplots of perception index (PSE differences) against production index (differences between /e/ and /ø/ in F1, F2 and F3). Solid lines represent the regression line and the shaded area represents 95% confidence intervals.	54
<b>Figure 14:</b> Experimental procedure associating a perceptual training phase with a pre- and post-test for speech production recording.	58

**Figure 15:** Representative speech signals of the test words “dé” and “deux” recorded in the production sessions (pre- and post-test). Top panel represents acoustic signals and bottom panel represents corresponding spectrograms with detected first, second and third formants (red dotted lines) and RMS power (blue solid line). Two black dashed lines show the vowel region detected based on RMS power. Two vertical yellow solid lines delimit the temporal window used for formant detection in the following analyses. \_\_\_\_\_ 59

**Figure 16:** Pre-post change in F1, F2 and F3 in each group (SOMA1, CTRL and SOMA2) with (a) the values averaged across speakers and across vowels /e/ and /ø/ and (b) the values averaged across speakers in each vowel /e/ and /ø/, separately. Error bars represent the standard error across participants. The mark \* correspond to  $p < 0.05$  by multiple comparisons in LME analysis. \_\_\_\_\_ 61

**Figure 17:** (a) Scatterplots of the per-speaker average pre-post change in formant related to the categorical boundary difference between the first and second halves of perceptual training, for F2 in the SOMA1 group and F2 and F3 in the SOMA2 group. The solid line represents the regression line and the shaded area represents 95% confidence intervals. (b) Psychometric functions during perceptual training averaged across participants, for the SOMA1 and for the SOMA2 group. Red dashed lines represent the one in the first half of the training and black solid lines represent the one in the last half of the training. Error bars represent the standard error across participants. \_\_\_\_\_ 63

**Figure 18:** Direction of formant change in F1-F2 (top panels) and F2-F3 (bottom panels) spaces for SOMA1 (left panels), CTRL (middle panels) and SOMA2 (right panels) groups. The blue circles on the unit circle represents the angles for each participant. Red solid lines indicate the mean vector representing the formant change. When the length of this vector is 1, all individual points are superposed and correspond to the same change angle. When the length of this vector is 0, all individual points are distributed uniformly in all directions. \_\_\_\_\_ 64

**Figure 19:** F2-F3 values of the vowels in the production session before perceptual training (blue squares for male speakers, red circles for female speakers) and of the edge sounds in the /e/-/ø/ continuum used for perceptual training (black triangles). Filled dots for /e/ sounds, empty dots for /ø/ sounds. Ellipses represent 95% confidence intervals separately for male and female productions and for each of the two vowels. \_\_\_\_\_ 65

**Figure 20:** Results from the perception tests and hearing profile in each hearing impaired (HI) participants (A) Hearing thresholds obtained in the speech-in-noise test. Vertical line labeled NH represents the averaged threshold from normal-hearing individuals ( $-8.5 \pm 0.06$  dB SNR). (B) Hearing profile across lifetime obtained by questionnaire. The entire length corresponds to their age. Filled area represents the duration after the detection of hearing-impairment. In this filled area, black part represents the duration of not using any hearing device, light grey part represents the duration of using a hearing-aid, and striped gray part represents the duration of using cochlear implants in either one or both ears. (C) Judgement probabilities in the vowel identification test. Red represents the condition with somatosensory perturbation (SOMA). Black represents the auditory-only condition (CTRL) \_\_\_\_\_ 72

**Figure 21:** Averaged response of judgement probabilities in the vowel identification test. NH corresponds to normal hearing group. HI-1, HI-2a and HI-2b corresponds to the hearing impairment groups represented in Figure 20. In each panel, left part represents judgement probabilities across each auditory stimulus and right part represents the judgement probabilities averaged across stimulus. Error bars represent standard errors across

participants. \_\_\_\_\_ 74

**Figure 22:** F2 Formant shift in the adaptation task \_\_\_\_\_ 79

**Figure 23:** (A) Hearing thresholds obtained in the speech-in-noise test. Vertical line labeled NH represents the averaged threshold from normal-hearing individuals. (B) Hearing profile across lifetime obtained by questionnaire. The entire length corresponds to their age. Filled area represents the duration after the detection of hearing-impairment. In this filled area, black part represents the duration of not using any hearing device, light grey part represents the duration of using a hearing-aid, and blue part represents the duration of using cochlear implants in either one or both ears. (C) Judgement probabilities in the vowel identification test. \_\_\_\_\_ 80

**Figure 24:** (A) Changes in F2 across trials (B) Averaged formants in baseline, hold and aftereffect phase. The left panel corresponds to HI-CI group, right is NH group. The error bars represent standard errors across the participants. \*:  $p < 0.05$ , \*\*:  $p < 0.01$ , and \*\*\*:  $p < 0.001$ . \_\_\_\_\_ 82

# List of Tables

<b>Table 1:</b> <i>Classification of severity of hearing impairment from the range of PTA value in the better ear (from World Report on Hearing, 2021).</i>	20
<b>Table 2:</b> <i>Formant values of synthesized stimuli used in the vowel identification task (Adapted from Trudeau-Fisette et al., 2019)</i>	45
<b>Table 3:</b> <i>Summary of statistical results based on linear mixed-effect model analysis and corresponding multiple comparisons.</i>	62

# List of Abbreviations

**CI** – Cochlear implants

**F1** – First formant

**F2** – Second formant

**F3** – Third formant

**HA** – Hearing aids

**HI** – Hearing-impaired

**NH** – Normal-hearing

**PSE** – Point of Subjective Equality

**PTA** – Pure Tone Average

**SNR** – Signal-to-noise ratio

**TMS** – Transcranial magnetic stimulation

**ULB** – Université Libre de Bruxelles

**WHO** – World Health Organisation

# Introduction

Individuals with hearing impairment can perceive speech with help of hearing devices such as hearing aids (HA) and/or cochlear implants (CI). Although these devices allow hearing-impaired (HI) individuals to master speech abilities concerning perception and production at levels possibly similar to normal-hearing (NH) individuals, their speech abilities can be dependent upon various factors such as age at implantation, type of the device used and duration of device usage.

The main focus of this thesis is to investigate the auditory-somatosensory interactions in speech perception and production in NH and HI individuals. Since speech-related somatosensory information associated with articulatory movement can modify the outcome of a speech perception task (Ito et al., 2009; Trudeau-Fisette et al., 2019), orofacial somatosensory inputs can also play a key role in the interaction between speech perception and production. Considering that the somatosensory role in speech perception could be acquired during the course of speech production development, the auditory-somatosensory interactions could vary in HI individuals depending on the various factors mentioned above. The current thesis tried to address these issues in four behavioral studies. The main experimental paradigm consists of using a robotic device to produce somatosensory stimulations on the facial skin in a precisely controlled way both in space and time.

This manuscript is split into eight chapters. Chapter 1 introduces the background about hearing impairment and its impact on speech perception and production. Chapter 2 describes the background about the functional link between speech perception and production mechanisms and introduces a possible hypothesis on how this interaction could form the basis for the contribution of orofacial somatosensory input in speech processing from the viewpoint of multisensory integration. Chapter 3 describes the methodology concerning the main tasks and stimuli used in the current work. Chapter 4 presents the first study that investigated the relationship between the effect of somatosensory inputs in speech perception and production performance in NH individuals. Chapter 5 presents the second study that investigated whether perceptual auditory training paired with somatosensory inputs changes speech production in NH individuals. Chapter 6 describes the third study which investigates the role of orofacial somatosensory inputs in speech perception in HI individuals and how this is related to different factors such as age of implantation, type of device and duration of device. Chapter 7 describes the fourth study which examines whether HI individuals show sensorimotor adaptation and how this could be related to factors affecting their hearing ability. Finally, chapter 8 proposes a general discussion about the four studies, future perspectives and conclusion.



# Hearing impairment and its impact on speech perception and production

---

## 1.1 Overview

Hearing plays a vital role in day-to-day functioning as it enables us to engage with the environment around us and to communicate with others. Indeed, its importance can be seen in the way a decline in hearing impacts the lives of HI individuals. It is often associated with social isolation (Mick et al., 2014), cognitive decline (Bisogno et al., 2021), depression (Li et al., 2014; Nordvik et al., 2018), decreased quality of life and communication difficulties (Dalton et al., 2003). The latest update of the Global Burden of Diseases study set up by the WHO (World Health Organization) estimated the prevalence of hearing impairment from 1990 to 2019 based on the severity of the impairment (Haile et al., 2021). They found that around 1.5 billion people experience a decline in hearing at some point during their course of life, with 1.57 billion people with a hearing impairment in 2019. In all, around 430 million people had a moderate to complete hearing impairment requiring some form of rehabilitation, a number which increased from the previous 225 million in 1990 and which is projected to double over by the year 2050 (around 700 million). Geographically, the largest number of people with moderate to complete hearing impairment resided in the Western-Pacific regions. This distribution varies across different regions due to different factors such as occupational noise exposure, preventable infections and health care access. Such variations across regions can lead to unaddressed hearing impairment which is detrimental as it causes considerable social and economic burden. It can occur across all age groups with a clear association between ageing and hearing impairment. Indeed, hearing impairment prevalence largely increased in adults aged 50 years and older (Haile et al., 2021). This can further contribute to dementia and cognitive decline in older individuals (Bowl & Dawson, 2019; Cardin, 2016). While hearing impairment prevalence is small in children less than five years old, hearing impairment can actually be severe in these young children – as it is also common in the elderly (Haile et al., 2021). Since hearing ability and auditory feedback provide crucial sensory information for speech communication, hearing impairment has severe detrimental consequences for the acquisition of speech perception and production. In particular, congenital HI individuals have serious difficulties to acquire

language, they display poor literacy and suffer reduced employment opportunities (Garramiola-Bilbao & Rodríguez-Álvarez, 2016; Matkin & Wilcox, 1999).

The advent of hearing devices (HA and CI) has revolutionized the treatment of hearing impairment. Specifically, early implantation of these devices in congenital HI individuals improves their ability of speech perception and production at an almost normal level. Individuals who acquire hearing impairment in later stages of life (acquired hearing impairment) can also benefit from hearing devices to regain their hearing and prevent deterioration of their speech ability. However, there is considerable variability in speech outcomes of HI individuals using hearing devices. These outcomes depend on different factors that can affect their speech acquisition such as the implantation age, type and usage duration of hearing device (HA or CI).

This thesis focuses on speech production and perception abilities of HI individuals using hearing devices (HA and/or CI). This chapter introduces a general background of hearing impairment and how speech perception and production of HI individuals are dependent on above-mentioned factors. Section 1.2 describes the basics of hearing mechanism and the types of hearing impairment. Section 1.3 describes the methods for diagnosis of hearing impairment. Section 1.4 describes classification of hearing impairment based on how and when it occurs. Section 1.5 describes treatment by hearing technologies with emphasis on HA and CI. Finally, section 1.6 describes speech perception and production abilities of HI individuals by focusing on factors such as implantation age, and type and usage duration of hearing device.

## **1.2 The auditory system and hearing impairment**

The mechanism of hearing consists of a series of steps that convert the acoustic signals from the surrounding environment to nerve impulses which are then transmitted to the brain to be decoded. The auditory pathway is classically separated into the “peripheral” and the “central” auditory system as shown in Figure 1. Briefly, the sound waves are picked up by the pinna in the outer ear and travel through the ear canal which leads to the eardrum. The incoming acoustic signal vibrates the eardrum (tympanic membrane) and these vibrations set the ossicles of the middle ear (malleus, incus, stapes) in motion. The function of the ossicles is to adapt and transmit the vibrations from aerial in the outside to liquid inside the cochlea in the inner ear. Pressure waves in the endolymph, the liquid inside the cochlea, set a membrane inside the cochlea, the basilar membrane, to vibrate, and these mechanical vibrations are converted by the

sensory cells of the cochlea into nerve impulses in the auditory nerve. The peripheral auditory system comprises the outer, middle and inner ear and the auditory nerve. The neural information from the auditory nerve is then transmitted to the auditory cortex, via a set of processing centers which constitute the central auditory system. This includes the cochlear nuclei, superior olivary nuclei, lateral lemniscus, inferior colliculus, medial geniculate nuclei, before reaching the auditory cortex in the temporal lobe of the brain (Peterson et al., 2023).

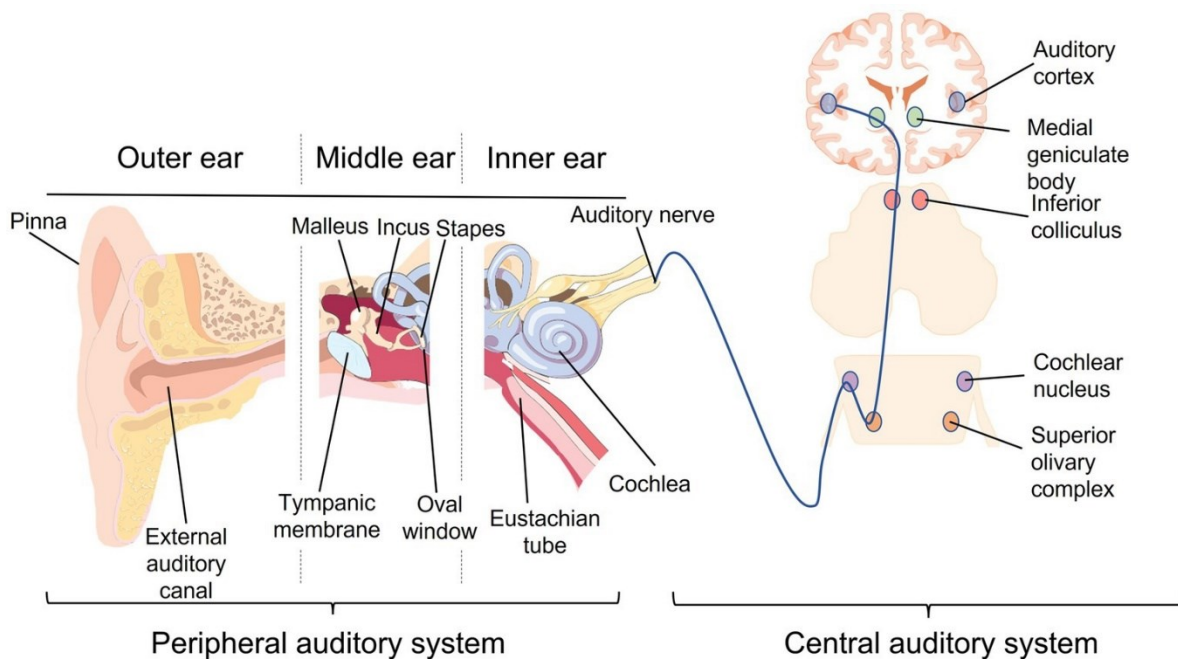


Figure 1: Auditory pathway. (Image from Graham et al., 2023)

Hearing impairment occurs when there is a problem in some part of this auditory pathway. A damage in the peripheral auditory system results in a difficulty to transmit acoustic signals to the central auditory system. This can be due to a problem in the outer or middle ear which results in a “conductive” hearing impairment where the sounds cannot pass to the inner ear. Common causes of conductive hearing impairment are otitis media, accumulation of earwax, presence of foreign objects and injury or defect in the outer ear, ear canal and ear drum. A problem in the inner ear and/or the auditory nerve results in a “sensorineural” hearing impairment. This can be caused due to aging, noise exposure, infections and diseases, head injuries and ototoxic medicines. A combination of conductive and sensorineural hearing impairment results in a “mixed” hearing impairment.

At a higher processing stage, a damage to the central auditory system results in a difficulty to analyse or interpret auditory information which is known as a “central auditory processing disorder”. Individuals with such an impairment typically have a functioning outer, middle and inner ear resulting in normal pure-tone audiometric thresholds, but they have problems with

speech perception in noise suggesting a disruption in the higher processing centres of the brain (Bellis & Bellis, 2015).

Depending upon their type, hearing impairment can sometimes be treated with medicines or surgeries. This is more common in the case of conductive hearing impairment, for which accumulation of ear wax can be removed by ear irrigation, dissolved by cerumenolytic agents or manually removed using curette or forceps (Schwartz et al., 2017). Ear infections can be treated by antibiotics, and abnormal ear growths can be removed by local surgery (Dwyer-Hemmings et al., 2019; Ramakrishnan et al., 2007). But rehabilitation with the help of hearing interventions is necessary when clinical treatment is not possible. This is typically the case for sensorineural or mixed hearing impairment and for central auditory processing disorders, where the usage of hearing devices is the most common solution. In the context of this thesis, the main focus is on individuals with sensorineural hearing impairment who use HA or CI.

### **1.3 Diagnosis of hearing impairment**

A first standardised method systematically used in diagnosis of hearing impairment is provided by pure-tone audiometry consisting in the evaluation of hearing threshold at various frequencies. In this method, a pure tone at a given frequency (e.g., 500, 1000, 2000 or 4000 Hz) is first presented at a sound level well above the estimated threshold and then decreased in 10-dB steps until it is no longer audible. The sound level is then increased in 5-dB steps or decreased in 10-dB steps depending on the participant's response about its audibility, until reaching the lowest sound level at which the participant responds at least 50% of the time that the sound is audible. This sound level is referred to as the hearing threshold for the corresponding frequency, and the "Pure Tone Average" (PTA) corresponds to the average of hearing thresholds for the set of tested pure tone frequencies. The severity of hearing impairment is classified based on this audiometric measurement. Table 1 shows the different grades of hearing impairment classified by the WHO and the relative functional experience in different hearing environments. A 'disabling' hearing loss is characterised when the hearing thresholds are greater than 35 dB in the better ear (*World Report on Hearing, 2021*).

*Table 1: Classification of severity of hearing impairment from the range of PTA value in the better ear (from World Report on Hearing, 2021).*

Severity	Thresholds in better ear (in dB)	Hearing experience in quiet environment	Hearing experience in noisy environment
Normal	<20	No difficulty hearing sounds	No/minimal difficulty hearing sounds
Mild	20 - 35	No difficulty with speech	May have difficulty with speech
Moderate	35 - 50	May have difficulty with speech	Difficulty with speech
Moderately-severe	50 - 65	Difficulty with speech, can hear raised voices	Difficulty with speech
Severe	65 - 80	Cannot hear speech, may have difficulty with raised voices	Extreme difficulty with speech
Profound	80 - 95	Extreme difficulty with raised voices	Cannot hear speech
Complete or total hearing loss/deafness	>95	Cannot hear speech and most environmental sounds	Cannot hear speech and most environmental sounds
Unilateral	<20 in better ear or >35 in worse ear	No difficulty unless sound is near the poorer ear. Difficulty with sound localisation	May have difficulty with speech and in locating sounds

Still, apart from pure-tone audiometry, rehabilitation is often dependent on measures that better reflect the sounds of the ‘real world’ and particularly speech measures (Hermann et al., 2019; van der Straaten et al., 2021). Speech evaluation of the hearing ability is related to speech perception and production performance in HI individuals. A first tool, which will be used further in this thesis, is the Digit-triplet test introduced by Smits et al. (2004). This test was initially developed in Dutch as an automated version of hearing screening via telephone and has been extended into several languages such as French, English, Polish and German (Hall, 2006; Jansen et al., 2010; Ozimek et al., 2009; Zokoll et al., 2012). In the test, three digits are presented inside background noise and the participants are asked to identify the digits. A different set of digits is presented at each epoch of the procedure. An adaptive procedure is used, where the signal-to-noise ratio (SNR) is changed based on the participants’ response in different steps to calculate the resulting “speech reception threshold”, that is the minimum SNR value ensuring audibility of the presented speech material. There are several advantages of

using digits as the target stimulus sound. Firstly, digits are frequently used and hence familiar words. Secondly, the test avoids higher-level cognitive aspects, not purely auditory, such as working memory and linguistic skills, when compared to the detection of words or sentences. Therefore, this method is a reliable functional measure to assess hearing thresholds, with reproducible scores from one evaluation to the next in a given participant. Interestingly, its outcomes strongly correlate with pure-tone audiometric thresholds (Potgieter et al., 2018) and other speech-in-noise measures (Smits et al., 2004). Hence, instead of using pure-tone audiometry, an adapted French version of the Digit-triplet test was used as a main tool for the measurement of hearing ability in the studies reported in this thesis.

## **1.4 The sources of hearing impairment**

Hearing impairment can occur at different stages of life during which individuals are exposed to multiple risk factors associated to, e.g., genetic, biological, behavioural or environmental causes. Given the heterogeneity of the causes, hearing impairment can be classified in different ways. Classification could be based on the degree or severity of the impairment, whether it affects one or both ears, if it affects both ears similarly or not, whether it happens gradually or suddenly, at which stage of life it occurs and how. This thesis mainly focuses on differences associated to the impairment occurrence. When hearing impairment is present at birth or soon after, it is called congenital hearing impairment. When it appears later, it is referred to as acquired hearing impairment.

Congenital hearing impairment could be due to genetic mutations involving connexin encoding gene (Estivill et al., 1998), intrauterine viral infections such as Cytomegalovirus and Rubella infections (Cohen et al., 2014; Grosse et al., 2008), low birth weight (Cristobal & Oghalai, 2008), lack of adequate oxygenation during birth (Leite et al., 2016) and the use of ototoxic medicines (Foch et al., 2018). The most common cause of congenital hearing impairment in newborns is ear infections (sometimes called as otitis media), often due to accumulation of fluid in the middle ear, ruptured ear drum and ear inflammation, and which results in a lost or degraded sound transmission through the middle ear (Aarhus et al., 2015; Monasta et al., 2012). Other infections are meningitis, mumps and measles (Cohen et al., 2014; Zeeshan et al., 2018). Acquired hearing impairment can occur at any point during the course of life and could be a result of disease, ageing, injury or lifestyle choices. Ageing, the most common cause of acquired hearing impairment, may produce a progressive decline of hearing sensitivity as the ear structures degenerate (Gordon-Salant, 2005; Wu et al., 2020). Over 65% of people aged over 60 experience some degree of hearing impairment (Haile et al., 2021). It is a common sensory

deficit in older adults associated with depression, cognitive decline, dementia resulting in vast health, societal and economic costs (Bowl & Dawson, 2019; Cardin, 2016; Stucky et al., 2010). Chronic pathologies such as diabetes and hypertension are also associated with decline in hearing as they cause sensorineural and microvascular damages disrupting the neurotransmission of auditory information (Cho et al., 2016; Duck et al., 1997). Additionally, prolonged exposure to excessive noise due to occupational activities or recreational activities such as regular usage of portable audio devices with a high volume may cause mechanical, metabolic and nerve damages to the ears, leading to a noise-induced hearing loss (Choi & Choi, 2015; Daniel, 2007). Head and ear trauma (Chen et al., 2018; Maillot et al., 2016), accumulation of earwax (Michaudet & Malaty, 2018), nutritional deficiencies (Emmett & Keith P West, 2015), smoking (Lin et al., 2020) and malformations or abnormal growths such as otosclerosis, acoustic neuroma, cholesteatomas (Weber & Klein, 1999) can also lead to acquired hearing impairment.

In the speech research framework, there is also a classification of deafness depending on the age in relation with the acquisition of speech and language. When deafness occurs before the acquisition of speech and language, it is called prelingual deafness. When deafness occurrence is posterior to acquisition, it is referred to as postlingual deafness. Since speech and language can develop from a very early age, typically one (Kuhl, 2004), the congenital vs. acquired and pre-lingually vs. post-lingually classifications mostly overlap. Since the thesis focuses on speech perception and production in HI individuals, we consider in the current work the classification between prelingual and postlingual deafness/impairment to describe the participants involved in the various studies.

## **1.5 Hearing interventions by hearing technologies**

HI individuals can benefit from a variety of hearing interventions such as using hearing devices (HA or CI) and from other communication strategies including sign language, lip-reading and cued speech. Communication requirements and individual preferences are considered to select a suitable hearing intervention. For example, sign language, which is a visuo-spatial mode of communication with its own grammatical and linguistic structure, is an efficient communication system in prelingually deaf individuals with insufficient access to sounds (Murray et al., 2019). Both lip-reading and cued speech, which is a visual mode of communication involving manual gestures superimposed on speech orofacial gestures, facilitate speech perception as well (Bergeson et al., 2005; Leybaert & LaSasso, 2010). While these interventions are effective, the

main focus in the present work is on participants who benefited from hearing intervention provided by hearing technologies, specifically HA and CI. These are commonly used devices that provide increased access to auditory information which is otherwise impeded due to decline in hearing. In particular, early detection of hearing impairment and timely use of such hearing devices benefit better speech perception, language development and overall quality of life. This section introduces how HA and CI compensate loss of hearing ability.

### 1.5.1 Hearing aids

HAs are small electronic devices that improve hearing by amplifying the sounds. The basic components of HA are a microphone, an amplifier and a speaker (see Figure 2A). The sound is received through the microphone which is then converted to electrical signals and sent to the amplifier. The power of the signal is increased in the amplifier – possibly differently along frequencies – and then it is sent to the ear through the speaker.

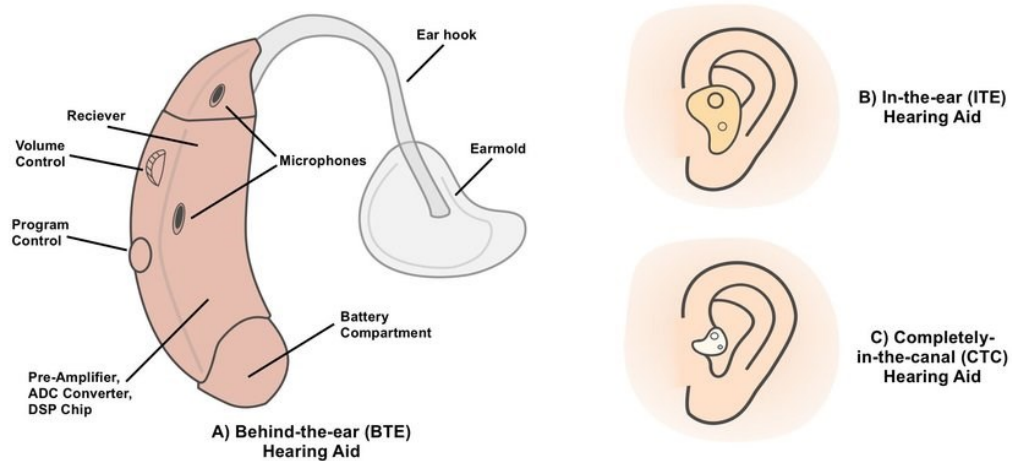


Figure 2: Parts of a hearing aid and its types. (Image from Gorman, 2018)

Figure 2 shows the different types of HAs based on the size and placement in the ear (Weber & Klein, 1999). Behind-the-ear (BTE) HA are worn behind the ear. In-the-ear (ITE) HA fit inside the outer ear and are relatively smaller in size. The other type is canal aids such as in-the-canal (ITC) or completely-in-canal (CIC) HA which are fit within the ear canal. Based on the way they work, HAs can be categorised into analog and digital (Kim & Barrs, 2006), depending on the format of the converted sound signals. Digital HAs enable complex signal processing, resulting in flexibility to customise the HA based on the user and specific listening environments. This includes additional features that help in fine-tuning the delivered sounds such as suppression of auditory feedback, noise reduction, amplifying softer sounds more than louder sounds (nonlinear compression) and usage of directional microphones to aid spatial



hearing (Kim & Barrs, 2006; Sprinzl & Riechelmann, 2010). Still, HAs provide limited benefit for speech perception in noisy environments as they amplify both the speech and noise signals (Duquesnoy & Plomp, 1983; Plomp, 1978).

### 1.5.2 Cochlear implants

Cochlear implantation is the commonly used intervention in case of a sensorineural hearing impairment when traditional amplification by HA provides insufficient benefits. It bypasses the damaged regions of the ear and directly stimulates the auditory nerve with electrical signals. This device mainly consists of an external portion placed behind the ear and a surgically implanted portion under the skin behind the ear (see Figure 3). The external portion has a microphone to detect sound, a speech processor, a transmitter to transmit the sounds to the internal portion and a battery to power the system. The internal portion consists of a receiver to receive signals from the transmitter and an electrode array that sends electrical signals to different regions of the auditory nerve. A surgical procedure is required to get a CI. After the implantation, it takes time and therapy to habituate to listening sounds with the CI.

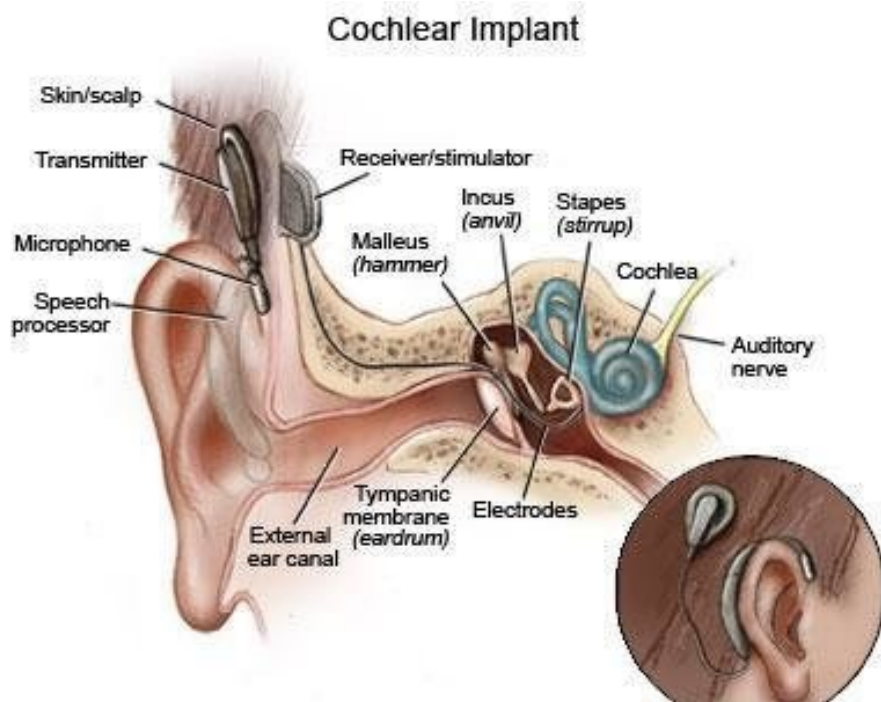


Figure 3: Components of cochlear implant (Image from Gupta et al., 2012)

CIs are useful and possibly required when there is small or no residual hearing ability left so that sounds cannot be amplified acoustically. Importantly, CIs are more beneficial for profoundly deaf individuals when compared to HAs (Mildner et al., 2006).

### **1.5.3 Other devices**

Although HAs and CIs are the most commonly used devices for people with hearing impairment, there are various other types of devices that can also help these people. When somebody with hearing impairment may not benefit from a HA and is not a candidate for a CI (in case of mixed or conductive hearing impairment), bone conduction and middle ear implants are an option. Bone conduction implants bypass the damaged middle ear and transmit sounds to the inner ear (Håkansson et al., 2019). Similarly, active middle ear implants mechanically stimulate the inner ear by vibrating the middle ear structures (Lassaletta et al., 2019). Apart from these, there are assistive listening devices which support HI individuals in loud or busy environments. This includes FM (frequency modulation) systems where the speaker speaks in a microphone connected to a receiver worn by the listener (Sprinzl & Riechelmann, 2010).

## **1.6 Impact of hearing impairment on speech perception and production**

As mentioned in the above section, hearing technologies (HA and CI) can help in restoring the hearing ability to acquire speech perception and production to a sufficient extent for daily life. However, HI individuals show varied outcomes in their speech abilities when compared with NH individuals (Bouchard et al., 2009; Bouton et al., 2012; Geers et al., 2003). Such variability in speech abilities can be due to an interplay of various factors such as duration of deafness (Bernhard et al., 2021), age at implantation (Geers et al., 2003), type of device used (Mildner et al., 2006), usage duration of the device (Blamey et al., 2013), cognitive factors (Pisoni, 2000), brain plasticity and reorganisations (Rouger et al., 2007; Strelnikov et al., 2010) and surgical approach (Finley et al., 2008; Lazard et al., 2012). This thesis mainly focuses on the impact of implantation age, type of hearing device and the usage duration of hearing devices. These factors are discussed in the present section.

### **1.6.1 Implantation age**

Age at implantation is an important factor affecting speech perception (Geers et al., 2003; Kileny et al., 2001; Nikolopoulos et al., 1999; Waltzman & Cohen, 1998). In case of congenital hearing impairment, receiving correct speech sounds within a critical period of 3.5-4 years (at best by second year) is crucial to ensure complete speech and language development (Kral & Eggermont, 2007; Kral & Sharma, 2012). If the hearing devices are implanted within such periods or before, then it increases the possibility to restore hearing abilities similar to those of typical hearing development. Geers et al. (2003) showed that children who received a CI below

the age of five could achieve rather good speech perception performances with an average level of 50% success in a large test battery including measures of phoneme, word and sentence recognition, but variations in performance could be quite large with a distribution of scores covering the whole range between 0 and 100%.

For speech production and language acquisition also, early access to speech sounds plays an important role (Connor et al., 2006; Karltorp et al., 2020). Connor et al. (2006) investigated whether receiving a CI in early years of life provides an added benefit in improving speech production and vocabulary outcomes by testing 100 children who had received CI between 1 and 10 years of age. They found that children who were implanted before the age of 2.5 years had significantly stronger outcomes than children who had received implants at a later age. Early implantation resulted in faster rates of speech production and vocabulary growth immediately after implantation. This burst of speech and vocabulary growth was reduced systematically with increasing age of implantation and was not observed in children older than seven years at implantation. This suggests that the impact of age of implantation in speech production outcomes is related to the developmental processes that occur at specific time periods.

All these findings converge to show that early implantation of hearing device is a critical factor to develop speech abilities, and hence that age of implantation is one of the important factors to study the development of speech perception and production.

### **1.6.2 Type of hearing devices**

Speech perception may differ based on the type of hearing device used. Mildner et al. (2006) carried out minimal-pair discrimination tasks in the perception of vowels and consonants in children equipped with CI vs. HA (in the following, CI vs. HA children). They found that both groups of children could discriminate vowels more correctly than the other sounds with relatively less confusions in distant vowels (example: /i/-/u/) compared to closer ones (example: /a/-/e/). However, the CI children performed relatively better in vowel discrimination compared to the other group, and they showed a consistent relationship between vowel distances and discrimination accuracy in contrast to the HA children. Globally, this suggests that the perceptual vowel space in the former group is closer to the target space for NH children. CI children also showed better performance in distinguishing consonants that differed in the manner, voicing and place of articulation. In contrast, HA children performed below chance level at these consonant discriminations. In all, the CI children performed significantly better

when compared with HA children in all tested sounds.

For speech production, CI children showed a more accurate consonant production than HA children in terms of consonant error patterns with less omission and substitution errors in produced consonants (Baudonck et al., 2010). They also have a relatively natural voice quality and overall higher speech intelligibility as judged by non-professionals (university students) and professionals (speech therapists) (Mildner et al., 2003). Similarly, Horga & Liker (2006) found that the CI children performed better than the HA children in terms of vowel differentiation, vowel intelligibility (except /a/), word accent production, voice and pronunciation quality, whereas the voice onset time and closure duration were not differentiated between CI and HA children.

Altogether, this shows the impact of type of device on speech abilities with CI users showing clearly better outcomes.

### **1.6.3 Usage duration of device**

The last important factor that we want to consider here is the duration of hearing device usage. Concerning speech perception, Blamey et al. (2013) carried out a retrospective multicentre study where they examined speech perception scores from 2251 post lingually deaf CI users. The evaluation method differed among centers (e.g., phoneme vs. word perception) but each CI user in each center provided comparative pre- vs. post-operative scores. The authors assessed what factors affected the speech perception performance of the CI users in an attempt to update and compare with their previous study (Blamey et al., 1996). In their previous study, the most important factor was the duration of deafness among other factors such as age at implantation, age at onset deafness, duration of implant usage and aetiology of deafness. Although those five main factors still remained in this study, they found that the duration of implant usage was the most significant factor with a relative reduction in the effects of duration of deafness and age. They attributed this shift to the evolving strategies behind care pathways provided for CI users such as better clinical management, less strict patient selection criteria, changes in surgical practises and improved devices.

For speech production abilities, Langereis et al. (1997) investigated the effect of CI on vowel production in three distinct time periods – pre-implantation, three months post-implantation and twelve months post-implantation in 20 post-lingually deaf adults. They found that after a year of usage of the CI there was a significant expansion of the vowel space reaching the normative range. They also showed less variability in productions and increase in the ability to make

phonological contrasts. When they compared the three-month post-implantation and twelve-month post-implantation performance, they found that some individuals performed poorly after three months of device use but they had improved after twelve months of use. This initial decrease in quality in vowel production could be due to the sudden interference of the auditory feedback in the first three months, but it may still improve with longer duration of usage as shown in the individual performance after twelve months of device use.

Grandon & Vilain (2020) investigated the development of fricative production in French-speaking CI children and compared it to NH children. They carried out a word repetition task and a picture-naming task with words consisting of fricatives (/f/, /s/, /ʃ/) as the target consonant, followed by either /i/ or /u/. They used spectral centre of gravity measures related to the acoustical focus of the fricative, together with standard deviations related to the range of frication noise, as acoustic correlates of the fricative productions. Overall, they found a fricative development in CI children, but the phonological representations were dependent upon the hearing ability of these children. Indeed, the results showed a lower centre of gravity with /s/ and smaller standard deviations with /f/ in comparison to NH children. Such differences in fricative production could be due to the fact that the perception of high-frequency sounds is challenging with CI. They also found that the duration of usage of CI resulted in continued development of fricative production suggesting that with adequate auditory experience with CI, these children can catch up to the performance of the NH children.

Taken together, there is an impact of usage duration of hearing devices on both speech perception and production in HI individuals.

# The role of orofacial inputs in speech perception and production

---

## 2.1 Overview

There is a growing interest in the somatosensory function in speech processing. The somatosensory function contributes not only in speech production, but also in speech perception, and we posit that this function could be central in the relationship between the production and perception systems. The current thesis investigates the role of the somatosensory function in speech production and perception in HI individuals in relation with its role in NH participants. This chapter presents the background about the orofacial somatosensory function in speech perception and production, and introduces our general hypothesis about the role of the somatosensory function in setting a relationship between speech production and perception, together with the related questions about the consequences for NH and HI individuals. Section 2.2 discusses the potential link between speech perception and production in terms of how articulatory movements can be involved in perceiving speech. Section 2.3 describes this perception-production link from the viewpoint of multisensory integration. It discusses how different sensory systems such as vision and somatosensation contribute to speech perception. Section 2.4 introduces a keypoint point of the current thesis, that is the involvement of the orofacial somatosensory system in speech from an anatomical and functional point of view. Finally, section 2.5 summarises the main topics of the thesis and specifies the general hypothesis and the underlying experimental questions.

## 2.2 Links between speech perception and production

### 2.2.1 The involvement of the motor system in speech perception

The link between speech perception and speech production has been investigated for several decades by considering the role of the speech motor system in speech perception in the framework of the so-called “Motor Theory of Speech Perception” (Liberman et al., 1967). This theory originated in the 1950s with the observation that there was a lack of one-to-one correspondence between the acoustic signal and the corresponding speech sound that was

perceived (Liberman et al., 1952). This study used the sound spectrograph with an instrument called as ‘pattern playback’ to make synthetic speech sounds, an instrument which was originally developed as a reading machine for the blind. The authors presented the synthesised unvoiced stops (/p/, /t/, /k/) followed by several vowels and asked the participants to identify the initial stop consonant as /p/, /t/, or /k/ regardless of the vowel they heard with it. They found that the same stop burst centred around 1440 Hz was identified as /p/ (produced with a labial constriction of the vocal tract) when it was followed by the vowels /i/ or /u/ and was identified as /k/ (produced with a velar constriction) when it was followed by the vowel /a/. This indicates that the acoustic cues were not invariant in nature, but were context driven because of coarticulation, that is the fact that the speaker, during the realisation of the consonant, anticipates the following vowel. This was further investigated in the perception of /d/ in /di/ and /du/. While the temporal pattern of the second formant in /d/ was different depending on the following vowel (rising pattern in /di/ and falling pattern in /du/), both sounds are perceived as the same consonant /d/ with an alveolar constriction in the vocal tract (Liberman et al., 1954). These findings were interpreted by the authors from the articulatory viewpoint, stating that the perception of these syllables was related to the articulatory movements rather than to the acoustic signal: the articulatory movement for the production of /d/ is consistent in /di/ and /du/, the sound differs, but the /d/ percept is the same, as is the articulation. This interpretation illuminates the central idea that the speech motor system is not only vital for articulating speech gestures but also for perceiving speech sounds based on the listener’s knowledge of the speech production mechanisms.

In the framework of the Motor Theory of Speech Perception, the involvement of the motor system in speech perception was hypothesized as the call to cognitive/mental “simulation” processes. Differently Fowler (1986) proposed that the actual articulatory gestures structuring the acoustic signal are “directly” recovered by the listener (the Direct Realist theory). In this theory, the listener does not perceive some information mediated by cognitive processes but rather the exact physical features which are the cause of the stimulation, i.e., for speech, the articulatory gestures that shape the phonetic segment. More recently, Schwartz et al. (2012) proposed the Perception-for-Action-Control Theory (PACT) according to which a speech unit is neither purely auditory nor motor but rather perceptuo-motor in nature. Thus, procedural knowledge from speech production together with sensory information from the auditory, visual and somatosensory systems which shape the articulatory gestures, contribute to the makeup of phonetic representations and, ultimately, of the speech perception process.

In parallel to these theoretical developments, the hypothesis of the motor system involvement

in speech perception has been strengthened by a series of neurocognitive findings concerning so-called “mirror neurons” in non-human primates (Rizzolatti et al., 1988), and, as a direct follow-up, a “mirror system” in humans (di Pellegrino et al., 1992). To cite one of the first and influential findings, the human premotor cortex which has long been known to be activated by hand grasping movements, has also been found to be activated when the same action was observed (di Pellegrino et al., 1992) – hence the term “mirror”, a single structure being involved in both producing an action and perceiving the same action made by somebody else. The existence of mirror neurons and mirror systems strongly support claims about the motor system involvement in speech perception. This has further been investigated in the speech domain. Thus, Watkins et al. (2003) showed that, when transcranial magnetic stimulation (TMS) was applied on the left motor cortex in the region associated to lip muscles, the amplitude of motor-evoked potentials was enhanced with the presentation of speech stimuli presented either auditorily or visually, but not with the presentation of non-speech stimuli. Fadiga et al. (2002) showed that, when TMS was applied on the left motor cortex in the region associated to anterior tongue muscles, motor-evoked potentials increased in a condition where the participant was presented with speech sounds that involved tongue movements (e.g. /r/), but not with speech sounds that did not involve tongue movements (e.g. /f/). Later on, fMRI studies showed that motor areas involved in speech production were activated during passive listening to specific speech sounds (Pulvermüller et al., 2006; Skipper et al., 2005; Wilson et al., 2004). Taken together, these findings clearly show the involvement of the motor system in speech perception and point towards a link between speech perception and production.

### **2.2.2 Experimental evidence in favor of a functional link between speech perception and production**

The functional nature of the perception-production link is supported by behavioural studies showing how speech perception and production abilities are related in many ways. Better auditory acuity in a speech discrimination task is associated in a given individual with larger articulatory and acoustic distances between speech categories and hence more precise targets in speech production (Franken et al., 2017; Perkell et al., 2004). Chao et al. (2019) related the perceptual boundary between two neighbour front vowels / $\epsilon$ / and / $\text{æ}$ / in American English with the production performance of these vowels. They found that the perceptual boundary in a given individual was closer to the vowel with the smaller variability in production. They also found a strong correlation among individuals between their perceptual boundary and their boundary in production. They concluded that speech production and speech perception systems are related



by a bidirectional link.

Studies involving the speech sensorimotor adaptation paradigm also support the hypothesis of a functional relationship between speech perception and production. In this paradigm, when a given sound produced by a speaker is played back to the speaker with altered acoustic parameters during a repetition task, it has been shown that the speakers adapt their production in the direction opposite to the imposed acoustic change (Houde & Jordan, 1998; Villacorta et al., 2007; see a review in Caudrelier & Rochet-Capellan, 2019). The interpretation is that the speaker is able to modify online the motor control process to partially recover from the perceptual adaptation and produce a modified target with acoustic properties eventually compatible with the perceptual template. When the acoustic perturbation is removed, the production stays modified for a while, displaying an “aftereffect”, and then gradually recovers from adaptation. Not all speakers display auditory-motor adaptation. We will exploit this paradigm on HI participants in one of the studies of this thesis (chapter 7).

Importantly, it has been shown that the perceptual representations of speech sounds were modified following speech motor adaptation tasks (Lametti et al., 2012; Nasir & Ostry, 2009; Shiller et al., 2009). Indeed, Shiller et al. (2009) showed that motor training on the production of fricative sounds with altered auditory feedback later modified the perceptual categorical boundary between the fricatives involved in production. Lametti et al. (2012) showed a similar perceptual modulation after sensorimotor adaptation in production for vowels. This was computationally accounted for by Patri et al. (2018) using a Bayesian joint model of speech production and speech perception, assuming a bidirectional link between production and perception.

The other way round, while speech motor adaptation appears to result in changes in speech perception, perceptual training with speech sounds also affects speech production. Cooper & Lauritsen (1974) showed that repeatedly listening to voiceless stop consonants (which are characterized by a large voice onset time, that is a large lag between plosive release and voicing onset), modified the further production of the corresponding consonants, inducing a decrease in produced voice onset time. Lametti et al. (2014) demonstrated that perceptual training modifying a vowel categorical boundary produced changes in the amount of speech motor adaptation in response to formant alterations in a sensorimotor adaptation paradigm. Altogether, these evidences confirm that the speech perception and production systems are closely linked.

### **2.2.3 Development of the speech perception-production link**

The perception-production link appears to be settled and/or shaped since early in development. Indeed, several studies provide evidence of sensorimotor influences in the processing of speech perception in infants in their first year of age. Thus, Bruderer et al. (2015) showed that 6-month-old infants displayed decreased performance in the discrimination of a non-native pair of speech sounds when the articulation relevant for the task was impeded by an appropriate pacifier; and DePaolis et al. (2011) showed that 9-month-old infants displayed differences in their auditory preference for sounds depending on their own production trends. Vilain et al. (2019) showed that the emergence of babbling in infants between 6 and 9 months could play a role in the extraction of vowel-independent representations of plosive place of articulation. Kuhl et al. (2014) showed that cortical maturation of the perceptuo-motor link from 7 to 11-12 months of age is related to speech production development, suggesting that the perception-production link capitalizes on motor experience for its development.

The development of the perception-production link has also been extensively studied in HI individuals (Blamey et al., 2001; O'Donoghue et al., 1999; Tye-Murray et al., 1995). The development of speech processing appears to depend on auditory experience during the periods of high neuroplasticity in the brain, that is, at least before 3.5-4 years of age (Kral & Sharma, 2012). If exposure to speech is missed during this period, then HI individuals can lose the possibility to acquire a typical level of hearing and fluent speaking (Cardon et al., 2012). As a result, the acquisition of perception-production link could also be impaired. Tye-Murray et al. (1995) showed that CI children with better perception of target speech features (voicing, nasality, duration, frication, and place of articulation) also displayed better intelligibility in production, and the production performance was related to the duration of CI experience. O'Donoghue et al. (1999) carried out a prospective study with profoundly deaf children at different stages of cochlear implantation such as pre-implantation, 2, 3, 4, and 5 years post-implantation. They showed that early speech perception abilities can affect later speech production skills, specifically speech perception abilities as early as two years post-implantation seemed to predict speech production in the three, four- or five-years post-implantation. All these studies converge to show that early sensorimotor experience is important in the development of speech communication including acquisition of the perception-production link.

## **2.3 Multisensory integration in speech**

### **2.3.1 Audio-visual integration and its development**

The involvement of the motor system in speech perception has also been discussed in the context of multisensory interactions involving different sensory systems such as auditory, visual and somatosensory. Audio-visual interaction is the most commonly investigated in this context, and a classic example is the McGurk effect (McGurk & Macdonald, 1976). In this effect, when a speech sound is paired with the visual component of another speech sound, it is typically perceived as a third different sound. For example, when an auditory /ba/ is presented with a visual /ga/, the participants mostly perceive it as /da/. This could be because the corresponding auditory stimulus for /ba/ has acoustic features rather close to those for /da/, whereas the visual stimulus for /ga/ has features similar to those for /da/. The /da/ response is hence more or less compatible with the discrepant input from the auditory and visual modalities, providing the resulting percept. This suggests that speech perception, considered as predominantly auditory, can be influenced by speech-related movement information from other sensory systems such as the visual system. Looking at the speaker's articulatory movements ("lipreading") also results in an increase of speech intelligibility in noise (Sumbly & Pollack, 1954). The information from visual cues enhances both detection and categorization of vowels and consonants in noisy conditions (Benoît et al., 1994; Erber, 1969; Grant & Seitz, 2000) and speeds up the phoneme detection process (Fort et al., 2010). The increase of speech intelligibility by visual information is actually prominent in HI individuals (Erber, 1971; Grant et al., 1998; Puschmann et al., 2019). Geers et al. (2003) showed that the global speech recognition performance in a battery of open-set perceptual tests reaches up to 80% in average when lipreading is added to listening in HI individuals – though with a high inter-individual variability. Altogether, this shows that visually acquiring the speaker's articulatory information affects speech perception, and that the effect is particularly strong for HI persons.

The development of audio-visual integration may take a long time depending on the maturation of individual sensory systems and also on the accumulation of sensory experience (Gori, 2015). McGurk & MacDonald (1976) already found age-related changes in visual integration in the processing of speech sounds in the McGurk effect. Indeed, adults showed a dominance of the visual modality while children showed a dominance of the auditory modality. Other studies also found that adults have an increased audio-visual gain compared to children (Lalonde & Holt,

2016; Ross et al., 2011). This is also evident in comparison between early and older children (5-8 y.o. vs above 10 y.o.), since early children did not show clear audio-visual integration, while older children showed integration similar to adults (Kaganovich & Schumaker, 2014; Sekiyama et al., 2014). These suggests the audio-visual integration in speech perception is likely acquired in a relatively late period (up to around 8-10 y.o.) compared with the onset of speech and language development (1 y.o.).

HI individuals depend on the critical period in which proper auditory inputs have to be received as described in section 1.6.1. The best timing is considered to be before 2 y.o. Early identification of hearing impairment and rapid interventions such as HA or CI help to develop perceptual and production performance and the resultant integration of different modalities (Kral & Sharma, 2012; Mayberry, 2010). When HI individuals missed to receive proper auditory inputs during this period, the development of audio-visual integration also suffers. This is observed in a study showing that CI children who were implanted after 2.5 years of age showed inconsistent or no McGurk effect (Schorr et al., 2005). Lachs et al. (2001) also showed that, in prelingually deaf CI children, children who acquired better listening-alone abilities also showed a better use of the complementary information provided by visual input. These observations suggest that the development of auditory-visual integration in HI individuals should also be affected by factors concerning implantation age and usage duration of hearing devices as described in section 1.6 for the development of auditory perception.

In summary, visually acquired articulatory movement information plays an important role in speech perception, and audio-visual integration in speech perception provides additional support in favor of the importance of the link between speech perception and production.

### **2.3.2 Audio-tactile integration**

In addition to audio-visual integration, the tactile sense also contributes to speech perception. Gick & Derrick (2009) showed that tactile stimuli using air puffs modified the perception of the voicing of plosive sounds. When air puffs were applied on the skin of either the hand or the neck in conjunction with the auditory perception of unaspirated voiced sounds (/ba/), the participants heard those sounds as the unvoiced aspirated /pa/ which is a sound produced with a burst of air 'aspiration'. In the case of HI individuals, tactile aids have been developed that benefit speech perception given sufficient training (Cowan et al., 1990; Weisenberger & Kozma-Spytek, 1991). In these devices, tactile cues can be presented in the form of

electrocutaneous stimulation (electrotactile devices) or vibrations (vibrotactile devices). These tactile cues consist of a signal input which is derived from particular speech features that may be important for understanding speech. Based on a similar concept, vibrotactile stimulation to the index finger or wrist also enhanced speech perception in noise in NH adults listening to CI-simulated auditory stimulation (Fletcher et al., 2018) and in CI participants (Fletcher et al., 2019; Schulte et al., 2023). These studies show that tactile information can also provide information associated with the presented speech sound and participate to speech perceptual processing.

The somatosensory system also provides articulatory-related information and this may also interact with the perception of speech sounds. In the Tadoma method, which is used by individuals with dual sensory loss such as deaf-blindness, produced speech sounds are perceived by assessing the articulatory movements, such as lip movement, jaw movement, oral and nasal airflow and laryngeal vibration, through the tactile sense of hands placed on the speakers' face (see review by Reed et al., 1984). Similarly to the McGurk effect, such tactile information appeared to alter speech perception when incongruent information from hands was paired with auditory information (Fowler & Dekle, 1991). Recent studies showed that orofacial somatosensory inputs associated with speech articulatory movements also alter the perception of speech sounds (Ito et al., 2009; Trudeau-Fisette et al., 2019). The interaction between auditory processing and orofacial somatosensory information associated with speech articulatory movements in speech perception is the main theme of the current thesis and is detailed in the following section.

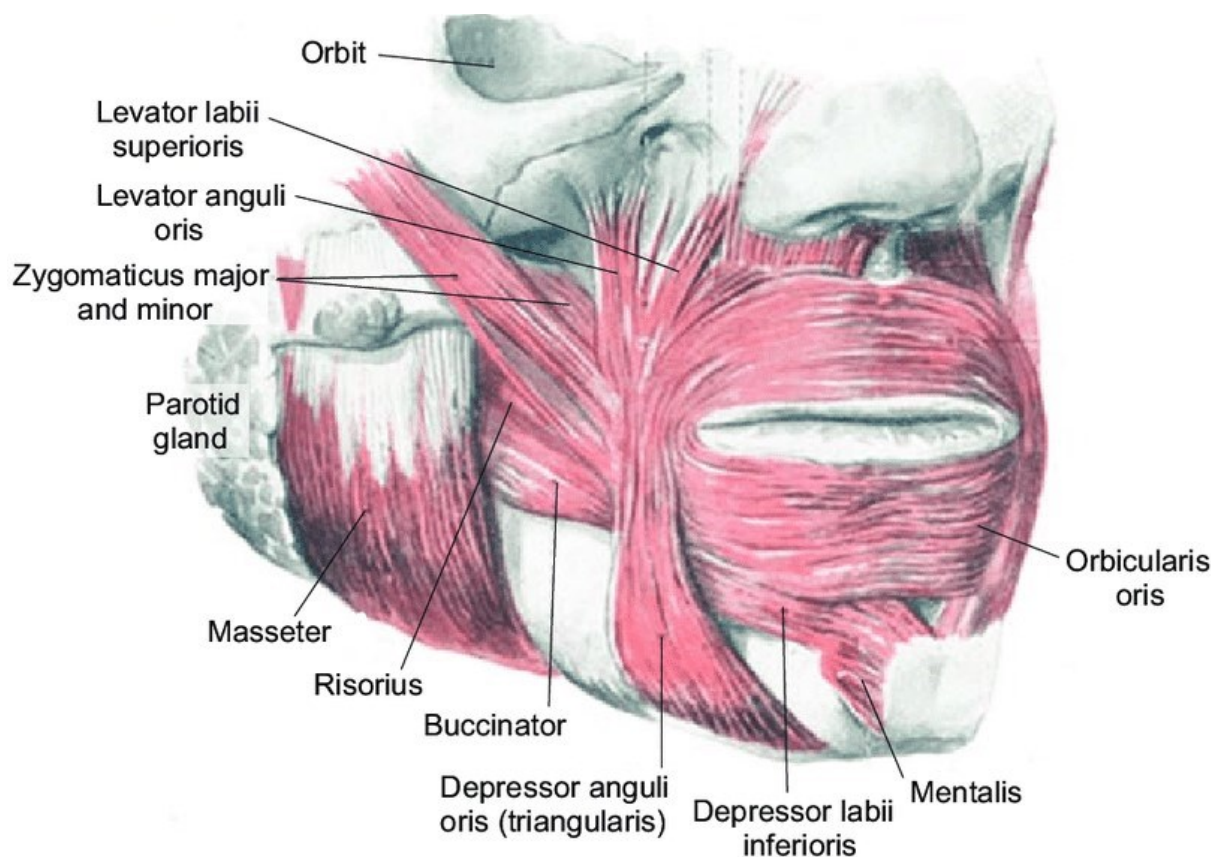
## **2.4 Orofacial somatosensory interaction in speech perception**

### **2.4.1 Source of kinaesthetic information from the orofacial system**

The term 'kinaesthesia' was coined by Bastian (1887) and it refers to the sense of movement. The sense of movement is widely investigated in the limb system with comparatively less attention given to the orofacial system. In limb motor control literature, it was initially thought that kinaesthetic information mainly arose from the joints as they are primarily related to limb movement. Boyd & Roberts (1953) showed the presence of kinaesthetic receptors in the knee joints of cats by recording afferent discharges from different positions and various movements. However, when a local anaesthetic was injected in the knee joint cavity, the movement ability remained unimpaired except for some perceived reduction of pressure sensations, and human

participants could still detect small changes in the angle of their knees (Clark et al., 1979). This suggests that the receptors in the knee joints may not be a main receptor for kinesthesia and hence other receptors such as muscle spindles and tendon organs may rather play a major role. A kinesthetic role of muscle spindles was assessed in a study using the tonic vibration reflex where an additional muscle contraction is induced due to a vibratory stimulation applied to the muscle. Goodwin et al. (1972) showed that the tonic vibration reflex induced by 100Hz vibrations applied to the biceps or triceps resulted in an illusion of movement in the elbows corresponding to the direction of the muscle stretch during the movement. In contrast, when the vibratory stimulation was applied on the elbow joints, no illusion was produced. As a result, muscle spindles and tendon organs are now considered to be a main source of kinesthetic information for limb motor control.

The possible sources of kinaesthetic information for speech motor control is still under debate. Figure 4 shows the orofacial muscles involved in speech production. Lips, jaw and tongue are involved in speech motor control. The jaw closing muscles are similar to the limb muscles with a large number of muscle spindles and the presence of a tonic vibration reflex (Hagbarth et al., 1976; Neilson et al., 1979). However, the situation for jaw opening muscles is more complex. While stretch reflex responses, which is another measure enabling to examine the role of muscle spindles, is induced for these muscles (Hellsing & Klineberg, 1982), the tonic vibration reflex is not (Hellsing, 1977). There is also a sparse density of jaw opening muscle spindles (Lennartsson, 1979). In lip muscles, although it has been considered for a long time that there were no muscle spindles (Stål et al., 1987, 1990), Omstead et al. (2023) recently showed the presence of muscle spindles in the orbicularis oris and zygomaticus muscles across a wide range of mammalian species, with a low density of muscle spindles in humans. However, several physiological studies still failed to show a typical function of lip muscle spindles. Neilson et al. (1979) did not find any evidence of stretch reflex in the lip muscles. Folkins & Larson (1978) showed an absence of additional muscle contraction when applying a tonic vibration stimulation in lip muscles (orbicularis oris and mentalis). In the tongue also, muscle spindles were found but they do not show a stretch reflex either (Neilson et al., 1979).



*Figure 4: Perioral Musculature (Image from Howard, 2019)*

This partly contradictory pattern in the physiology of orofacial muscles is also seen in nerve innervation. For sensory nerves, the orofacial system is innervated by the three branches of the trigeminal nerve – ophthalmic, maxillary and mandibular nerve in relation to the corresponding areas on the face (Figure 5). The maxillary and the mandibular nerve are the nerves for speech motor control. Although the trigeminal nerve is generally considered as a sensory nerve, one exception is the innervation of the jaw closing muscles, for which the mandibular nerve has both sensory and motor connections. This is similar to the nerve innervation for the limb muscles and consistent with the physiological finding that the jaw closing muscles show reflex responses similar to those of limb muscles as mentioned earlier. On the other hand, the lip muscles are innervated by two separate nerves respectively for motor and sensory function. While the trigeminal nerve serves for sensory function, the facial nerve serves for motor function. This is also similar in the tongue muscle, in which the hypoglossal nerve is for motor function and the lingual nerve is for sensory function. The innervation in the lips and tongue muscles is hence different from what happens for the jaw closing muscles and the limb muscles, which seems consistent with the lack of reflex related to muscle spindles for lips and tongue. This whole pattern suggests that the function of muscle spindles in the orofacial system, besides the jaw closing muscles, may not have a similar function as that of the limb system, and

hence that the kinaesthetic information in the orofacial system may come from other sensory structures.

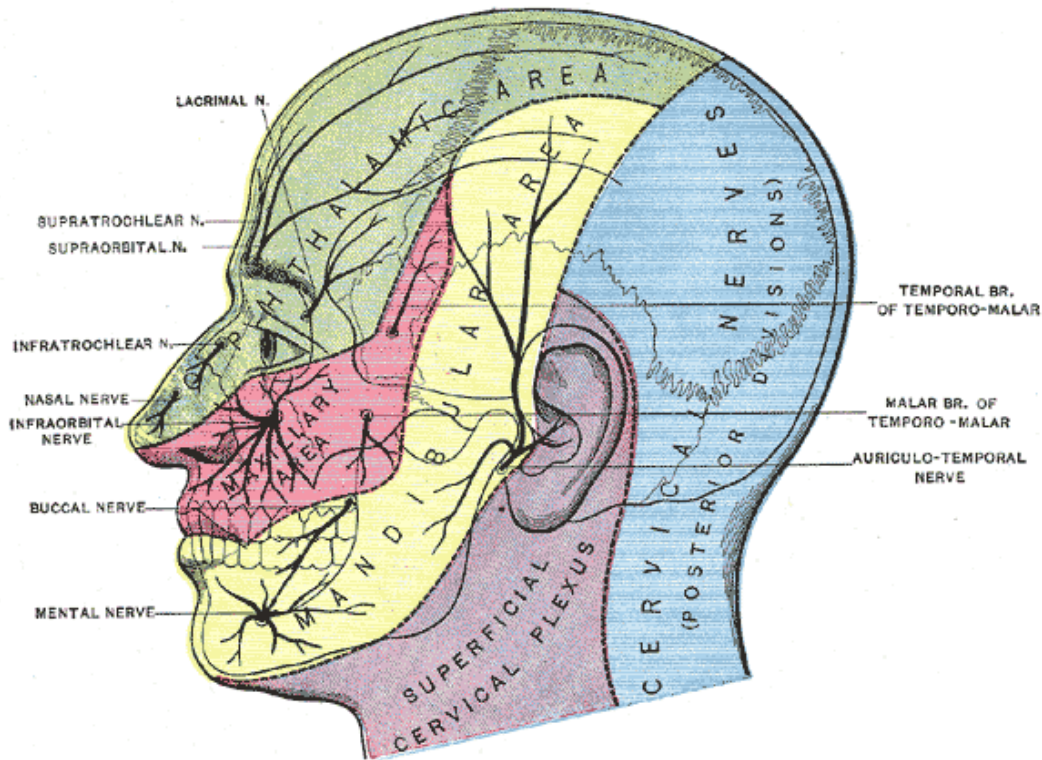


Figure 5: Orofacial nerves (Image from 'Trigeminal Nerve', 2024)

Cutaneous mechanoreceptors can be one possible source providing movement-related information. In the study of limb motor control, the cutaneous mechanoreceptors at the back of the hands respond to skin patterns associated with joint movements (Edin & Abbs, 1991). Externally applied skin-strain patterns in the finger produces movement illusions (Edin & Johansson, 1995). Cutaneous mechanoreceptors in the human thigh region also provide information related to knee-joint movements (Edin, 2001). The cutaneous mechanoreceptors in the orofacial region are similar to the ones found in hands (Trulsson & Johansson, 2002). Micro-electrode recordings made from the infraorbital nerve innervating the orofacial region showed activation when the facial skin was deformed due to speech movements, chewing and external application of facial skin-stretch (Johansson et al., 1988; Nordin & Thomander, 1989). This is also evidenced by observing the reflex (Ito & Gomi, 2007) and speech motor adaption (Ito & Ostry, 2010) when skin lateral to the lip is stretched. This suggests that deforming the orofacial skin provides movement-related information for speech sensorimotor processes.



## **2.4.2 Movement related somatosensory inputs intervene in speech perception**

Crucially, the articulatory information provided by somatosensory inputs associated with facial skin deformation, described in detail in the previous section, happen to modulate the perception of corresponding speech sounds (Ito et al., 2009). In a vowel identification task contrasting the sounds of ‘head’ and ‘had’, which differ in jaw opening-closing, when a facial skin deformation was externally applied in the upward direction during the presentation of the sounds, the participants perceived ‘head’ rather than ‘had’. This change was turned into the opposite when the facial skin deformation was applied in the downward direction and was not induced when the facial skin was deformed in a direction (horizontally backward) irrelevant to the studied auditory contrast. Similarly, Trudeau-Fisette et al. (2019) showed that, when Canadian French participants had to identify sounds in an acoustic continuum between the front mid-vowels /e/ (unrounded) and /ø/ (rounded), the participants perceived the sound more as /e/ when a skin stretch stimulation was applied in the backward direction, a stimulation likely to reproduce somatosensory inputs related to the backward articulatory movement for the production of /e/. This effect was consistently obtained for both adults and children. Ogane et al. (2020) also showed that orofacial somatosensory inputs can affect speech perception at a higher processing level, that is, word segmentation for lexical decision. They showed that an orofacial somatosensory stimulation applied on both sides of the participant’s mouth in the upward direction changed the place of word segmentation depending on the stimulation timing in relation to the tested acoustic utterance.

Given such a functional role of somatosensory inputs in the perception of speech sounds, this auditory-somatosensory link may be utilized in the learning processes of speech perceptual representations. Ohashi & Ito (2019) showed that somatosensory inputs applied during a speech motor training task can play a role in the perceptual change following the speech motor training phase. Exploiting the paradigm in which motor training with altered auditory feedback changes the perceptual representation of the corresponding speech sounds (Lametti et al., 2014; Shiller et al., 2009 see section 2.2.2), the study showed that an additional somatosensory stimulation applied during the altered auditory feedback motor training phase modified the change in speech perception depending on the direction of somatosensory stimulation associated with facial skin deformation. This suggests that receiving somatosensory inputs together with auditory inputs during speech motor learning can be important for the learning of speech perception processes. This possibility was supported in the study of visuo-motor control, since Wong et al. (2012) showed that the motor learning performance was improved following perceptual training with

somatosensory inputs obtained by passive limb movement paired with visual inputs. This idea has been further examined in the changes of speech perception following perceptual training with specific pairs of auditory-somatosensory inputs (Ito & Ogane, 2022). In this study, the participants first received a perceptual training phase in which they had to identify a vowel continuum between /ε/ and /a/. When the stimulus sound was accompanied with an upward somatosensory stimulation in the training phase, the perceptual boundary between /ε/ and /a/ was shifted toward /ε/ following training. Altogether, these data show the role of orofacial somatosensory inputs in the acquisition or recalibration of speech perceptual representations. As shown in the development of audio-visual integration, auditory-somatosensory integration could evolve along with the development of speech production and perception. Trudeau-Fisette et al. (2019) showed that the size of the somatosensory effect in vowel perception was larger in adults than in children. However, it is still unknown how this auditory-somatosensory integration mechanism is developed in speech disorders including HI individuals. It can be expected that HI individuals with hearing devices can also acquire this auditory-somatosensory integration mechanism, and that the amount of interaction could vary depending on their production and perception ability as it is the case in the development of audio-visual integration (see section 2.3.1).

## **2.5 Research questions and Hypothesis**

The main goal of this thesis is to further investigate the role of the speech-related somatosensory function in speech perception and production, with a specific focus on the way hearing impairment may modify this role. A general hypothesis is that the orofacial somatosensory system plays an important role in the link between speech production and speech perception. By focusing on HI individuals who use hearing devices that may be either HA or CI or both, we explore the possibility that the development of auditory-somatosensory integration may be based on sensorimotor experience related to speech abilities in HI individuals. As shown in chapter 1, their speech abilities concerning hearing and production can be dependent on different factors such as implantation age, type of hearing device used and the duration of usage of hearing device. In the present chapter, we have seen that these factors might also contribute to the development of the perception-production link and of audiovisual integration. Still, the role of somatosensory inputs in speech perception and the way auditory-somatosensory integration occurs in HI individuals has never been explored to our knowledge. Further investigation with NH individuals will help to establish better experimental procedure for further tests on HI participants. For this aim, this thesis is organized in four successive studies.

The first two studies aim at better characterizing the perception-production link in NH participants, addressing the following questions:

1. Is there a relationship between somatosensory effects in speech perception and production performance in NH individuals?
2. Can orofacial somatosensory inputs paired with auditory sounds during perceptual training change speech production performance in NH individuals?

Capitalizing on the experimental response to these two questions, which increase our understanding on the way the somatosensory system shapes the perception-production link when the auditory system functions well, the two last studies explore this link when it dysfunctions in HI participants, with a specific focus on the role of implantation age, type of used hearing device and duration of hearing device usage, addressing the following questions:

3. Do orofacial somatosensory inputs also play a role in speech perception in HI individuals and what factors related to their hearing ability affect auditory-somatosensory integration?
4. Do HI individuals show speech sensorimotor adaptation and what factors related to their hearing ability affect adaptation performance?

These questions are explored in four related studies described in chapter 4, chapter 5, chapter 6 and chapter 7 respectively.

# Methodology

---

## 3.1 Overview

A major component of the current work is the use of articulatory-related somatosensory stimulation from the orofacial system associated with facial skin deformation. Using this kind of stimulation allows to investigate somatosensory effects without engaging the participant in a motor action, which enables to dissociate somatosensory from motor effects in the study of perception and production tasks.

This chapter presents in detail the paradigm and tools associated to this somatosensory stimulation method, together with the main speech tasks and stimuli used in the present studies. Section 3.2 provides information about the ethics approval obtained for the current work. Section 3.3 describes the main speech tasks in this thesis. Section 3.4 describes the main somatosensory stimulation devices. Section 3.5 describes the modified Digit triplet test to evaluate the participants hearing ability using speech stimuli. Finally, section 3.6 describes the experimental setup for the studies.

## 3.2 Ethics

The ethics approval for the four studies in the thesis was obtained from the local ethical committee of the Université Grenoble Alpes [Comité d’Ethique pour la Recherche, Grenoble Alpes (CERGA-Avis-2021-8)]. All the participants signed the approved consent form before starting the experiments. The participants were all French speakers. The NH participants reported no neurological, hearing or speech deficits. The HI participants had bilateral hearing impairment and used HA and/or CI. All HI participants were able to speak correctly, as informally checked along conversations before the experimental procedure. They were also asked to fill a hearing assessment questionnaire (see Appendix 1) to obtain information regarding their hearing such as onset of deafness, their used hearing devices and duration of usage.

Additionally, the fourth study was carried out in collaboration with the Université Libre de Bruxelles (ULB) and with the Centre Comprendre et Parler in Belgium who helped with the

recruitment of HI participants. The participants who took part in the fourth study signed the consent form approved by the Erasme-ULB Hospital-Faculty Ethics Committee.

### 3.3 Speech Tasks

The somatosensory effect in speech perception has been examined with the English vowels / $\epsilon$ / and / $a$ / (Ito et al. 2009) and with the French vowels / $e$ / and / $\emptyset$ / (Trudeau-Fisette et al., 2019). Considering that the target participants for the current work were native French speakers, we adopted the vowels / $e$ / and / $\emptyset$ / for the main speech tasks of the thesis.

These vowels are front mid-high vowels that are acoustically and articulatory close (see Figure 6). Articulatorily, these two vowels are contrasted by horizontal articulatory movements, that is, spreading vs. rounding of the lips and forward vs. backward movement of the tongue (Lindblom & Sundberg, 1971; Schwartz et al., 1993). Acoustically these two vowels are contrasted in the second formant (F2) and to a lesser extent in the third formant (F3). F2 and F3 for / $e$ / are higher than for / $\emptyset$ / while the first formant (F1) is mostly similar for the two vowels (Figure 6). These target vowels were used for both perception and production tasks.

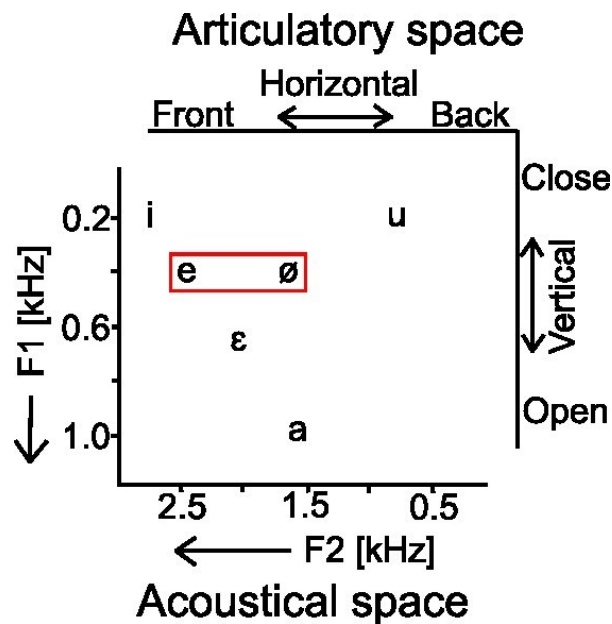


Figure 6: Vowel space illustrating the / $e$ /-/ $\emptyset$ / contrast

For the perception task, a vowel identification test was carried out using an eight-member acoustic continuum between / $e$ / and / $\emptyset$ /. This continuum, adapted from the study by Trudeau-Fisette et al. (2019), was synthesized using an articulatory synthesizer based on the Maeda's model (Maeda, 1990) as in Ménard & Boe (2004). In this model, a sagittal section of the vocal tract shape is defined by seven parameters – jaw opening, tongue dorsum, tongue body, tongue

apex position, larynx height, lip protrusion and lip height. Then, the formants can be calculated using this vocal tract shape model. The continuum was obtained by controlling the first five formants (F1, F2, F3, F4 and F5) as follows: the F1 and F5 were kept constant while F2, F3 and F4 were equally shifted upwards in steps of 30, 40 and 50 Hz respectively, from one edge (/ø/) to the other (/e/). The resultant stimuli were isolated vowels of 300 ms duration. Table 2 shows the formant values for each member of the /e/-/ø/ continuum.

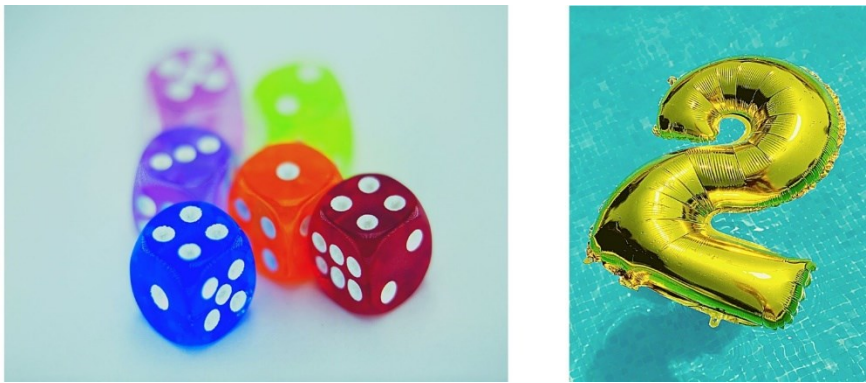
*Table 2: Formant values of synthesized stimuli used in the vowel identification task (Adapted from Trudeau-Fisette et al., 2019)*

Stimuli Number	F1	F2	F3	F4	F5
1 (/e/)	364	1922	2509	3550	4000
2	364	1892	2469	3500	4000
3	364	1862	2429	3450	4000
4	364	1832	2389	3400	4000
5	364	1802	2349	3350	4000
6	364	1772	2309	3300	4000
7	364	1742	2269	3250	4000
8 (/ø/)	364	1712	2229	3200	4000

In the test, one speech sound from this /e/-/ø/ continuum was presented to the participant. The participant was asked to identify if the heard speech sound was the vowel /e/ or /ø/ by pressing the corresponding key on a keyboard. For the analysis, the probability of /ø/ vs. /e/ responses was calculated for each of the 8 stimuli. Using these judgement probability values, individual psychometric functions were estimated to obtain the point of subjective equality (PSE) at the 50% crossover boundary value. This PSE value representing the category boundary between the two vowels was used to evaluate perceptual performance. This perception test was carried out both with and without somatosensory stimulation in order to measure the somatosensory effect in speech perception by comparing corresponding PSE values (see section 3.4).

For speech production, we recorded the target vowels /e/ and /ø/ produced by the participants. In this recording, the target vowels were embedded in the French words “dé” (English “dice,” /de/) and “deux” (English “two,” /dø/). We preferred to embed the vowels in a CV word rather than keeping them isolated to make the task more natural and familiar for the participants. The

number of repetitions was varied depending on the studies. For the purpose of possible future studies with children, the recording was carried out using a picture naming task where the corresponding pictures (Figure 7) were presented on a monitor. For the analysis, F1, F2 and F3 frequencies were extracted using PRAAT (Boersma & Weenink, 2001). The obtained formants were temporally averaged in a 40 ms time window. This time window was set at different temporal positions within the vowel utterance depending on the studies. In study 1, the time window was centered on the middle of the manually detected period of target vowel production (see Figure 12(a)). In study 2, since we observed possible effects at a later time during the vowel realization, we set the time window later in the vowel period (see chapter 5).



*Figure 7: Picture corresponding to “Dé” (left panel) and “Deux” (right panel) used in the picture naming task*

### **3.4 Somatosensory stimulation associated with facial skin deformation**

The current somatosensory stimulation method has been originally developed in Ito et al. (2009) and then used in several studies (e.g., Ogane et al., 2020; Trudeau-Fisette et al., 2019). This method is based on the assumption that orofacial skin deformation provides a kinesthetic information as introduced in section 2.4.1. The benefit of this method is to modulate somatosensory processing alone. This allows to investigate direct involvement of the somatosensory function dissociated from its role in the motor system. The schematic view of the experimental setup is represented in Figure 8(a). A small robotic device (Phantom 1.0, 3D Systems) is used to produce a force leading to stretch the facial skin. The robotic arm has a wire with plastic tabs (2 x 3 cm) at the other end. These plastic tabs are attached to the lateral side of the oral angle with double-sided tape.

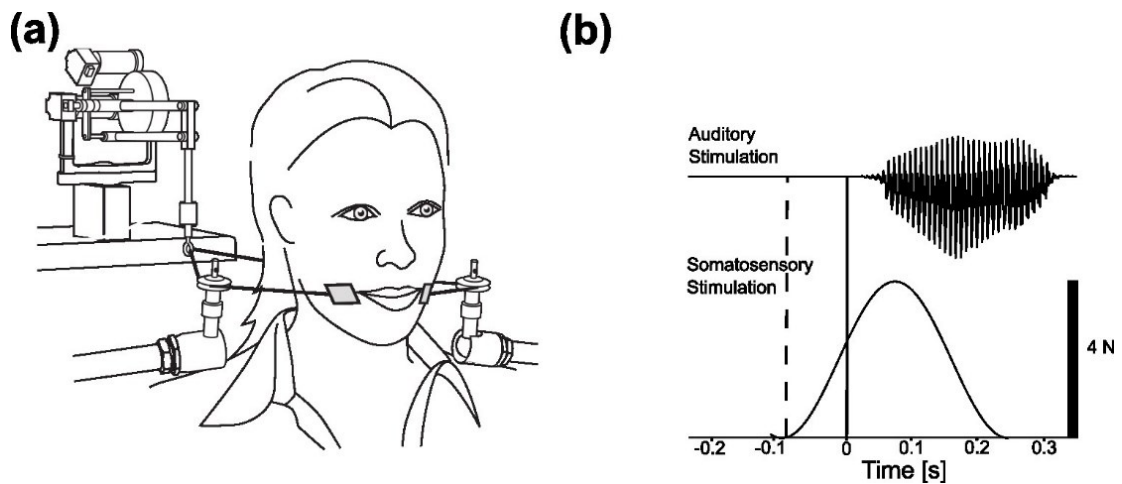


Figure 8: (a) Experimental setup for vowel identification tests with somatosensory stimulation. Reproduced with permission from Ito & Ostry (2010). (b) Temporal relationship between the somatosensory stimulation (onset, dotted line) and the auditory stimulation (onset, plain line).

By combining this somatosensory stimulation with the vowel identification task mentioned above, we investigated the somatosensory effect in speech perception. During the vowel identification task, the skin stretch was applied in the backward direction given the assumption that this direction of skin stretch can represent somatosensory inputs corresponding to the articulatory movement for the production of the lip-spread vowel /e/ as done in Trudeau-Fisette et al. (2019). To attempt to mimic as precisely as possible the somatosensory inputs for the production of /e/, the stimulation temporally followed a single cycle of a 3 Hz sinusoid, starting in amplitude with a force at 0N, a maximum at 4N and ending at 0N. This provides a duration of the somatosensory stimulation (333ms) which is comparable with the duration of the vowel stimuli (300ms). Considering that the articulatory movement precedes the produced sound, we set a 90-ms lead of the somatosensory stimulation relative to the onset of the auditory stimuli as done in Ito et al. (2009, 2014). Figure 8(b) represents the temporal pattern of this stimulation relative to the stimulus sound. We tested two conditions: auditory-alone condition (control) and auditory-somatosensory condition (skin-stretch). These two conditions were alternated every 8 trials, in which the 8 members of the synthesised /e/-/ø/ continuum were presented in random order. In total, 24 blocks of 8 trials were tested. The PSE values were calculated in each condition (control vs. somatosensory) by following the method described in section 3.3. The difference in PSE between the two conditions was used to assess the somatosensory effect in speech perception as done in the previous studies (Ito et al. 2009, Trudeau-Fisette et al. 2019).



### 3.5 Modified version of French Digit triplet test

Hearing ability was evaluated by measuring hearing thresholds in noise. We carried out a modified version of the French Digit Triplet Test. In the test, the participants were asked to identify what digit was presented by pressing a numeric keyboard. We applied monosyllabic French digits (1 – 9) following a prefixed carrier phrase “les numéros” (“The numbers” in English) using French Digit Triplet Test material in Jansen et al. (2013). The stimuli were recorded by a female French speaker. Each digit sound was embedded in a stationary speech-shaped noise. The background noise started 500ms before and stopped 500ms after the stimulus. In the test, the noise level was fixed at 70dB SPL and the signal-to-noise ratio (SNR) of the stimulus sound was controlled by the participant’s response using an adaptive procedure.

By expecting the large variability of hearing threshold in noise in the HI participants and in order to complete the test with a fixed number of trials, we modified the original test as follows. The test was divided in two parts named digit-singlet phase and digit-triplet phase as shown in Figure 9. The digit-singlet phase was used for rough estimation with large SNR (4 dB) steps. A digit-triplet phase was then used for detailed search with small SNR (1 dB) steps. In the digit-singlet phase, single French digits (1-9) were presented in ten trials with pseudo-random order. We started from 0 dB SNR and changed the SNR every trial based on a simple up-down procedure with 4-dB steps. Based on our empirical observation that the responses converged after four trials in most of the participants as shown in the example in Figure 9, we took an average of the SNRs in the last six trials for a first rough estimation of the hearing threshold. The obtained value was used as the initial value in the following digit-triplet phase. The procedure of this digit-triplet phase was similar to the one in Jansen et al. (2013). The SNR was decreased in 1-dB steps when all three digits were correctly identified and increased by 1 dB if there was at least one incorrect response. The final hearing threshold was defined as the average of the last seventeen SNR values.

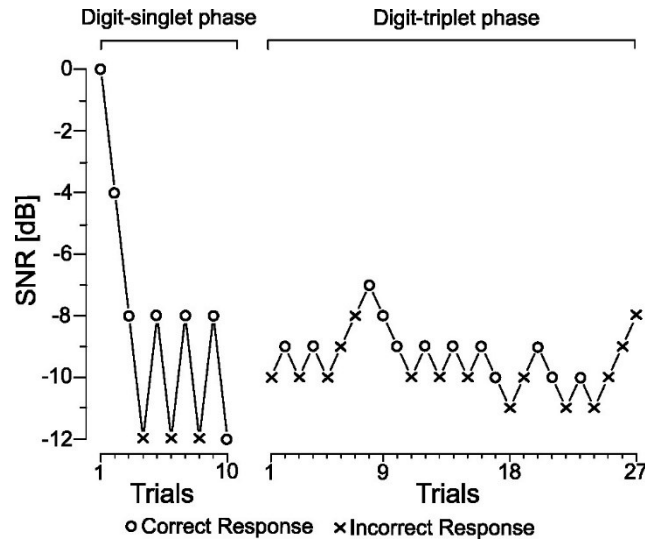


Figure 9: Representative example of a participant's responses in the speech-in-noise test. The left panel represents the digit-singlet phase (4 dB step change), and the right panel represents the digit-triplet phase (1 dB step change). Circles represent correct responses and crosses represent incorrect responses.

### 3.6 Instrumental configuration of the experimental setup

All experiments besides the test in ULB were carried out in a sound-proof room to reduce the effects of environmental noise. Figure 10 represents the equipment that were involved in studies 1, 2 and 3. The participants were seated in front of a computer monitor and a loudspeaker (Behritone C50A). To facilitate the experiment for the HI participants without having to adjust their hearing devices, we used a loudspeaker for presenting auditory stimuli, instead of headphones that were used in our previous studies (Ito et al. 2009, Trudeau-Fisette et al. 2019, Ogane et al. 2020). A robotic device (Phantom Premium 1.0, SensAble Technologies) was set behind the participants for somatosensory stimulation as described in section 3.4.

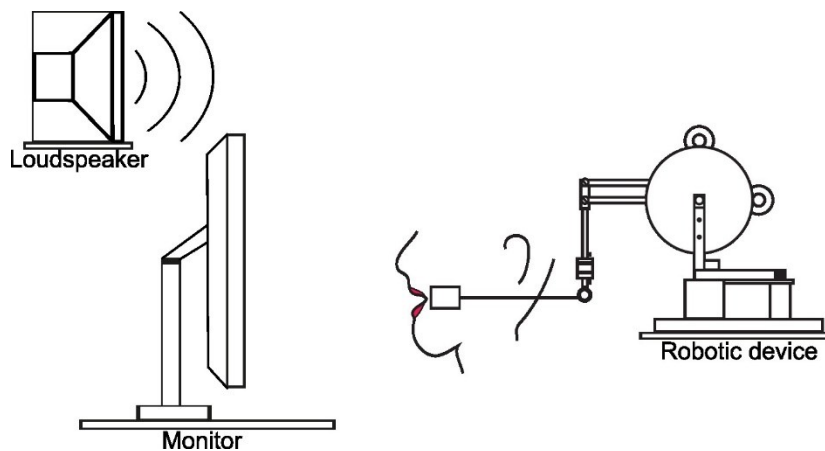


Figure 10: Experimental setup for the vowel identification test with somatosensory stimulation. This setup was also used in the speech-in-noise test by removing the tabs from the facial skin (no physical connection between the robotic device and the participant).

# Relationship between speech perception and production performance in individuals with normal-hearing

---

## 4.1 Introduction

Orofacial somatosensory inputs modify the perception of speech sounds (Ito et al., 2009). This has been consistently observed in both adults and children, where the size of the somatosensory effect was smaller in children compared to adults (Trudeau-Fisette et al., 2019). Such auditory-somatosensory integration likely develops alongside speech production acquisition. Considering that differences in developmental trajectories result in individual variability in speech production, our assumption is that the variability in speech production might also be related to the variability in the magnitude of the somatosensory effect in speech perception. The current study explores this assumption. We examined whether the somatosensory effect in speech perception varies depending on individual characteristics of speech production.

The experiment consists of a within-subjects procedure, in which participants are involved in both a speech perception task with somatosensory stimulation and a speech production task. The somatosensory effect in speech perception was quantified by the amount of category shift in a vowel identification test between the target vowels /e/ and /ø/, as performed in Trudeau-Fisette et al. (2019). For characterizing speech production performance, we recorded the corresponding vowels /e/ and /ø/ from the same participants, and evaluated the acoustic distance in F1, F2 and F3 between those vowels. Then, we carried out correlation analyses between acoustic distances in speech production and the amplitude of the somatosensory effect in speech perception. This analysis was applied in F1, F2 and F3 respectively. As known from previous studies, we expected that F2 would predominantly represent the acoustical contrast between /e/ and /ø/, while F1 should be stable from one category to the other (Ménard et al., 2008). F3 may also separately represent the articulatory lip rounding contrast for high front vowels (Lindblom & Sundberg, 1971; Schwartz et al., 1993). Therefore, we expected to observe a correlation with F2 and possibly F3 but not with F1.

This study is already published in the Journal – “Auditory, Perception and Cognition”.

Ashokumar, M., Guichet, C., Schwartz, J.-L., & Ito, T. (2023). Correlation between the Effect of Orofacial Somatosensory Inputs in Speech Perception and Speech Production Performance. *Auditory Perception & Cognition*, 6(1-2), 97–107.  
<https://doi.org/10.1080/25742442.2022.2134674>

## 4.2 Methods

### 4.2.1 Experimental Procedure

Nineteen native speakers of French participated in both a speech perception test with somatosensory stimulation and a speech production test. These two tests were carried out in a single session. The perception test with somatosensory stimulation was followed by the production test. The speech perception test with somatosensory stimulation was described in section 3.4 and the production test in section 3.3. In the production task, six repetitions were recorded for each task word. We also carried out a modified version of the French Digit Triplet test to measure the hearing threshold in noise (see section 3.5).

### 4.2.2 Statistical Analysis

As presented in section 3.3, PSE values were extracted using the estimated psychometric function and compared between the conditions (control and skin-stretch) using repeated-measures ANOVA. The amplitude of the somatosensory effect in speech perception was quantified using the difference in PSE between the skin-stretch and control conditions.

The speech production index was computed separately for F1, F2 and F3 as the acoustic distance in Hz between mean values of the six /e/ vs. /ø/ utterances.

Correlation analyses with Pearson correlation coefficients between perception and production performance among the 19 participants were then separately computed between the index of somatosensory effect in speech perception (PSE difference between conditions) and each of the three indices of speech production (average /e/ vs. /ø/ distance) as characterized respectively by F1, F2 and F3.

## 4.3 Results

The perceptual data successfully replicated the results obtained by Trudeau-Fisette et al. (2019), showing that perception of the /e/-/ø/ contrast was modified by the somatosensory perturbation associated with facial skin deformation. Figure 11(a) shows a representative example of the judgment probability for each of the 8 stimuli and the estimated psychometric function of a

representative participant. The estimated psychometric function was shifted towards the right (i.e. towards /ø/) in the skin-stretch condition, indicating that more stimuli were identified as the vowel /e/ in this condition. Figure 11(b) shows the mean PSE values in control and skin-stretch conditions. A one-way repeated-measures ANOVA showed a significant difference between control and skin-stretch conditions ( $F(1,18) = 7.42, p = 0.01$ ).

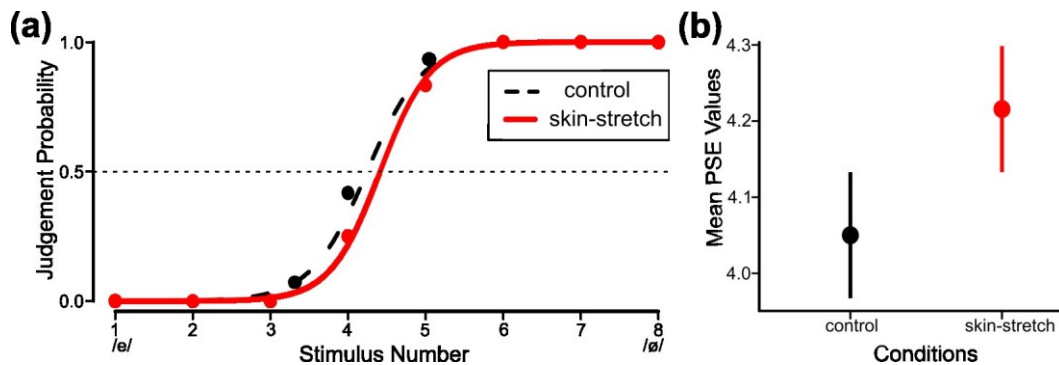


Figure 11: (a) Judgment probabilities and estimated psychometric function for a representative participant. The horizontal dotted line shows the 50<sup>th</sup> percentile of the judgment probability. The point where it crosses the estimated psychometric function is the PSE which represents the category boundary between the two vowel targets. (b) Mean PSE values in control and skin-stretch conditions. Error bars show the standard error.

Considering production data, Figure 12(a) represents the temporal pattern of the sound signal for one utterance of the word “dé” and “deux” (top panel) and the corresponding spectrogram (bottom panel). Horizontal solid lines represent the three formants (F1, F2 and F3). As shown in this representative example, F1 is similar, but F2 and F3 are different between the two vowels. Figure 12(b) represents the formant values averaged across participants, with the corresponding standard errors. Notice that the variability in formant values from one utterance to another of the same vowel is rather low for all participants (standard deviations of formant values for each vowel category and each participant and then averaged across participants are around 15 Hz for F1, 50 Hz for F2 and 60 Hz for F3 for both /e/ and /ø/). The acoustic distance between mean values for /e/ and /ø/ stays very small in F1, but large in F2 and F3, with higher F2 and F3 values for /e/ than for /ø/. As expected, the acoustical contrast between /e/ and /ø/ hence appears as well represented in F2 and in F3.

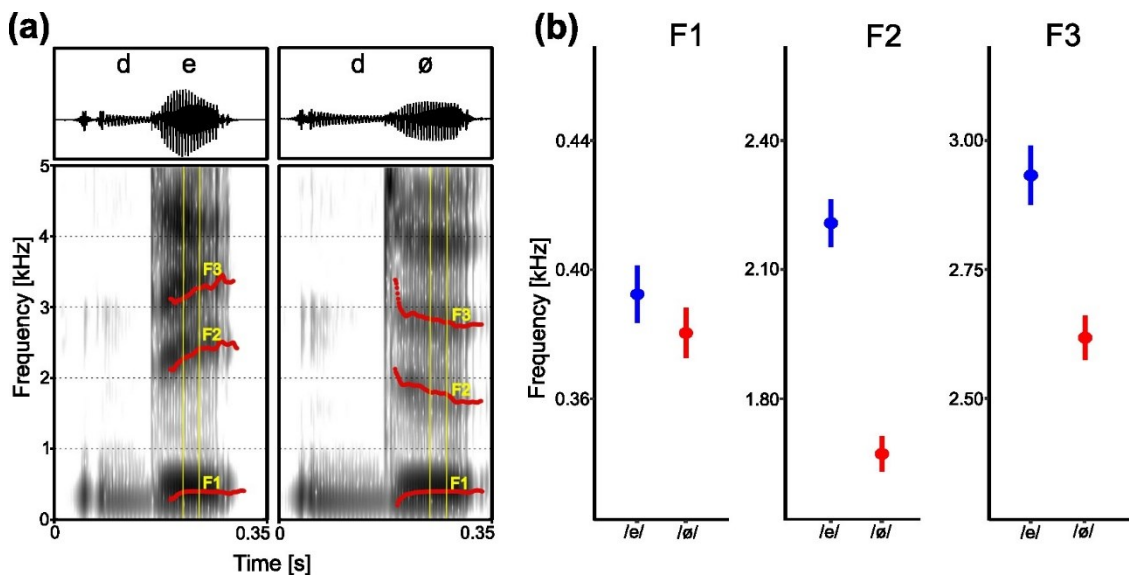


Figure 12: (a) Temporal pattern of target words 'dé' and 'deux' (top panel) and corresponding spectrograms (bottom panel). The trajectories of the three formants (F1, F2 and F3) extracted using PRAAT are displayed by thick solid lines. The 40ms time window used for formant measurement is temporally set at the middle of the vowel and marked by the two vertical thin solid lines. (b) Averaged formant values across participants. Error bars show standard error.

We then correlated the somatosensory effect in vowel identification (PSE difference) with speech production indices that are acoustic distances in F1, F2 or F3 between the target vowels /e/ and /ø/. Figure 13 represents scatter plots together with the regression line and 95 % confidence interval for each correlation display. While there is no significant correlation with F1 ( $r(17) = 0.17$ ,  $p = 0.49$ ), there is a significant correlation with F2 ( $r(17) = 0.48$ ,  $p = 0.03$ ) and correlation is marginally significant with F3 ( $r(17) = 0.41$ ,  $p = 0.08$ ). This indicates that larger somatosensory effects in speech perception are obtained for participants who have a larger difference between /e/ and /ø/ in F2 and/or F3 in speech production. Since the target vowels are contrasted in F2 and in F3 as shown in Figure 12, these correlation results suggest that the production of the contrast for the corresponding vowels can be related to the somatosensory-auditory interaction in speech perception.

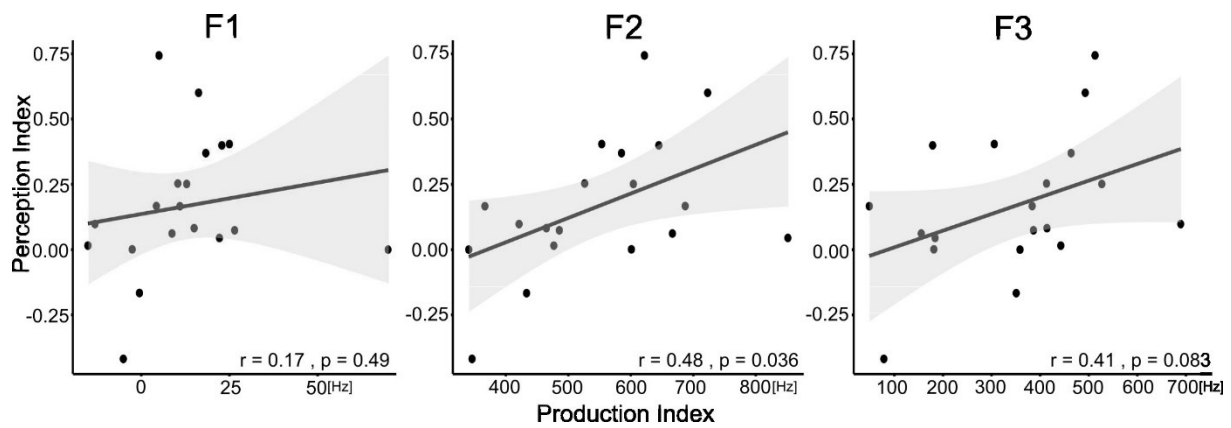


Figure 13: Scatterplots of perception index (PSE differences) against production index (differences between /e/ and /ø/ in F1, F2 and F3). Solid lines represent the regression line and the shaded area represents 95% confidence intervals.

## 4.4 Discussion

In the present study, we examined whether the somatosensory effect in speech perception correlates with individual variability of speech production. We successfully replicated the findings by Trudeau-Fisette et al. (2019) showing that the category boundary between /e/ and /ø/ was shifted towards /ø/ when a somatosensory stimulation associated with rearward skin stretch was applied. Importantly, we found a reliable correlation between the size of this somatosensory effect and the acoustic distance between /e/ and /ø/. This correlation was seen in F2 and marginally in F3, but not in F1.

The results from the current study confirm our expectation that the somatosensory effect and, accordingly, the auditory-somatosensory integration mechanism in speech perception might vary in relation with the participants' production performance. They shed light on the way this relation seems to operate, with larger acoustic distances between vowel targets leading to a larger somatosensory effect. A possible interpretation, in line with the previously cited studies showing a link between sensory and motor precision (Franken et al., 2017; Ghosh et al., 2010; Perkell et al., 2004), could be that if /e/ and /ø/ targets are far apart, this should correspond to a large articulatory gesture with a large movement of the face, associated with a large contrast in somatosensory configurations between the targets. Individuals with such a large acoustic distance would hence be more sensitive to the somatosensory contrast between the vowels, and in consequence, display a larger somatosensory effect in perception. The correlation between the somatosensory effect in perception and the acoustic distance between /e/ and /ø/ targets in production is significant for F2 but only marginally for F3, which could be related to the larger acoustic variation between the two vowels in F2 than in F3 (see Figure 12(b), showing that the

mean distance is 547 Hz in F2 vs. 345 Hz in F3, with also lower standard deviations in F2 than in F3). This is likely related to the fact that F3 is known to be a correlate of lip rounding for high (/i/ vs. /y/) rather than for mid-high (/e/ vs. /ø/) vowels (Lindblom & Sundberg, 1971).

The correlation between speech production performance and the magnitude of the somatosensory effect in speech perception fits well with the assumption that the development of the somatosensory function and of auditory-somatosensory integration in speech perception could be related to the acquisition of speech production. This possibility is also in line with the findings by Trudeau-Fisette et al. (2019) that the somatosensory effect in the perception of speech sounds was larger in adults than in children. Indeed, the not-yet-mature speech production in children would result in having an auditory-somatosensory integration still under development, and hence a smaller somatosensory effect in speech perception for children compared to adults. Interestingly, such developmental differences in auditory-somatosensory integration related to speech production abilities could have important consequences considering HI individuals. As mentioned in section 1.6, the presence of auditory experience during critical periods is important to acquire both hearing and speaking abilities and they are very dependent on age of implantation and usage period of hearing devices. We therefore anticipate that investigating the somatosensory effect in HI individuals equipped with hearing devices could provide further information on such underlying mechanisms.



# Speech perceptual training with somatosensory inputs changes speech production in individuals with normal hearing

---

## 5.1 Introduction

Speech production involves the integration of both sensory and motor processing. Previous studies have shown that participants adjust their motor behavior in response to an experimentally applied alteration of either or both auditory and somatosensory inputs (Houde & Jordan, 1998; Lametti et al., 2012; Tremblay et al., 2003). Interestingly, sensorimotor adaptation to speech motor training also induces changes in speech perception (Nasir & Ostry, 2009). Since speech perception and production are interactive in nature the reverse may also be true. That is, speech perceptual training may also affect speech production. Orofacial somatosensory inputs play an important role in speech motor control and speech learning. Since receiving specific auditory– somatosensory inputs during speech perceptual training alters speech perception, similar perceptual training could also alter speech production.

We explore this assumption that perceptual training using similar combination of somatosensory stimulation and vowel sounds could induce a change in the production of the corresponding vowels. To test this assumption, we carried out a vowel identification task on French participants using a perceptual training with the /e/-/ø/ continuum as done in Ito & Ogane (2022). A somatosensory stimulation produced by a backward facial skin deformation was applied in synchrony to the sounds during the training phase in one group of participants (somatosensory group) but not in the other group (control). Before and after the training phase, the participants had to utter the /e/ and /ø/ vowels. We recorded their productions and compared F1, F2 and F3 values of the uttered vowels before and after training. We expected that the combination of auditory training with these vowels and somatosensory stimulation associated with backward skin stretch could induce a possible modulation in speech production performance for the corresponding vowels. This would result in changes in production following the training phase mainly in the F2 – F3 space and mainly for the somatosensory

group compared with the control group who trained with auditory sounds alone. At this stage, we had no specific assumption about the possibility that the learning effect would be similar for both vowels in the training phase, or more focused on either of them. Hence, we systematically assessed possible differences in production for each vowel following the training phase.

Moreover, to better assess the nature of the coordination between sounds and somatosensory stimulation required for producing an effect of perceptual training on further speech production, we also tested another somatosensory condition with a different relationship between sounds in the perceptual training phase and target vowels in the production phase. In this other condition, we tested the same /e/-/ø/ vowel contrast in production, but replaced the “horizontal” contrast in the perceptual training phase by a “vertical” contrast involving the categorization of an acoustic continuum between /i/ and /e/. The same horizontal rearward direction of the skin stretch perturbation was used, so that there was a possible incompatibility between the natural auditory dimension (vertical) and the somatosensory directivity (horizontal). At the current state of our knowledge very little is known about how close the pairing between somatosensory stimulation and the tested auditory dimension needs to be in order to induce a change in production. This other somatosensory condition was designed to shed more light on this issue.

This study has been accepted in the “Journal of Speech, Language and Hearing Research”

Ashokumar, M., Schwartz, J.-L., & Ito, T. (2023). Changes in speech production following perceptual training with orofacial somatosensory inputs. [In press, Journal of Speech, Language and Hearing Research]
--

## 5.2 Methods

### 5.2.1 Experimental Procedure

Forty-five French native speakers (age range 18-41, 33 female) participated in the study. They were submitted to a perceptual training task based on a vowel identification task and a vowel production task before and after the training task. We tested three different groups based on the nature of the stimulation: purely auditory (CTRL) or involving an additional somatosensory stimulation (SOMA1 and SOMA2). Fifteen participants were assigned to each of the three training conditions.

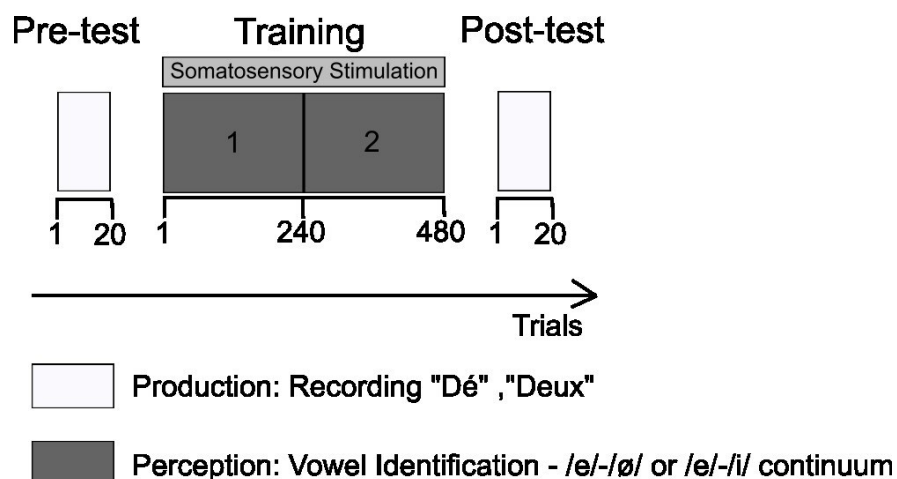


Figure 14: Experimental procedure associating a perceptual training phase with a pre- and post-test for speech production recording.

The test was carried out in the order: production session (pre-test), perceptual training session, and production session (post-test) as shown in Figure 14. The production task was explained in the section 3.3. We recorded ten times each of the task speech sounds “dé” and “deux”. The perceptual training session consisted in the vowel identification test with somatosensory stimulation. The procedure is similar to the one explained in section 3.4. In this study, the somatosensory stimulation, when present, was applied in all trials in the training phase. In total, 480 trials (60 blocks of 8 trials) were tested. We first tested two groups of participants using the /e/-/ø/ continuum, respectively with somatosensory stimulation (SOMA1) and without the stimulation (CTRL). We also tested another group of participants using the /e/-/i/ continuum with somatosensory stimulation (SOMA2). The /e/-/i/ continuum was produced using Praat (Boersma & Weenink, 2001) based on the same edge /e/ sound as in the /e/-/ø/ continuum. From this edge, F1 values were equally shifted downwards in steps of 17 Hz, and F2, F3 and F4 values were equally shifted upwards in steps of 20, 123, and 18 Hz respectively, while F5 was kept constant. The five formants in the resultant edge vowel /i/ are 247, 2062, 3372, 3677, 4000 Hz.

Note that we did not carry out a pure auditory testing of the perceptual boundary before/after the training task in order to avoid changes in production due to repetitive hearing of the stimulus sounds in such a perception test.

## 5.2.2 Statistical Analysis

All the analyses were directed towards the changes in formant values from the pre- to the post-test, depending on group and participant.

For each participant and for each utterance recorded in the pre- or post-test, the first three formants were extracted using PRAAT (Boersma & Weenink, 2001). We first detected the onset and offset of the vowel in the task utterances using the RMS power. The vowel onset was determined as 30ms after the closest minimum point before the peak RMS power during the vowel production. In our observation, this point was close to the release burst for the consonant. The vowel offset was determined as the point becoming 40% of the peak RMS power. A representative example is shown in Figure 15 with RMS power. The detected onset and offset (vertical dashed lines) appropriately corresponded to the onset and offset of the vowel sound. We visually checked that this temporal detection process worked well and consistently in this set of quite constrained and simple CV forms with always the same consonant and only two possible vowels.

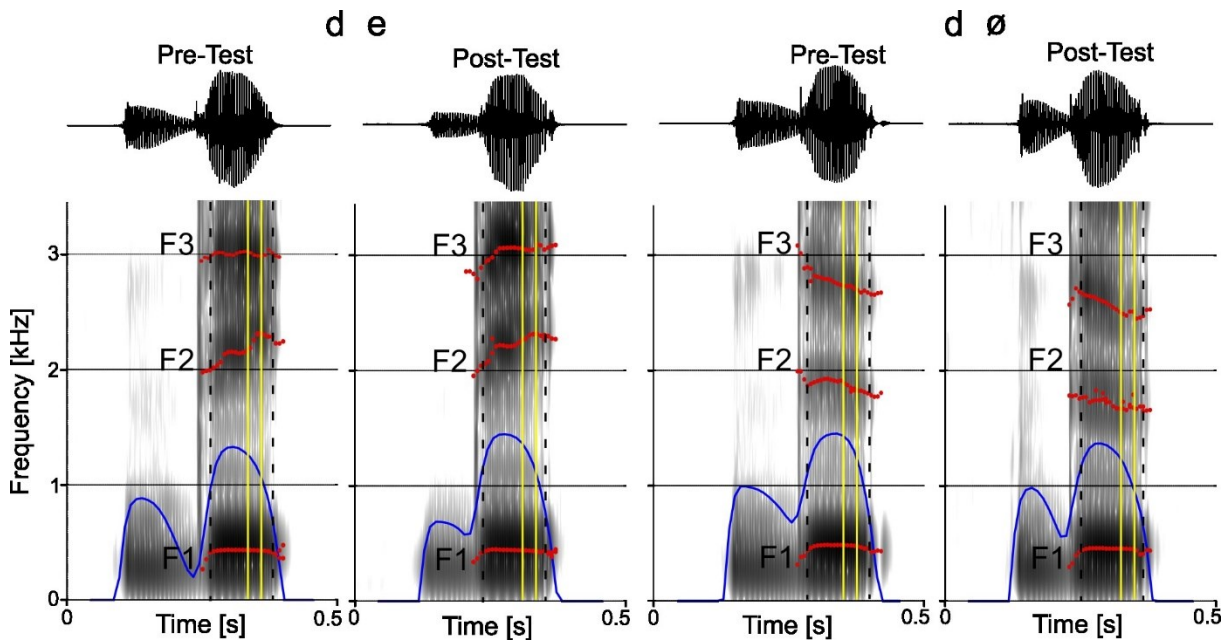


Figure 15: Representative speech signals of the test words “dé” and “deux” recorded in the production sessions (pre- and post-test). Top panel represents acoustic signals and bottom panel represents corresponding spectrograms with detected first, second and third formants (red dotted lines) and RMS power (blue solid line). Two black dashed lines show the vowel region detected based on RMS power. Two vertical yellow solid lines delimit the temporal window used for formant detection in the following analyses.

A 25-ms time window was centered at 75% of this vowel duration (displayed by the range between the two yellow vertical lines in Figure 15). This time point was based on a preliminary analysis showing that the latter portion of the vowel production, less dependent on coarticulation associated to the prefixed consonant /d/, showed a relatively larger difference between pre- and post-tests than in earlier portions. As a matter of fact, we also tested two other temporal positions respectively centered at 25% and 50% of the vowel duration, but the corresponding data did not lead to statistically significant effects, hence in the following we

only report the data for the 75% time point. To reduce individual variability, the extracted formants were normalized for each speaker using z-scores in each vowel and each formant separately.

We first assessed possible pre- vs. post-test changes separately for each vowel and each formant. To assess a global change across the vowels, we also evaluated pre- vs. post-test changes for each formant, averaged across the vowels /e/ and /ø/. All these analyses were based on Linear Mixed Effect analysis using the lme and nlme package (Pinheiro et al., 2023) in R. In these analyses we evaluated models predicting averaged z-scored formant values from fixed effects of group (CTRL, SOMA1, SOMA2), pre/post training and their possible interaction. The random intercepts were intended to account for participant dependent changes in intertrial variability between pre- and post-test. The anova() function in R was used in nested model comparison to determine whether fixed effect interaction was supported. The anova() function was also applied to model output to perform F-tests of the significance of effect contributions. We also applied multiple comparison using the glht function in the multcomp package according to the method by Hothorn et al. (2008) in R. In this method, the comparison matrices are between pre- and post-test for each group and between groups for each pre- and post-test. Then, we examined whether the observed production changes were related to the perceptual change during the training phase. For this aim, we first estimated the psychometric function of the identification scores along the 8-member continuum, respectively in the first and second half of the training session (240 stimuli each altogether, see Figure 14). This provided estimates of the category boundary as the 50% value of the psychometric function, and the effect of training was estimated by the difference in the corresponding categorical boundaries from the first to the second half of the training session. From there, we carried out correlation analyses with Pearson correlation coefficients between the pre- vs. post-test formant changes averaged across the two vowels in production for each speaker and the first- vs. second-half psychometric boundary shift in the perceptual training session for the corresponding speaker.

Considering the complex portrait of the formant changes in vowel production in the groups, we also applied directional statistics using the CircStat toolbox (Berens, 2009) in MATLAB. We examined which direction of the formant change from pre- to post-test was predominantly induced in each group in the F1-F2 and F2-F3 vowel spaces, respectively. For this aim we calculated the directional angles of the averaged pre- to post-test production differences in the F1-F2 and F2-F3 space for each group, and tested the departure from uniformity in the distribution of these angles using the Rayleigh test. We also applied a Watson-Williams test to compare between two groups only when these two groups showed non-uniformity in the

distribution of the directional angles.

Finally, to assess whether there was also a change in the sound amplitude and vowel duration between pre- and post-test, we extracted the peak amplitude of the RMS power and vowel duration for each speech sound trial. We examined the difference between pre- and post-test by a linear mixed effect model with post-hoc multiple comparison tests.

### 5.3 Results

Figure 15 shows a representative example of the recorded utterances “dé” and “deux” in pre- and post-tests in the SOMA1 group. Red dots represent the extracted formants. As shown in this figure, from pre- to post-test, F3 of /e/ seems to increase while F2 of /ø/ seems to decrease.

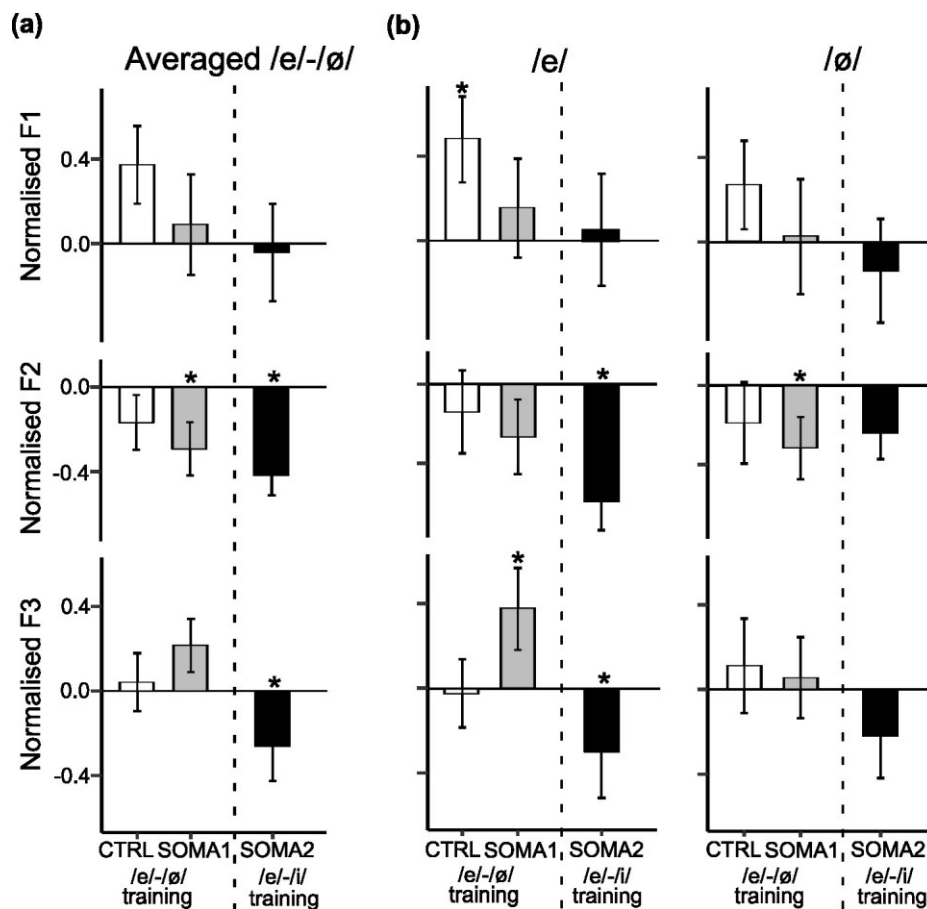


Figure 16: Pre-post change in F1, F2 and F3 in each group (SOMA1, CTRL and SOMA2) with (a) the values averaged across speakers and across vowels /e/ and /ø/ and (b) the values averaged across speakers in each vowel /e/ and /ø/, separately. Error bars represent the standard error across participants. The mark \* correspond to  $p < 0.05$  by multiple comparisons in LME analysis.

We first evaluated the differences averaged across the two vowels. Figure 16(a) shows the difference between pre- and post- test in each formant and each group. In F1, we did not find any significant difference in the interaction effect ( $F(2,1718) = 1.91, p = .15$ ) nor in the main effect (Group  $F(2,42) = .0015, p = .99$  and pre-post test  $F(1,1718) = 2.66, p = .1$ ). The averaged

value showed larger increase in CTRL, although it did not reach a significant level [ $z = 2.46, p = .098$ ] in multiple comparison. In F2, there was no significant difference in the interaction ( $F(2,1718) = 2.22, p = .11$ ), nor in the Group ( $F(2,42) = .001, p = .99$ ), but we found a significant difference in pre- vs. post-test ( $F(1,1718) = 37.76, p < .0001$ ). Multiple comparison showed a significant decrease in SOMA1 ( $z = -3.69, p < .01$ ), and in SOMA2 ( $z = -5.15, p < .001$ ) but not in CTRL ( $z = -2.12, p = .21$ ). In F3, there was a significant interaction effect ( $F(2,1718) = 6.29, p < .01$ ). A significant change was seen in SOMA2 ( $z = -2.80, p = .039$ ), but not in SOMA1 ( $z = 2.16, p = .20$ ) nor in CTRL ( $z = .43, p = .99$ ). Table 3 shows a summary of these statistical results.

Table 3: Summary of statistical results based on linear mixed-effect model analysis and corresponding multiple comparisons.

Formant	group	pre-post	interaction
<b>F1</b>	$F(2,42) = .0015, p = .99$	$F(1,1718) = 2.66, p = .1$	$F(2,1718) = 1.91, p = .15$
<b>F2</b>	$F(2,42) = .001, p = .99$	<b><math>F(1,1718) = 37.76, p &lt; .0001</math></b>	$F(2,1718) = 2.22, p = .11$
CTRL		$z = -2.12, p = .21$	
SOMA1		<b><math>z = -3.69, p &lt; .01</math></b>	
SOMA2		<b><math>z = -5.15, p &lt; .001</math></b>	
<b>F3</b>	$F(2,42) = .0002, p = .99$	$F(1, 1718) = .014, p = .91$	<b><math>F(2,1718) = 6.29, p &lt; .01</math></b>
CTRL			$z = .43, p = .99$
SOMA1			$z = 2.16, p = .20$
SOMA2			<b><math>z = -2.80, p = .039</math></b>

We also evaluated the change in production separately for each vowel. Figure 16(b) shows the corresponding differences between pre-test and post-test values. The changes appear to differ between groups. In the CTRL group the change was seen mostly in F1, increasing for both vowels though the only significant change concerns /e/ [ $z = 2.95, p = .025$ ]. In the SOMA groups, the changes rather concern F2 and F3. F2 decreases for both vowels. In the SOMA1 group, the change was significant for /ø/ [ $z = -2.83, p = .037$ ], but not significant for /e/ [ $z = -2.38, p = .12$ ]. In the SOMA2 group, the change was significant for /e/ [ $z = -5.22, p < .001$ ], but not for /ø/ [ $z = -2.09, p = .226$ ]. F3 increases only for /e/ in the SOMA1 group ( $z = 3.23, p < .01$ ) and F3 decreases in both vowels in the SOMA2 group, although only the change for /e/ was significant [ $z = -2.89, p = .03$  for /e/ and  $z = -1.93, p = .31$  for /ø/]. Altogether, these results suggest that introducing somatosensory inputs during perceptual training affected F2 and F3.

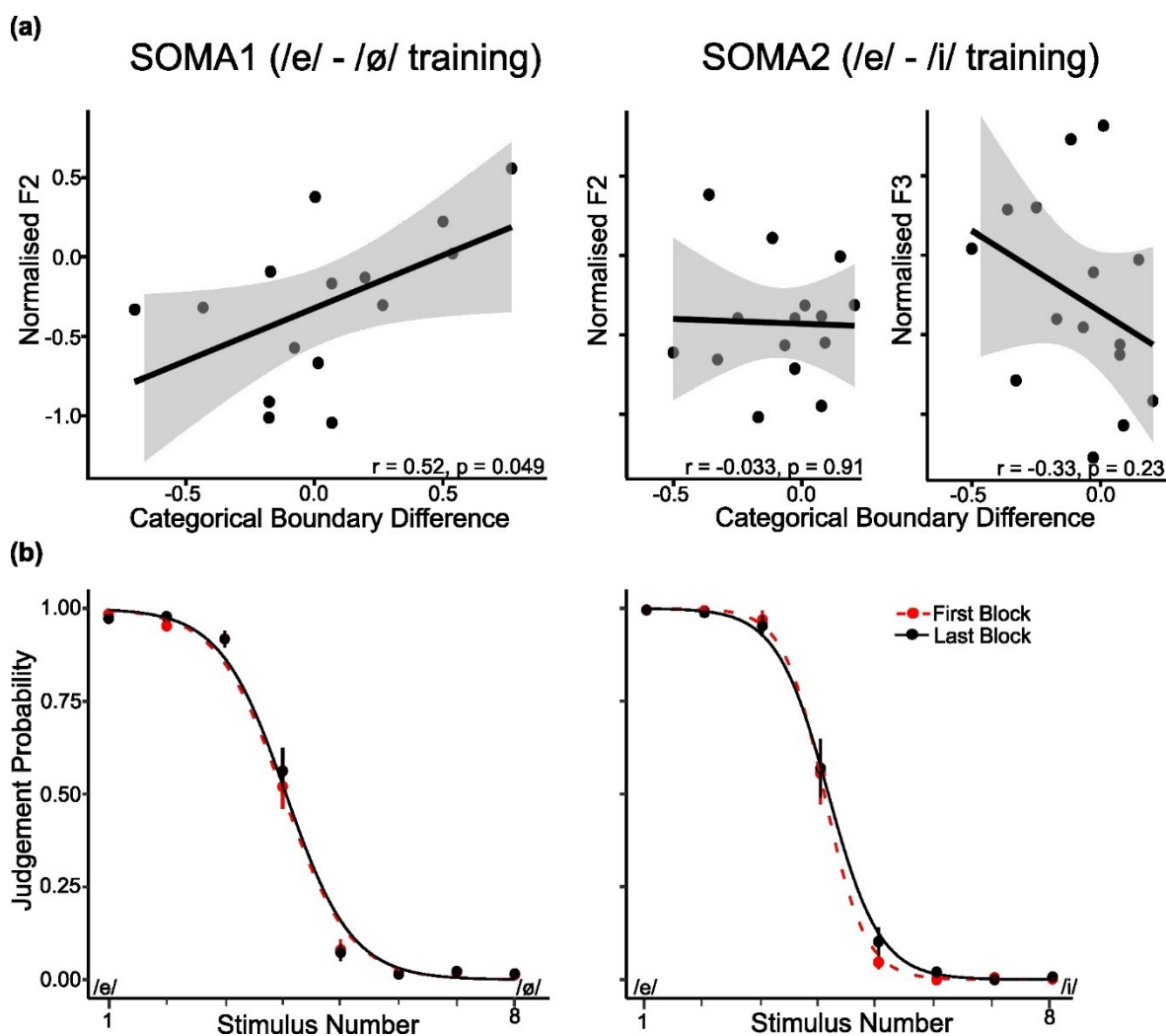


Figure 17: (a) Scatterplots of the per-speaker average pre-post change in formant related to the categorical boundary difference between the first and second halves of perceptual training, for F2 in the SOMA1 group and F2 and F3 in the SOMA2 group. The solid line represents the regression line and the shaded area represents 95% confidence intervals. (b) Psychometric functions during perceptual training averaged across participants, for the SOMA1 and for the SOMA2 group. Red dashed lines represent the one in the first half of the training and black solid lines represent the one in the last half of the training. Error bars represent the standard error across participants.

Although we did not find any reliable difference in the amount of perceptual change between the first and the second half of the training phase (SOMA1:  $F(1,14) = .22$ ,  $p = .65$ ; SOMA2:  $F(1,14) = 2.51$ ,  $p = .14$ ) as seen in Figure 17(b), we carried out a correlation analysis among participants between the production change averaged over /e/ and /ø/ and this amount of perceptual change in training since the change in perception was actually distributed in both directions of perceptual shift across participants. This analysis was applied to acoustic parameters which showed a significant difference for the pre- vs. post production task in the previous analysis, that is for the F2 change for the SOMA1 group and for both F2 and F3 changes for the SOMA2 group. Figure 17(a) represents the scatterplot with the regression line



and 95% confidence interval. There was a significant correlation for F2 in SOMA1 ( $r(13) = .52, p = .049$ ), which suggests that the perceptual change along training might be related to the change in production. In SOMA2, we did not find any significant correlation in both F2 and F3. This is perhaps due to the fact that the range of variations of the amount of perceptual change during training was much smaller in SOMA2 than in SOMA1 (see Figure 17(a)).

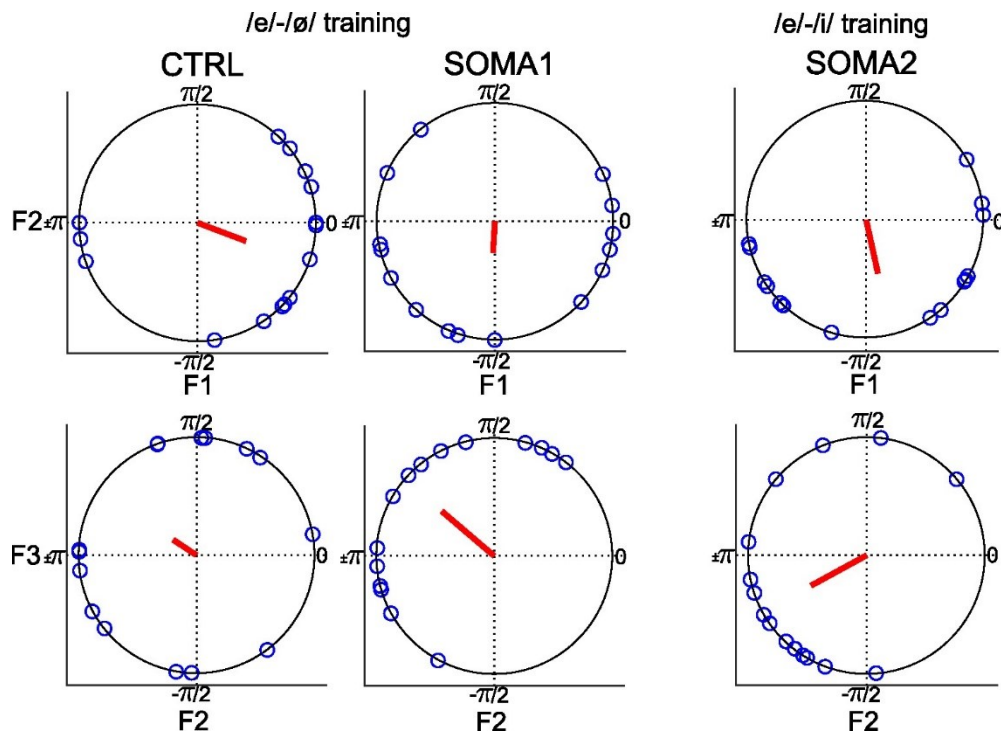


Figure 18: Direction of formant change in F1-F2 (top panels) and F2-F3 (bottom panels) spaces for SOMA1 (left panels), CTRL (middle panels) and SOMA2 (right panels) groups. The blue circles on the unit circle represents the angles for each participant. Red solid lines indicate the mean vector representing the formant change. When the length of this vector is 1, all individual points are superposed and correspond to the same change angle. When the length of this vector is 0, all individual points are distributed uniformly in all directions.

In an attempt to globally embrace the changes in several formants at the same time, we applied directional statistics in the F1-F2 and F2-F3 formant spaces. Figure 18 shows the corresponding angles of the changed direction for the participants in each group. The blue circles put on the unit circle present the angles in each individual. The red solid lines represent the vector of the mean changed direction. In the two SOMA groups, the mean change direction in the F1-F2 space consistently points along the F2 axis in the negative direction. The individual angles are compatible with a uniform distribution in SOMA1 ( $p = .3$ ), but not in SOMA2 ( $p = .04$ ). In contrast, the mean change direction in the F2-F3 space was different between the two groups, pointing along the direction of F2 negative and F3 positive in SOMA1 and the direction of F2 negative and F3 negative in SOMA2. A Rayleigh test showed a non-uniform distribution in

both groups ( $p = .004$  and  $p = 0.01$ , respectively). Since both are non-uniform distributions, we further examined if the direction of change is different between SOMA1 and SOMA2. A Watson-Williams test showed a significant difference ( $p < .05$ ), suggesting that the training effect on production is different between the two groups. On the contrary, in the CTRL group the change directions are rather unimodal in F1-F2 ( $p = .05$ ) but not in F2-F3 ( $p = .4$ ). These results are consistent with the previous ones and globally support the conclusion that the directions of pre- to post-test change are different between the SOMA and CTRL groups.

We then wondered whether the F1 increase in the CTRL group from pre- to post-test, likely related to an increase in jaw opening, could be accompanied with an increase in vocal effort resulting in an increased amplitude during the post-test production. To assess this possibility, we examined the change in amplitude between pre- and post-test averaged across the vowels using multiple comparison based on Linear Mixed-effect Model as done in the formant analysis. We found no significant change in all three groups, CTRL: difference  $-.011 \pm .014$  (SE),  $z = -.082$ ,  $p = .95$ ), SOMA1: difference  $-.0024 \pm .014$  (SE),  $z = -.17$ ,  $p = .99$ ) and SOMA2: difference  $-.021 \pm .014$  (SE),  $z = 1.455$ ,  $p = .61$ ). Hence the F1 increase in the CTRL group is not associated to an increase in voice intensity. In addition, we carried also out the analysis with the vowel duration. We did not find any reliable change between pre- and post-tests averaged across the vowels in each group; CTRL: difference  $-7.75 \pm 5.12$  (SE),  $z = -1.51$ ,  $p = .58$ ; SOMA1: difference  $4.92 \pm 4.68$  (SE),  $z = 1.05$ ,  $p = .86$ ; SOMA2: difference  $-8.97 \pm 4.70$  (SE),  $z = -1.91$ ,  $p = .32$ ). Hence, altogether, it appears that perceptual training did not affect vowel parameters other than formant frequencies.

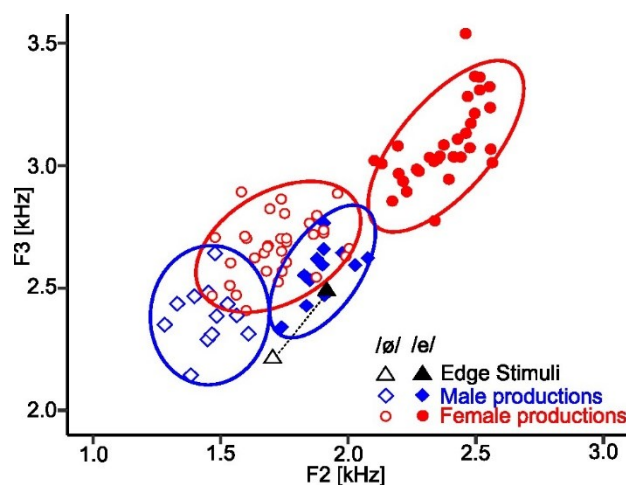


Figure 19: F2-F3 values of the vowels in the production session before perceptual training (blue squares for male speakers, red circles for female speakers) and of the edge sounds in the /e/-/ø/ continuum used for perceptual training (black triangles). Filled dots for /e/ sounds, empty dots for /ø/ sounds. Ellipses represent 95% confidence intervals separately for male and female productions and

*for each of the two vowels.*

Finally, we checked the conformity of the speakers' production of the /e/-/ø/ contrast with edge acoustic stimuli in the perceptual training /e/-/ø/ continuum. Figure 19 shows F2-F3 values obtained in the production task in the pre-test (baseline) for the whole set of the 45 speakers in the three groups, together with F2-F3 values for the acoustic stimuli in the continuum. We observe a qualitative coherence between the synthetic perceptual stimuli and the speakers' productions with a joint increase in F2 and F3 from /ø/ to /e/, with of course the typical increase in formant values for female speakers in the group.

## **5.4 Discussion**

The current study showed that repetitive exposure to paired auditory-somatosensory inputs during perceptual training modified to a certain extent the content of subsequent speech productions. Perceptual training in the /e/-/ø/ region associating auditory synthetic stimuli with orofacial somatosensory stimulation due to horizontal backward facial skin deformation led to formant modifications in the later production of the corresponding vowels, with a global small decrease in F2 averaged across the two vowels and larger for /ø/ and a slight increase in F3 for the vowel /e/. The change in F2 was significantly correlated with the size and direction of the perceptual shift during the training phase. Perceptual auditory training in the /e/-/i/ region involving the same backward orofacial somatosensory stimulation also led to formant modifications in the later production of the corresponding vowels, though with a different pattern. Indeed, there was a similar decrease in F2 but now a decrease in F3, and these changes were larger for /e/ than for /ø/, and they displayed no correlation with the perceptual shift during the training phase. In contrast, none of these appeared when the training phase was purely auditory, with only a significant increase in F1 for both vowels. Directional analysis confirmed the existence of average unimodal directional changes in the F2-F3 space, though differently in the two somatosensory groups SOMA1 and SOMA2, while the only visible change in the control group, related to F1, occurred in the F1-F2 space. Taken together, these findings support the assumption that paired auditory-somatosensory inputs may play a significant role in speech learning. Still, they raise of course some questions.

The first question that arises concerns the direction of the main change observed in both SOMA groups, that is the global decrease in F2. In the SOMA1 group, the somatosensory stimulation associated with backward skin deformation during perceptual training should drive the participants' percepts towards /e/ rather than /ø/ (Ashokumar et al., 2023; Trudeau-Fisette et

al., 2019). Interestingly, this later drives their production in the other direction, in a kind of compensatory strategy. This is actually compatible with previous findings concerning the perceptual change following perceptual training (Ito & Ogane, 2022). In this study, perceptual training was carried out using a vowel identification task between / $\epsilon$ / and /a/. Somatosensory stimulation associated with upward facial skin deformation appeared to induce a perceptual shift towards / $\epsilon$ /, that is, the participants perceived / $\epsilon$ / more than /a/ in the training phase. But a resultant perceptual change after training was induced in the opposite direction, that is, the participants perceived /a/ more than / $\epsilon$ /. The present study extends this result to a production after-effect, the learning drift towards / $\epsilon$ / due to the somatosensory stimulation later inducing production shifts in the opposite direction for both vowels, that is, towards / $\emptyset$ / by decreasing their F2 value. Interestingly, this after-effect of auditory-somatosensory perceptual learning on later vowel production seems to depend on the direction and size of the perceptual drift in the learning phase, as shown by the correlation analysis reported in Figure 17(a).

In the SOMA2 group where training was carried out with a different vowel continuum, that is / $\epsilon$ /-/i/, a similar F2 decrease was also induced, with a larger decrease in the production of / $\epsilon$ /. Interestingly, this suggests that the learning effect could be local, since only / $\epsilon$ / in the production task is involved in the preliminary perceptual training phase. In addition, we did not find any correlation between the perceptual shift during the training task and further changes in the production session. As shown in Figure 17(a), the variability of the perceptual change during training was relatively small in the SOMA2 group, suggesting a relatively small effect of the somatosensory stimulation in the perceptual training task. This may be due to the unmatched auditory-somatosensory pair involved in this group, with a horizontal somatosensory stimulation associated to a vertical articulatory contrast in auditory stimulation. Still, although the somatosensory effect in the current perceptual training task is small, the somatosensory stimulation might change the processing of the vowel / $\epsilon$ / in both perception and production, hence the larger change for the production of / $\epsilon$ /.

The next question concerns the possible articulatory interpretation of the F2 and F3 changes in the SOMA1 and SOMA2 group. Our expectations, presented in the introduction section, were that changes would possibly be induced in F2 and/or F3, which are the main phonetic correlates of the contrast between the target vowels in this experiment. But, surprisingly, F3 in production appears to be shifted after the learning phase in opposite directions in the SOMA1 group and in the same direction in the SOMA2 group. The consistent decrease of F2 in both groups could be related to either a slight backward move of the tongue or a slight global increase in lip rounding. On the other hand, an increase in lip spreading likely induces the increase in F3, decreasing the

size of the front cavity and increasing the associated front resonance (Lindblom & Sundberg, 1971; Schwartz et al., 1993). Therefore, it seems possible that the modulation in production likely associates both a tongue retraction component responsible for the F2 decrease in both groups and the F3 decrease for the SOMA1 group and an additional lip spreading mechanism producing the F3 increase in the SOMA2 group. At this stage, we must admit that we do not have a convincing explanation of why such a difference between the two groups occurs.

Then comes the question of the interpretation of the F1 increase after pure auditory training in the control group. As shown in previous studies (Cooper & Lauritsen, 1974; Cooper & Nager, 1975; Jamieson & Cheesman, 1987), the repetitive exposure to auditory stimuli alone could indeed affect later production mechanisms. For example, Cooper & Lauritsen (1974) showed that repetitive exposure to consonants resulted in a decreased voice onset time in the later production of the corresponding consonants. Another possibility could be that the F1 increase was just related to an increase in vocal effort, but this assumption has been discarded by the lack of amplitude change following the training session. Therefore, the F1 increase, which seems to occur in both vowels, just likely reflects a global increase in jaw opening (supraglottal) without any difference in vocal effort (glottal). Interestingly, the target vowels /e/ and /ø/ are both mid-high and hence involve the control of a rather high position of the jaw and tongue. The increase in F1 and underlying jaw opening in this context hence appear as a kind of progressive decrease in articulation precision, perhaps just due to the repetition of many utterances of the same vowels. In this context, the fact that the change in F1 following the training session is not seen in the somatosensory group suggests that another effect of the somatosensory stimulation following training could have been to maintain the articulatory precision of the height position for this group compared to controls.

Hence, altogether, the present study provides evidence for the role of the somatosensory modality in the effect of perceptual learning on speech production. Adding a somatosensory stimulation to the auditory task in the perceptual learning phase provides both an increase in the precision of the constrictions required for realizing the production task, and some variations or biases in the definition of the task and the generation of the associated articulatory gestures. The exact way these variations are produced is still not completely clear. Importantly, it appears that (1) there can be an effect of the somatosensory stimulation even if it is not relevant for the perceptual task (as in the SOMA2 group) and (2) there could be some local effects depending on whether the vowel to produce is part of the perceptual training task or not. A number of open questions hence remain at this stage. Firstly, a number of additional associations between somatosensory stimulation and auditory classification should be tested in the future.

Furthermore, it could also be asked whether the effect of the somatosensory stimulation is active, that is related to the perceptual task at work in the training phase, or passive, that is just due to the somatosensory stimulation independent on a training task. It could even be wondered whether the effect of the somatosensory stimulation required any auditory stimuli and task to be displayed in a further production task. These questions pave the way for future studies on the exploration of the role of the somatosensory modality in changes in speech production following perceptual training.

It is noticeable that the current effects in production following perceptual training were relatively small. This is probably related to the limitations of the present study. We just evaluated the formants of the produced sound, without direct information of possible changes in articulatory configurations, and formant measurements may not have enough sensitivity to efficiently represent the articulatory changes following perceptual training. The measurement of articulatory movements is hence required for a better understanding of the somatosensory learning effect in production change. In addition, the current study tested vowels in a different form in the training phase (synthetic isolated vowels) and in the production task (French words). Previous studies suggest that speech motor learning could be fundamentally local (Rochet-Capellan & Ostry, 2011) with limited generalization of motor learning effects to sounds and utterances different from those used in training (Ito & Ostry, 2010). This could have limited the size of perceptual training effects seen in the post-test production phase in the present study, the more so since the production of actual words such as “dé” and “deux” can be well-learned and anchored in the speakers’ repertoire, and hence difficult to modify through perceptual training with other forms of the task sounds. Nevertheless, the current findings are in line with the more general idea that receiving specific pairs of auditory-somatosensory stimulation could be of importance for speech learning. Importantly, the present empirical findings could be used or adapted for possible practical applications in speech training or rehabilitation.

In summary, introducing somatosensory inputs during a perceptual training phase induces changes in speech production. The tendency of changes was different between tested groups, but consistently, the changes in both somatosensory groups were seen in the values of F2 and F3 that are acoustically predominant in the contrast of tested vowels. Repetitive exposure to paired auditory-somatosensory stimulation can play a role in speech learning for both perception and production. This is in line with previous findings demonstrating the role of orofacial somatosensory inputs in speech learning involving the perception-production link (Ito & Ogane, 2022; Ohashi & Ito, 2019).

# Somatosensory function in speech perception in individuals with hearing impairment

---

## 6.1 Introduction

HI Individuals acquire and maintain speech perception and production performance with the help of hearing devices such as HA and/or CI. Recent studies have shown that orofacial somatosensory inputs also participate in speech perception (Ito et al. 2009, Trudeau-Fisette et al 2019). However, it is still unknown how such somatosensory inputs are involved in speech perception in individuals with hearing impairment.

We explored auditory-somatosensory interactions in HI individuals and assessed whether the somatosensory effect in speech perception can be found in HI individuals and how the effect is related to their hearing profile. To study participants who are native speakers of French, we applied a similar paradigm as the one used in Trudeau-Fisette et al. (2019). The experimental paradigm exploits a vowel identification task using a speech continuum between the French lip-spread vowel /e/ vs. lip-rounded vowel /ø/. Considering that these two vowels are contrasted by horizontal labial gestures, the somatosensory stimulation associated with facial skin stretch was applied in the backward direction. The target somatosensory effect in speech perception was quantified as the difference in judgement probability between vowel auditory categorization scores with and without somatosensory stimulation. By testing HI participants with variable use of hearing devices (CI and/or HA), we examined whether and how the somatosensory effect in speech perception was related to their basic hearing performance represented by hearing thresholds in noise screened in a separate test, and to their hearing profile.

## 6.2 Methods

### 6.2.1 Experimental procedure and data analysis

We tested twenty-five NH individuals (aged 18-31) and twenty-one individuals with bilateral hearing impairment wearing HA and/or CI (aged 18-74). Participants of both groups were native French speakers and participated in the same experimental procedure.

The main experimental procedure was a vowel identification test with somatosensory stimulation as described in section 3.4. We also assessed the participants' hearing threshold by using the modified version of the French Digit Triplet test as described in section 3.5. Note that we did not test thresholds higher than 0 dB SNR in this test. In addition to this, HI participants were asked to fill out a hearing assessment questionnaire to get information about their hearing profile such as the onset of deafness, their used hearing devices and duration of usage. These two tests together with the filling of the questionnaire were carried out in one day. We first carried out the vowel identification task. Then the speech-in-noise test followed after removing the setup for somatosensory stimulation. HI participants were asked to fill the questionnaire before those two tests. We compared between the somatosensory effect in speech perception and hearing ability represented by hearing threshold in noise and hearing profile concerning implantation age, type of hearing device and the usage duration of hearing device.

### 6.3 Results

We first report the results from the NH participants. Their hearing thresholds reached  $-8.5 \pm 0.1$  dB SNR. This level is slightly higher than the thresholds obtained in previous studies:  $-10.5 \pm 0.3$  dB SNR in French participants (Jansen et al., 2010) and  $-11.2 \pm 1.3$  dB SNR in Dutch participants (Smits et al., 2004), probably because of a difference in stimulus presentation environment (headphones vs. loudspeaker).

Considering the vowel identification task, we successfully replicated the results from Trudeau-Fisette et al. (2019) about the somatosensory effect in this task. The results averaged across participants are shown in the top panel of Figure 21. As shown here, NH participants identified the vowels correctly in the auditory-only condition. In the condition with somatosensory stimulation, the perceptual boundary between /e/ and /ø/ slightly shifted toward /ø/, indicating that the NH participants identified more sounds as /e/. Repeated-measures ANOVA showed a significant increase in judging the vowel as /e/ in the SOMA condition ( $F(1,24) = 6.7, p = 0.02$ ). Hence, as shown in the first study reported in the current thesis, the somatosensory effect in speech perception can be consistently induced in the current open-air situation using a loudspeaker rather than headphones.

Let us now describe the results for the HI participants. Their performance was quite varied in both the vowel identification test and the hearing threshold screening. Figure 20 summarizes



those results in the two perception tests together with the hearing profile for each HI participant. This figure is organized around two criteria that appeared to us as able to let emerge different portraits of HI participants. Firstly, their ability to categorize the vowels auditorily is more or less directly related to their hearing thresholds, and secondly, the somatosensory effect in speech perception seems to appear only for specific profiles of participants with hearing impairment.

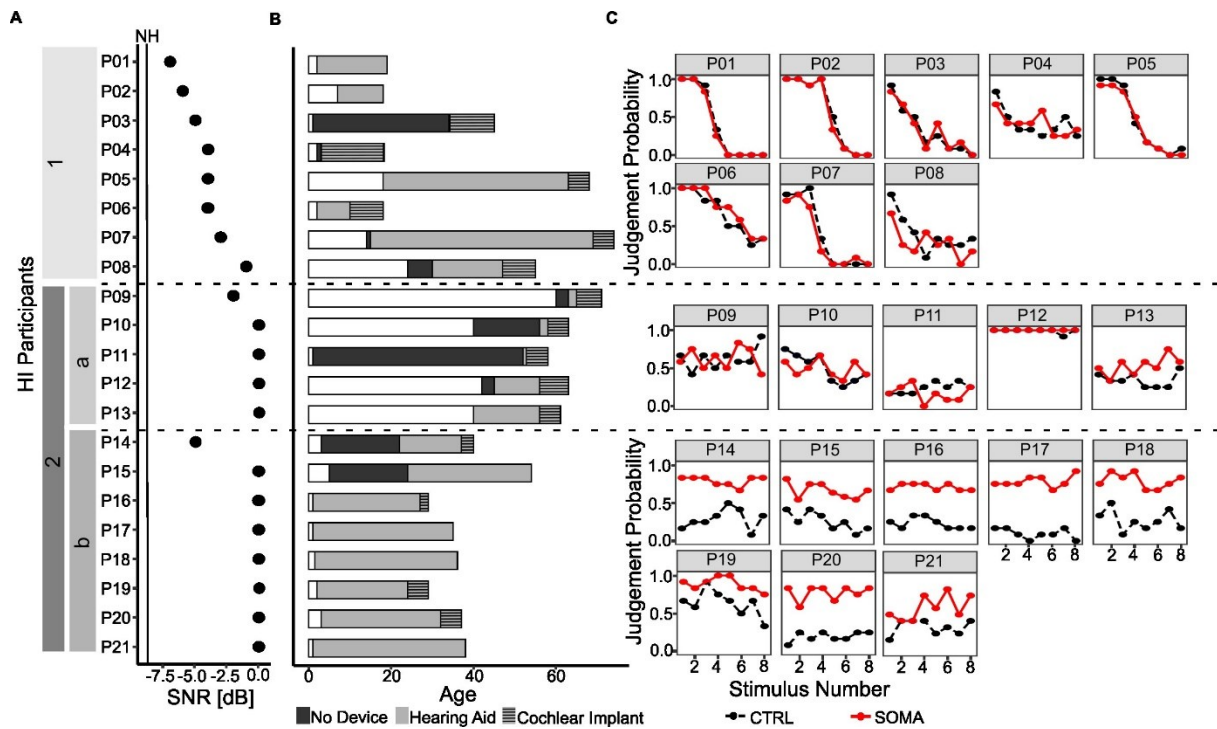


Figure 20: Results from the perception tests and hearing profile in each hearing impaired (HI) participant (A) Hearing thresholds obtained in the speech-in-noise test. Vertical line labeled NH represents the averaged threshold from normal-hearing individuals ( $-8.5 \pm 0.06$  dB SNR). (B) Hearing profile across lifetime obtained by questionnaire. The entire length corresponds to their age. Filled area represents the duration after the detection of hearing-impairment. In this filled area, the black part represents the duration of not using any hearing device, the light grey part represents the duration of using a hearing-aid, and the striped gray part represents the duration of using cochlear implants in either one or both ears. (C) Judgement probabilities in the vowel identification test. Red points and lines represent the condition with somatosensory perturbation (SOMA). Black ones represent the auditory-only condition (CTRL)

For the first criterion, the vowel identification performance visible in Figure 20(c) (black dots and lines) shows that more than half of HI participants (13 of 21, from P09 to P21) were not able to identify the vowels correctly in the auditory-based condition (without somatosensory stimulation). Indeed, they showed mostly flat pattern of judgement probabilities across the stimulus continuum. We categorized those thirteen participants as a second group HI-2 contrasting them with a first group HI-1 of eight participants (P01 to P08). Participants in this first group HI-1 displayed more or less sigmoidal patterns of auditory classification progressively switching from one category to the other, particularly for some participants (P01,

P02, P05 and P07) who showed patterns similar to those of NH participants, while some others still showed some difficulty to identify the vowels.

These patterns of auditory categorization seem rather well correlated with the participants' hearing threshold in noise displayed in Figure 20(a). Indeed, most participants in HI-2 (besides two exceptions) showed a high hearing threshold in noise ( $\geq 0$  dB SNR). On the other hand, participants in HI-1 showed relatively low hearing thresholds between -7 and -1 dB SNR, though still higher than those of NH participants ( $-8.5 \pm 0.1$  dB SNR, displayed by the vertical solid line in the figure). It hence appears that a sufficiently low hearing threshold in noise can be required to identify the vowel correctly, in line with the general idea that a minimal hearing ability is required to capture the sound characteristics correctly.

For the second criterion, let us consider the amount of the somatosensory effect that was qualified as a change of judgement probabilities between auditory-somatosensory (SOMA) and auditory-alone (CTRL) conditions in Figure 20(c) (red vs. black dots and lines). It appears that, while the participants in HI-1 were able to identify the vowels in a relatively correct way, they did not show any clear differences between the two conditions. Instead, a strong perceptual bias due to the somatosensory stimulation was seen in several participants in HI-2. We further divided HI-2 into two sub-groups. Eight participants who showed a perceptual bias were grouped in HI-2b and the others in HI-2a. The participants in HI-2b all displayed judgement probabilities in the SOMA condition much higher than the ones in the CTRL condition, indicating that the participants perceived the stimulation more as /e/ when the somatosensory stimulation was applied. In contrast, participants in HI-2a showed no clear difference between conditions. These two sub-groups have a clear difference in their hearing profile. Indeed, Figure 20(b) displays the participants' hearing profile over their lifetime. For each participant, the entire length of the bar corresponds to the age. The gray and striped gray parts represent the duration of the usage of hearing devices. The gray part corresponds to HA and the striped gray part corresponds to CI. The black part represents the duration without any hearing devices after the detection of hearing-impairment. It appears that HI-2a were all elderly ( $> 55$  years) and most of them were diagnosed post-lingually. In contrast, HI-2b were diagnosed pre-lingually and had a long duration of usage of a hearing device ( $\geq 18$  years).

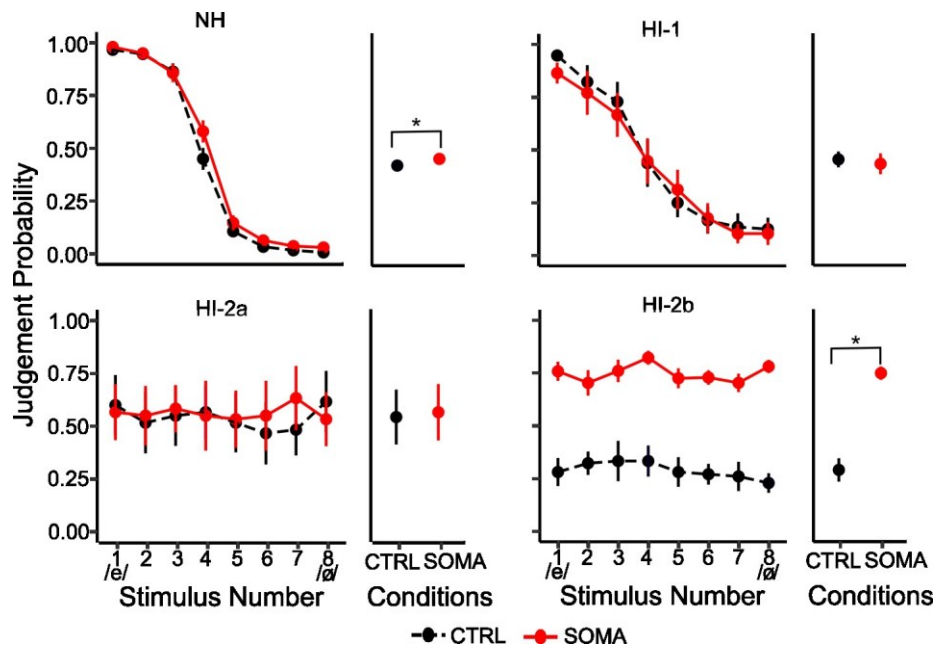


Figure 21: Averaged response of judgement probabilities in the vowel identification test. NH corresponds to normal hearing group. HI-1, HI-2a and HI-2b corresponds to the hearing impairment groups represented in Figure 20. In each panel, left part represents judgement probabilities across each auditory stimulus and right part represents the judgement probabilities averaged across stimulus. Error bars represent standard errors across participants.

This division in three groups is summarized in Figure 21 that displays the patterns of responses in the CTRL and SOMA conditions averaged separately over the participants of the three groups, compared with those of the NH participants. The patterns are rather coherent and clearly display the different behaviors of the participants in these groups. Repeated-measures ANOVA separately applied in each group showed no difference between the conditions in HI-1 ( $F(1,7) = 1.3, p = 0.3$ ) nor in HI-2a ( $F(1,4) = 0.3, p = 0.6$ ). However, participants in HI-2b showed a strongly significant difference between CTRL and SOMA ( $F(1,7) = 74.9, p < 0.001$ ).

## 6.4 Discussion

In the current study, we examined the role of somatosensory inputs in speech perception in HI individuals. This was assessed in the context of the change in vowel identification performance due to somatosensory stimulation. We adopted the finding in Trudeau-Fisette et al. (2019) that the addition of a backward somatosensory perturbation in an auditory vowel identification test between the French vowels /e/ and /ø/ biases the identification of the presented sound towards the percept /e/. We successfully replicated this finding in NH individuals with the open-air presentation of auditory stimulus. When we applied the same procedure to HI individuals, we found that the somatosensory effect in speech perception varied a lot and that these variations could be related to their hearing thresholds in noise, evaluated in a separate test, and to their

hearing-impairment profiles. We categorized our participants into three groups. The first group consisted of participants who had a relatively lower hearing threshold in noise and were able to auditorily identify the vowel sounds with reasonable precision. It appears that this group did not show a consistent somatosensory effect. The second group consisted of participants with a high hearing threshold in noise ( $\geq 0$  dB SNR), who were not able to auditorily identify the vowel sounds in the continuum. This group was further divided into two subgroups based on the somatosensory effect displayed in the vowel identification test. Participants in the first subgroup did not show any perceptual bias when the somatosensory stimulation was applied. They consisted of elderly participants who did not dispose of a hearing device usage before the age of 40 years old or later, and had a relatively short duration of experience with these devices (6 years at best, 21 years at most). The second subgroup showed a large perceptual shift with somatosensory stimulation. Most of the participants in this subgroup were pre-lingually deaf, most of them were equipped of hearing devices very early ( $\leq 3$  years old for 6 over 8 participants) and with a long duration ( $> 18$  years) of hearing device usage. Taken together, these findings indicate that the somatosensory effect in speech perception can be developed in HI individuals although it is dependent on their hearing ability and history. We will now attempt to interpret how these three varied portraits of somatosensory effect might have emerged.

The only participants displaying a large perceptual bias induced by the somatosensory stimulation are those of the group HI-2b. The participants in this group, mostly prelingually impaired and equipped early with a long duration of usage of their equipment, used mostly HA with no or very reduced use of CI. Since HA simply amplify the sounds coming to the ear while CI provide direct stimulation to the auditory nerve after bypassing the damaged regions of the ear, it is likely that their auditory system does not perform well, hence their high hearing thresholds in noise and their poor vowel identification auditory performance. Still, their long duration of usage of hearing aids, mostly from the prelingual age, may have habituated these participants to hear speech sounds in poor hearing conditions, and have driven them to rely to a large extent on any available additional information. This could explain the large bias they display when experiencing the somatosensory stimulation while listening to the speech sounds in the continuum. This could be related to the studies about audio-visual integration, mentioned in section 2.3.1 showing that, for CI users at least, early implantation (typically before 2.5 years of age) is required to obtain efficient integration (Schorr et al., 2005). Six over the eight participants in this group benefitted from hearing aids at an early age ( $\leq 3$  years) possibly compatible with the “critical period” that seems optimal for acquiring efficient multisensory

integration behavior (see also Kral & Sharma, 2012). Participants in this subgroup display a case of “somatosensory dominance” well in line with data on visual dominance in the McGurk effect in CI users (Desai et al., 2008; Rouger et al., 2008; Schorr et al., 2005). Interestingly, they display a unique situation (never explored in previous studies) where it seems possible for these participants to completely recognize a vowel just based on the somatosensory input, since the auditory information provides essentially no valid information in the categorization task for these subjects. In the current study, we could unfortunately not find adult participants who received CI before the mentioned critical period, probably due to the fact that receiving CI in infants was less popular twenty years ago and more. Since prelingually deaf children using CI show better speech perception and language development when compared to those using conventional HA (Bittencourt et al., 2012), we can expect that individuals who receive CI before those critical periods could develop real auditory-somatosensory integration associating efficient audition and significant somatosensory bias.

In this context, the results displayed by participants in the HI-1 group may appear puzzling. Indeed, these participants were able to identify the vowels with some precision, but showed no somatosensory effect. They had relatively low hearing threshold in noise, suggesting that their auditory system was to some extent more efficient than in the other HI participants. Still, their audition remained poorer than in the NH group, considering both their hearing threshold and identification curves often more blurred than in typical hearing subjects. This could induce a lower reliance on auditory information in this group compared to the NH group. Capitalizing on principles of sensory reliance claiming that interactions between different sensory systems are governed in a statistically optimal way so that senses that relay better information contribute more to perception (Alais & Burr, 2004; Ernst & Banks, 2002), we would predict a role of the somatosensory stimulation at least as large in the HI-1 group as in the NH one. A possible explanation could be that, even though their audition performs rather well, the participants in this group might have to pay more attention in hearing to achieve the proposed experimental task. Indeed, attentional effects are known to intervene in audio-visual integration.: audiovisual integration in speech perception, as evaluated by the size of the McGurk effect, was drastically decreased when the NH participants were concurrently performing an unrelated task, be it auditory, visual (Alsius et al., 2005) or tactile (Alsius et al., 2007). We suppose that allocation of attentional resources to the listening task might have reduced or totally blocked the somatosensory effect in speech perception for this group.

Finally, participants in the HI-2a group were unable to identify the vowels correctly and showed no somatosensory bias effect. This could be due to an age-related effect, since all the participants in this group were over the age of fifty-five years old and older than all participants in HI-2b. The data concerning audio-visual integration in speech perception in older adults with hearing impairment are contradictory, sometimes claiming that integration declines with age (Musacchia et al., 2009; Tye-Murray et al., 2007), others that it actually increases in seniors (Sekiyama et al., 2014; Rosemann & Thiel, 2018). Another interpretation could be that the reduced and late experience these participants had with HA (see Figure 20(b)) did not allow them to pay sufficient attention to the somatosensory modality. In any case, the total absence of somatosensory effect in this group with very poor audition shows that somatosensory cues are much more fragile than visual ones, which was rather expected. Altogether, it remains unclear how aging affects the somatosensory effect in aging NH or HI individuals, and considering the rather limited number of participants in each group in the present study, further investigation is required.

In summary, we found that HI participants may present a somatosensory effect in speech perception but that this effect is highly variable and appears to be dependent on their hearing ability (as represented by their hearing threshold in noise) and on their hearing profile. Future studies should be devoted to better understanding the development and evolution along life of the auditory-somatosensory integration in relation to the hearing ability.

# Speech sensorimotor adaptation in individuals with hearing impairment

---

## 7.1 Introduction

The importance of auditory feedback for learning and maintaining speech production has been frequently examined using an experimental adaptation model based on real-time formant modulation of altered auditory feedback (Houde & Jordan, 1998; Villacorta et al., 2007). When a produced vowel sound is played back with altered formants during a vowel repetition task, speakers adapt their production in the direction opposite to the imposed formant change which gradually returns to original after the removal formant perturbation as an aftereffect. Although this is consistent in NH individuals and in some disorders such as blind speakers (Trudeau-Fisette et al., 2017), Parkinson's disease (Mollaei et al., 2013) and persons who stutter (Daliri et al., 2018), it is still unknown whether HI individuals using hearing devices also show adaptation to real-time formant modulation. Previous studies have shown that the CI individuals can adapt to other types of altered feedback such as delayed auditory feedback and frequency altered feedback (Taitelbaum-Swead et al., 2019) suggesting that such participants could also adapt to formant perturbations. We examined whether HI participants using hearing devices show speech sensorimotor adaptation and how this is related to their speech perception ability and hearing profile.

To this end, we assessed sensorimotor adaptation with an altered auditory feedback task with the target vowel /ø/. The second formant of the productions was shifted in real time. We compared the F2 values across the course of the training in the baseline, hold and after-effect phases. In addition to this, speech perception ability was also evaluated by measuring hearing thresholds in noise and performing vowel identification tests with the target vowels /e/ and /ø/. These measures were compared with the adaptation performance and the hearing profile of HI individuals.

## 7.2 Methods

### 7.2.1 Experimental procedure

We tested fourteen NH and fourteen HI participants. HI participants were either HA or CI users for a long time (>10 years). All participants participated in a within-subjects procedure involving a speech motor adaptation test using altered auditory feedback together with speech perception tests consisting of a speech-in noise test and a vowel identification test.

Firstly, the hearing thresholds in noise were evaluated with the modified version of the French Digit Triplet test described in section 3.5. The hearing threshold in noise was provided by the average of the last seventeen SNRs in the test. Note that in this study, contrary to study 3, we explored thresholds higher than 0 dB SNR when necessary. Then the participants performed a vowel identification task between the French vowels /e/ and /ø/ as described in section 3.3. After these two evaluation tests, the main sensorimotor adaptation test consisted in an altered auditory feedback task. The participants were asked to utter the French word “Deux” with the target vowel /ø/ for 150 times. Across the trials, the second formant of the produced sounds was varied in real time as shown in Figure 22. The first 20 trials provided the baseline phase with unaltered feedback, the next 50 trials comprised a ramp phase with gradual increase in F2, then subsequent 50 trials provided the hold phase with the maximum formant shift amounting to 25% relative to the participant’s second formant. This was followed by the after-effect phase consisting of 30 trials with unaltered feedback. A masking noise at 70dB was applied to prevent the effect of environmental noise and audition of the original unaltered bone-conducted sound.

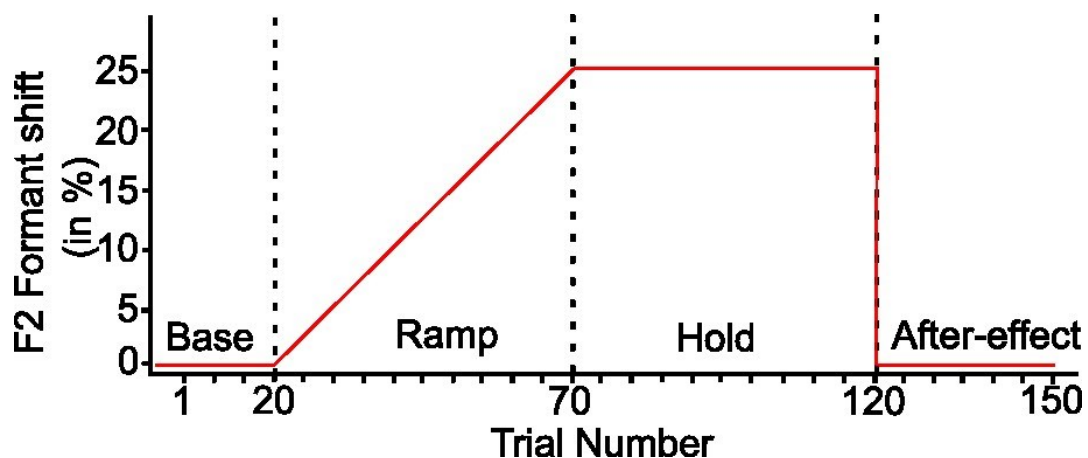


Figure 22: F2 Formant shift in the adaptation task



## 7.2.2 Statistical Analysis

We evaluated the second formant, that was normalized by dividing by the mean value over the last 10 trials of the baseline phase with unaltered feedback. We quantified the adaptation amount by averaging the normalized amplitude at the end of the hold phase (average of the last 10 trials) and at the beginning of the aftereffect phase (average of the first 10 trials). Repeated-measures ANOVA was applied in each of the groups. We also assessed how these measures concerning the adaptation to altered auditory feedback in HI participants were related to hearing ability, evaluated by both the nature of the participant's hearing assistance (CI vs. HA) and by the results of the speech-in-noise and the vowel identification tests.

## 7.3 Results

In the NH participants, the hearing threshold in the speech-in-noise test amounted to  $-9.2 \pm 0.05$  dB SNR. This is close to the thresholds in a previous study ( $-10.5 \pm 0.3$  dB SNR) in French participants (Jansen et al., 2010). The HI participants showed varied performances in both the speech-in-noise test and the vowel identification test. Figure 23 summarises the results from the speech perception tests and the hearing profile for each HI participant. Considering their hearing profile and speech perception abilities, we divided the participants into two groups.

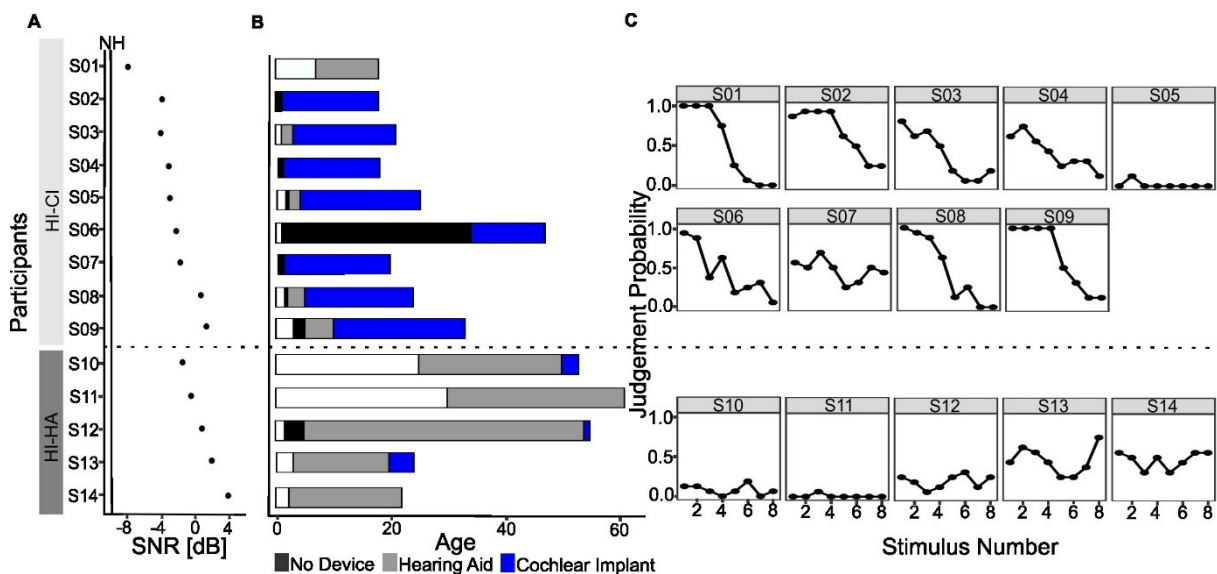


Figure 23: (A) Hearing thresholds obtained in the speech-in-noise test. Vertical line labeled NH represents the averaged threshold from normal-hearing individuals. (B) Hearing profile across lifetime obtained by questionnaire. The entire length corresponds to their age. Filled area represents the duration after the detection of hearing impairment. In this filled area, the black part represents the duration of not using any hearing device, the light grey part represents the duration of using a hearing aid, and the blue part represents the duration of using cochlear implants in either one or both ears. (C) Judgement probabilities in the vowel identification test.

The first group (HI-CI: nine participants) consisted in long-term CI users ( $\geq 13$  years) except S01 who was a HA user with a mild hearing impairment and is included in this group because of similar overall performance. The second group (HI-HA: five participants) consisted in long-term HA users ( $> 16$  years). The first group showed hearing thresholds in noise higher than the ones by NH participants ( $-2.5 \pm 0.9$  dB SNR), but they were able to identify the vowels in the vowel identification test. In contrast, the second group showed even higher thresholds ( $1.1 \pm 0.9$  dB SNR) and had difficulty to identify the vowels correctly. Since the adaptation task using altered auditory feedback requires an ability to detect small differences between the intended produced sounds and actually heard sounds, HI-HA participants may not dispose of the hearing ability required for adaptation, contrary to HI-CI ones. One-way repeated measures ANOVA showed that the second formant was significantly different across the three phases in NH participants ( $F(2, 26) = 54.9, p < 0.001$ ) and in HI-CI participants ( $F(2, 16) = 5.3, p < 0.05$ ) but not in HI-HA participants ( $F(2, 8) = 0.3, p = 0.7$ ). Further analysis was focussed on the NH and HI-CI groups as shown in Figure 24, which shows the changes in F2 productions over the course of trials together with the averaged F2 values in the baseline phase, the hold phase and the after-effect phase. Post-hoc analyses with Tukey's HSD showed that the second formant showed a significant decrease in the hold phase when compared with the baseline phase in NH ( $z = -10.1, p < 0.001$ ) and in HI-CI participants ( $z = -3.2, p < 0.01$ ). However, the formant was significantly different between the hold and after-effect phase in NH ( $z = -2.5, p < 0.05$ ), but not in HI-CI participants ( $z = -2.3, p = 0.07$ ). The formant in after-effect was also significantly different from the one in the baseline in NH ( $z = 7.5, p < 0.001$ ), but not in HI-CI participants ( $z = 0.9, p = 1$ ). Given the difference in averaged standard deviation during the after-effect phase between NH ( $0.05 \pm 0.01$  (SE)) and HI-CI participants ( $0.08 \pm 0.01$  (SE)), the difference in after-effect between the two groups can be due to more dispersion in the HI-CI group. Individual one-sided t-tests between baseline and hold phase were carried out to verify if the adaptation was induced in individual bases. This analysis of individual behaviors showed that 6 out of 9 HI participants and 13 out of 14 NH participants showed a significant F2 decrease. Although the rate of 66% of participants showing adaptation is comparable with previous findings in adaptation studies (Lametti et al., 2012), this rate is smaller than the one in NH participants (93%) in the current data. This smaller adaptation ratio in HI participants can result from the relatively smaller amplitude of adaptation in the averaged data. It could also be due to the unequal size of the groups (14 NH vs. 9 HI-CI).

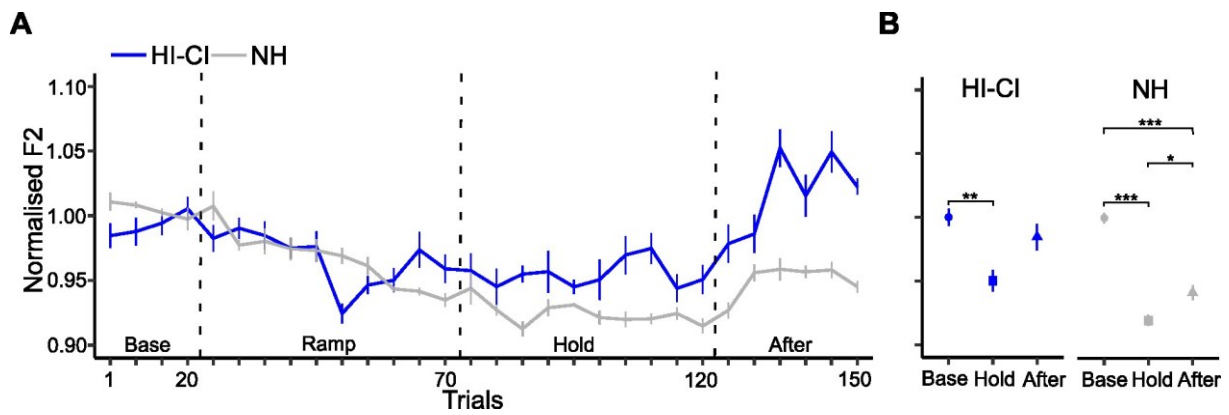


Figure 24: (A) Changes in F2 across trials (B) Averaged formants in baseline, hold and aftereffect phases. The left panel corresponds to the HI-CI group, the right one to the NH group. The error bars represent standard errors across the participants. \*:  $p < 0.05$ , \*\*:  $p < 0.01$ , and \*\*\*:  $p < 0.001$ .

## 7.4 Discussion

This study assessed whether HI participants using hearing devices show speech sensorimotor adaptation and whether this is dependent on factors affecting their hearing. We first confirmed that our experimental setup with altered auditory feedback induces an adaptation in NH participants as shown in previous studies (Trudeau-Fisette et al., 2017). Using this setup, we found that HI participants showed an adaptation only in the HI-CI group but not in the HI-HA group. The nature of auditory input from these two devices can be different because hearing aids simply provide amplification to the sounds entering the ear in comparison to cochlear implants which bypass the damaged ear regions and directly stimulate the auditory nerve as described in section 1.5. As a result, both speech perception and production outcomes are relatively better in CI users when compared to HA users (see section 1.6.2 for details). This difference in hearing between HA and CI was seen in the current result of hearing threshold in noise and vowel identification performance. In addition to the importance concerning the type of devices, the usage duration can also be important as discussed in section 1.6.3. While some participants in the HI-HA group have used CI for a short period, they still did not show any adaptation behavior. This indicates that HI individuals need some experience with their devices to be able to learn (or relearn) how to process auditory information supplied by the CI. Depending upon their level of habituation with their device it could be reflected in their speech perception and production performance.

Sensory reliance can also affect speech motor adaptation using altered auditory feedback. The previous studies showed that not all participants showed an adaptation (Ohashi & Ito, 2019; Rochet-Capellan & Ostry, 2011). Lametti et al. (2012) demonstrated that, when both somatosensory and auditory feedback are modulated simultaneously during a speech production task, adaptation was induced based on their sensory preference: either or both auditory and somatosensory adaptation. This is what happens with HI-CI individuals in the current study where the rate of 66% of participants showed an adaptation. This overall relative consistency between NH and HI individuals suggests that hearing impairment does not impede sensorimotor adaptation, provided that audition recovery is sufficient. This is in phase with some other studies showing that HI subjects equipped with CI can indeed adapt to small acoustic variations in their environment, as shown in the study by Scarbel et al. (2017) in which HI-CI participants displayed phonetic convergence effects with formant modification of vowel production to follow the production of a speaking partner.

In summary, HI-HA participants have a reduced ability to precisely distinguish between the minute speech sound differences compared with HI-CI ones. A certain period of CI experience may be required to achieve this level of auditory performance. These findings highlight the importance of auditory feedback in controlling production mechanisms, and suggest that HI individuals may display behavior similar to NH individuals in terms of sensorimotor adaptation in speech motor control if they dispose of a sufficient level of auditory feedback.

## General Discussion

---

### 8.1 The somatosensory function in the perception-production link in NH and HI persons, in a developmental perspective

At the beginning of this work in section 2.5, we raised 4 questions that have been addressed in the four studies of this thesis. We will discuss the corresponding results, with a specific focus on the way auditory-somatosensory interactions for speech perception develop along with the development of speech abilities.

The first two questions aimed at better characterizing the role of the somatosensory function at the pivot of the perception-production relationships, in NH participants. Studies 1 and 2 do provide new interesting light on this question, confirming, as we expected at the beginning of this work, that the auditory-somatosensory interaction is actually bidirectional in respect to perception and production. Indeed, Ito et al. (2009) had shown in their pioneer study that somatosensory inputs intervene in speech perception. Study 1 interestingly shows that the size of the perceptual shift due to somatosensory inputs is related to the precision of production performance, concluding that the larger the somatosensory contrast between two vowels in production, the larger the somatosensory effect in the perception of the same contrast. This suggests a first developmental mechanism, in which the acuity of the acquisition process of speech production would favor the role of somatosensory inputs in speech perception. This is actually in line with the findings by Trudeau-Fisette et al. (2019) that the somatosensory effect in speech perception was larger in adult than in children, suggesting that altogether, auditory-somatosensory interaction in speech perception seems to develop along with the development of speech abilities.

The results of Study 2 also increase our knowledge about the bidirectionality of the perceptuo-motor link underlying the somatosensory function in speech processes. Indeed, we show for the first time that speech perceptual training using paired auditory-somatosensory inputs affects speech production. The precise way this effect of auditory-somatosensory perceptual training further modifies speech motor control is not completely clear yet, and we introduce in the discussion in section 5.4 a number of open questions in this paradigm. But the important point

remains, that the way auditory and somatosensory inputs are presented and combined may modify the perceptual representations underlying the motor control specification. Once again, this shows that learning processes are involved and able to constantly monitor the link between perception and production, and that the somatosensory system is involved in these learning processes.

On this basis, the next two questions concerned the way perceptuo-motor relationships could develop in HI participants, around two paradigms, that were the somatosensory effect in speech perception and sensorimotor adaptation. We had a special interest and concern with the way auditory experience and hearing abilities would intervene in these two studies. It appeared that these abilities played a crucial role in the results. Indeed, studies 3 and 4 showed that HI hearing ability to acoustically identify small differences in speech sounds strongly depended on the type of hearing device. While CI users correctly identified the vowel continuum in both studies, HA users had difficulty to identify the vowels correctly. This suggests that the quality of auditory inputs HI persons get from CI allows them to recover enough hearing ability to be able to perform auditory-related functions relatively well, especially because in both studies they need to be able to precisely recognise the minute differences in the presented speech sounds. As described in section 1.6.2, CI users acquire in consequence better precision in speech production.

Accordingly, it could be expected that CI users would show a larger somatosensory effect in speech perception. Actually, the results in study 3 are rather the opposite, since a somatosensory effect in speech perception was seen in some HA users and with a large amplitude, but not in CI users. A possible reason could be a difference in hearing experience, since most of the HA users in that study corrected their hearing since prelingual periods, while CI users received their devices during relatively later periods. This suggests that although the development of auditory-somatosensory interaction in speech perception can be related to the development of speech production mechanism, hearing experience from prelingual ages can be more important to acquire auditory-somatosensory integration in HI individuals. As we also suggested in section 6.4, another factor could play an important role in the lack of somatosensory effect in speech perception in CI users in spite of their good hearing ability, that is the amount of attention they probably put on auditory processing to ensure this good level of auditory performance. Interestingly, audiovisual interactions are classically considered as excellent in HI participants whatever their used hearing device. It could hence suggest that auditory-somatosensory

integration requires more attention than audiovisual integration. This also opens the route to a topic that we did not address in these studies, that is the role of attention and executive mechanisms in all these tasks, and also possibly the way the somatosensory system could facilitate certain perceptual tasks and decrease auditory effort, known to be extremely important in auditory handicap.

Finally, study 4 showed that when the hearing ability was good enough to detect small formant differences between vowels, HI individuals showed adaptation to altered auditory feedback training. This suggests that sensorimotor learning ability could be similar between NH and HI individuals. Connecting this result to the results of study 2 in NH participants, the question arises whether HI individuals could also show adaptive changes in speech production following perceptual learning using paired auditory-somatosensory stimulation as examined in NH individuals in study 2. If receiving paired auditory-somatosensory inputs during perceptual training as in speech production is a key for this adaptive change, HI individuals could display adaptive changes regardless of their hearing profile since this sensory experience can be received in all individuals who acquired speaking. However, if having an efficient auditory-somatosensory interaction mechanism is a key for such adaptation following auditory-somatosensory perceptual training, then adaptive changes could also depend on the HI participants' hearing profiles as in Study 3. Long-time HA (and possibly CI) users from prelingual ages may show the changes in the production following the perceptual training, but not the others. Further investigation is required.

Globally, the results about sensorimotor interactions and the role of the somatosensory system in speech perception in HI individuals confirm the crucial role of auditory experience in these processes, though direct evidence of the role of sensorimotor developmental processes in the two corresponding studies is less clear.

## **8.2 Future Works**

The current work has implications for the developmental aspects of auditory-somatosensory integration in speech perception. This can be further investigated in HI children of different age groups. Given the limited auditory-somatosensory interaction in our findings it may be interesting to investigate if the developmental trajectory of auditory-somatosensory integration is similar to that of audiovisual integration by testing participants in and around the critical ages mentioned previously (see section 2.3.1). It could be expected that the size of the somatosensory

effect is small in HI children similar to results shown in Trudeau-Fisette et al. (2019). But what mainly needs to be assessed in such a study is what other factors contribute to the development of auditory-somatosensory integration. It would be interesting to study what contributes to sensory experience that is needed to develop auditory-somatosensory integration. In our findings, we have already shown that device-related factors affect auditory-somatosensory integration. It may be worthwhile to investigate other factors such as the effect of communication strategies used in their day-to-day functioning. Could lipreading (visual speech processing) abilities intervene and interact with auditory-somatosensory integration? Could it be that better lipreaders show larger auditory-somatosensory interactions because they can integrate information from different senses efficiently or is it the other way round, in which since they already rely on visual inputs, they don't use somatosensory inputs as much? These questions could provide insights on how auditory-somatosensory integration may evolve at different stages in life.

Findings from the current work also have a potential for clinical applications such as speech therapy and rehabilitation procedures. Tactile stimulation has already been used for actual practical implications. The PROMPT "Prompts for Restructuring Oral Muscular Phonetic Targets" speech therapy is a tactile-kinaesthetic based method. The speech therapist manually provides tactile cues by gently moving the face and neck to guide patients to produce the correct articulatory movements for producing speech sounds. However, even though those stimulations are applied on the relevant articulator, the learner has to learn a mapping between receiving tactile cues and recovering intended sounds to be produced. In contrast, using articulatory-related somatosensory stimulation can be more efficient to instruct, to educate and to rehabilitate articulatory movement. As shown in the second study, speech production was changed by applying specific pairs of auditory-somatosensory stimulation. Finding an appropriate pair and direction of change can be used for acquiring or correcting articulatory movement in a proper manner.

The finding from the second study can also be useful for educational applications such as developing tools for foreign language learning using somatosensory-based production training. Articulatory based pronunciation training is a method that is used for teaching foreign languages where the teachers coach students based on movement of the articulators adequate to produce speech sounds. This has been investigated in terms of visual articulatory feedback to assist with pronunciation learning (Kühnert & Pillot-Loiseau, 2022; Suemitsu et al., 2015). Similarly,



orofacial somatosensory inputs could also help by providing articulatory-related information. Additionally, since the findings from the second study are relevant for practical applications, it could be further complemented by obtaining articulatory data such as the amount of movement of lips, tongue or jaw in response to perceptual training with somatosensory inputs.

### **8.3 Conclusions**

The current work expands the knowledge on the existing literature on the role of orofacial somatosensory inputs in speech perception, both by direct exploration on NH participants, and by investigating this in HI individuals. The presence of a degraded auditory system provides a unique situation for exploring how the auditory-somatosensory integration mechanism may develop in such cases. We found that the somatosensory function in speech perception was varied in HI individuals. Moreover, we were able to associate these variations to different hearing profiles giving rise to a possible criterion that can be considered in future investigations. The variation in auditory-somatosensory integration could be explained in relation to the development of hearing and production ability in HI individuals.

The findings shed further insights on the auditory-somatosensory integration mechanism based on sensorimotor experience related to speech acquisition. It also emphasises the role of orofacial somatosensory inputs in the perception-production link. The possible future implications can be practical applications for both clinical and educational purposes including an intervention with HI individuals.

# Appendix 1

## QUESTIONNAIRE POUR ÉVALUER LA DÉFICIENCE AUDITIVE

Titre de l'étude : Exploitation de la fonction somatosensorielle chez les sujets malentendants (HI) pour comprendre et réhabiliter la boucle perceptuo-motrice dans la communication vocale

Sujet n° :

Nom :

Age :

Date du test :

Année de naissance

:

Sexe :

Merci de remplir le questionnaire ci-dessous dans le cadre de l'expérience.

N'hésitez pas à ajouter autant de détails/explications que nécessaire à chacune des questions ci-dessous :

1. À quel âge votre perte auditive a-t-elle été détectée ?
2. Quel type de perte auditive avez-vous ? <ul style="list-style-type: none"><li>• Perte auditive neurosensorielle</li><li>• Perte auditive conductive</li><li>• Perte auditive mixte</li></ul>
3. Quelle est la cause de votre perte auditive ? <ul style="list-style-type: none"><li>• Perte d'audition due au bruit</li><li>• Perte auditive congénitale</li><li>• Perte auditive héréditaire</li><li>• Perte auditive liée à une maladie (due à une maladie de Ménière, à une méningite, etc.)</li><li>• En raison de médicaments</li><li>• En raison d'un traumatisme</li><li>• Perte auditive soudaine</li><li>• En rapport avec l'âge</li><li>• Autre, veuillez préciser</li></ul>
4. Quel est votre degré de perte auditive à l'oreille droite ? <ul style="list-style-type: none"><li>• Perte auditive légère</li><li>• Perte auditive modérée</li><li>• Perte auditive grave</li><li>• Perte auditive profonde</li><li>• Normal/Non perte d'audition</li></ul>

<p>5. Quel est votre degré de perte auditive à votre oreille gauche ?</p> <ul style="list-style-type: none"> <li>• Perte auditive légère</li> <li>• Perte auditive modérée</li> <li>• Perte auditive grave</li> <li>• Perte auditive profonde</li> <li>• Normal/Non perte d'audition</li> </ul>
<p>6. À quel âge avez-vous été équipé d'implants cochléaires/aides auditives ?</p> <ul style="list-style-type: none"> <li>• Oreille gauche (préciser le cas échéant implant ou aide) :</li> <li>• Oreille droite (préciser le cas échéant implant ou aide) :</li> </ul>
<p>7. À quelle fréquence utilisez-vous vos appareils auditifs/implants cochléaires ?</p>
<p>8. Quel type/modèle d'implant cochléaire/aide auditive utilisez-vous ?</p> <ul style="list-style-type: none"> <li>• Oreille gauche :</li> <li>• Oreille droite :</li> </ul>
<p>9. Quand avez-vous subi une évaluation auditive pour la dernière fois ?</p>
<p>10. Entendez-vous souvent dans vos oreilles des bruits (tintement aigu, bousculade, etc.) qui durent plus de 5 minutes (acouphènes) ?</p>

# References

- Aarhus, L., Tambs, K., Kvestad, E., & Engdahl, B. (2015). Childhood Otitis Media: A Cohort Study with 30-year Follow-Up of Hearing (The HUNT study). *Ear and Hearing, 36*(3), 302–308. <https://doi.org/10.1097/AUD.0000000000000118>
- Alais, D., & Burr, D. (2004). The Ventriloquist Effect Results from Near-Optimal Bimodal Integration. *Current Biology, 14*(3), 257–262. <https://doi.org/10.1016/j.cub.2004.01.029>
- Alsius, A., Navarra, J., Campbell, R., & Soto-Faraco, S. (2005). Audiovisual integration of speech falters under high attention demands. *Current Biology: CB, 15*(9), 839–843. <https://doi.org/10.1016/j.cub.2005.03.046>
- Alsius, A., Navarra, J., & Soto-Faraco, S. (2007). Attention to touch weakens audiovisual speech integration. *Experimental Brain Research, 183*(3), 399–404. <https://doi.org/10.1007/s00221-007-1110-1>
- Ashokumar, M., Guichet, C., Schwartz, J.-L., & Ito, T. (2023). Correlation between the Effect of Orofacial Somatosensory Inputs in Speech Perception and Speech Production Performance. *Auditory Perception & Cognition, 0*(0), 1–11. <https://doi.org/10.1080/25742442.2022.2134674>
- Bastian, H. C. (1887). The “muscular sense”; its nature and cortical localisation1. *Brain, 10*(1), 1–89. <https://doi.org/10.1093/brain/10.1.1>
- Baudonck, N., Dhooge, I., D’haeseleer, E., & Van Lierde, K. (2010). A comparison of the consonant production between Dutch children using cochlear implants and children using hearing aids. *International Journal of Pediatric Otorhinolaryngology, 74*(4), 416–421. <https://doi.org/10.1016/j.ijporl.2010.01.017>
- Bellis, T. J., & Bellis, J. D. (2015). Chapter 30—Central auditory processing disorders in children and adults. In M. J. Aminoff, F. Boller, & D. F. Swaab (Eds.), *Handbook of Clinical Neurology* (Vol. 129, pp. 537–556). Elsevier. <https://doi.org/10.1016/B978-0-444-62630-1.00030-5>
- Benoît C., Mohamadi, T., & Kandel, S. (1994). Effects of Phonetic Context on Audio-Visual Intelligibility of French. *Journal of Speech, Language, and Hearing Research, 37*(5), 1195–1203. <https://doi.org/10.1044/jshr.3705.1195>
- Berens, P. (2009). CircStat: A MATLAB Toolbox for Circular Statistics. *Journal of Statistical Software, 31*, 1–21. <https://doi.org/10.18637/jss.v031.i10>
- Bergeson, T. R., Pisoni, D. B., & Davis, R. A. O. (2005). Development of Audiovisual Comprehension Skills in Prelingually Deaf Children With Cochlear Implants. *Ear and Hearing, 26*(2), 149–164.
- Bernhard, N., Gauger, U., Romo Ventura, E., Uecker, F. C., Olze, H., Knopke, S., Hänsel, T., & Coordes, A. (2021). Duration of deafness impacts auditory performance after cochlear implantation: A meta-analysis. *Laryngoscope Investigative Otolaryngology, 6*(2), 291–301. <https://doi.org/10.1002/lio2.528>
- Bisogno, A., Scarpa, A., Di Girolamo, S., De Luca, P., Cassandro, C., Viola, P., Ricciardiello, F., Greco, A., Vincentiis, M. D., Ralli, M., & Di Stadio, A. (2021). Hearing Loss and Cognitive Impairment: Epidemiology, Common Pathophysiological Findings, and Treatment Considerations. *Life, 11*(10), 1102. <https://doi.org/10.3390/life11101102>

- Bittencourt, A. G., Torre, A. A. G. D., Bento, R. F., Tsuji, R. K., & Brito, R. de. (2012). Prelingual deafness: Benefits from cochlear implants versus conventional hearing aids. *International Archives of Otorhinolaryngology*, *16*(3), 387–390. <https://doi.org/10.7162/S1809-97772012000300014>
- Blamey, P., Arndt, P., Bergeron, F., Bredberg, G., Brimacombe, J., Facer, G., Larky, J., Lindström, B., Nedzelski, J., Peterson, A., Shipp, D., Staller, S., & Whitford, L. (1996). Factors Affecting Auditory Performance of Postlinguistically Deaf Adults Using Cochlear Implants. *Audiology and Neurotology*, *1*(5), 293–306. <https://doi.org/10.1159/000259212>
- Blamey, P., Artieres, F., Başkent, D., Bergeron, F., Beynon, A., Burke, E., Dillier, N., Dowell, R., Fraysse, B., Gallégo, S., Govaerts, P. J., Green, K., Huber, A. M., Kleine-Punte, A., Maat, B., Marx, M., Mawman, D., Mosnier, I., O'Connor, A. F., ... Lazard, D. S. (2013). Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants: An update with 2251 patients. *Audiology & Neuro-Otology*, *18*(1), 36–47. <https://doi.org/10.1159/000343189>
- Blamey, P., Barry, J., Bow, C., Sarant, J., Paatsch, L., & Wales, R. (2001). The development of speech production following cochlear implantation. *Clinical Linguistics & Phonetics*, *15*(5), 363–382. <https://doi.org/10.1080/02699200010017823>
- Boersma, P., & Weenink, D. (2001). PRAAT, a system for doing phonetics by computer. *Glott International*, *5*, 341–345.
- Bouchard, M.-E., Ouellet, C., & Cohen, H. (2009). Speech Development in Prelingually Deaf Children with Cochlear Implants: Speech Production in Implanted Prelingually Deaf Children. *Language and Linguistics Compass*, *3*(1), 1–18. <https://doi.org/10.1111/j.1749-818X.2008.00079.x>
- Bouton, S., Serniclaes, W., Bertoncini, J., & Colé, P. (2012). Perception of Speech Features by French-Speaking Children With Cochlear Implants. *Journal of Speech, Language, and Hearing Research*, *55*(1), 139–153. [https://doi.org/10.1044/1092-4388\(2011/10-0330\)](https://doi.org/10.1044/1092-4388(2011/10-0330))
- Bowl, M. R., & Dawson, S. J. (2019). Age-Related Hearing Loss. *Cold Spring Harbor Perspectives in Medicine*, *9*(8). <https://doi.org/10.1101/cshperspect.a033217>
- Boyd, I. A., & Roberts, T. D. M. (1953). Proprioceptive discharges from stretch-receptors in the knee-joint of the cat. *The Journal of Physiology*, *122*(1), 38-58.1.
- Bruderer, A. G., Danielson, D. K., Kandhadai, P., & Werker, J. F. (2015). Sensorimotor influences on speech perception in infancy. *Proceedings of the National Academy of Sciences*, *112*(44), 13531–13536. <https://doi.org/10.1073/pnas.1508631112>
- Cardin, V. (2016). Effects of Aging and Adult-Onset Hearing Loss on Cortical Auditory Regions. *Frontiers in Neuroscience*, *10*. <https://www.frontiersin.org/articles/10.3389/fnins.2016.00199>
- Cardon, G., Campbell, J., & Sharma, A. (2012). Plasticity in the developing auditory cortex: Evidence from children with sensorineural hearing loss and auditory neuropathy spectrum disorder. *Journal of the American Academy of Audiology*, *23*(6), 396–411; quiz 495. <https://doi.org/10.3766/jaaa.23.6.3>
- Caudrelier, T., & Rochet-Capellan, A. (2019). Changes in speech production in response to formant perturbations: An overview of two decades of research,. In S. Fuchs, J. Cleland, & A. Rochet-Capellan (Eds.), *Speech production and perception: Learning and*

- memory*. (Vol. 6, pp. 15–75). Peter Lang. <https://hal.science/hal-02426327>
- Chao, S.-C., Ochoa, D., & Daliri, A. (2019). Production Variability and Categorical Perception of Vowels Are Strongly Linked. *Frontiers in Human Neuroscience*, *13*, 96. <https://doi.org/10.3389/fnhum.2019.00096>
- Chen, J. X., Lindeborg, M., Herman, S. D., Ishai, R., Knoll, R. M., Remenschneider, A., Jung, D. H., & Kozin, E. D. (2018). Systematic review of hearing loss after traumatic brain injury without associated temporal bone fracture. *American Journal of Otolaryngology*, *39*(3), 338–344. <https://doi.org/10.1016/j.amjoto.2018.01.018>
- Cho, Y., Kim, D. H., Choi, J., Lee, J. K., Roh, Y.-K., Nam, H.-Y., Nam, G.-E., Kim, D.-W., Lee, S.-H., Lee, C.-W., Han, K., & Park, Y.-G. (2016). Glomerular Filtration Rate and Urine Albumin to Creatinine Ratio Associated With Hearing Impairment Among Korean Adults With Diabetes. *Medicine*, *95*(17), e3423. <https://doi.org/10.1097/MD.00000000000003423>
- Choi, S. H., & Choi, C.-H. (2015). Noise-Induced Neural Degeneration and Therapeutic Effect of Antioxidant Drugs. *Journal of Audiology and Otology*, *19*(3), 111–119. <https://doi.org/10.7874/jao.2015.19.3.111>
- Clark, F. J., Horch, K. W., Bach, S. M., & Larson, G. F. (1979). Contributions of cutaneous and joint receptors to static knee-position sense in man. *Journal of Neurophysiology*, *42*(3), 877–888. <https://doi.org/10.1152/jn.1979.42.3.877>
- Cohen, B. E., Durstenfeld, A., & Roehm, P. C. (2014). Viral Causes of Hearing Loss: A Review for Hearing Health Professionals. *Trends in Hearing*, *18*, 2331216514541361. <https://doi.org/10.1177/2331216514541361>
- Connor, C. M., Craig, H. K., Raudenbush, S. W., Heavner, K., & Zwolan, T. A. (2006). The Age at Which Young Deaf Children Receive Cochlear Implants and Their Vocabulary and Speech-Production Growth: Is There an Added Value for Early Implantation? *Ear and Hearing*, *27*(6), 628. <https://doi.org/10.1097/01.aud.0000240640.59205.42>
- Cooper, W. E., & Lauritsen, M. R. (1974). Feature processing in the perception and production of speech. *Nature*, *252*(5479), Article 5479. <https://doi.org/10.1038/252121a0>
- Cooper, W. E., & Nager, R. M. (1975). Perceptuo-motor adaptation to speech: An analysis of bisyllabic utterances and a neural model. *The Journal of the Acoustical Society of America*, *58*(1), 256–265. <https://doi.org/10.1121/1.380655>
- Cowan, R. S. C., Blamey, P. J., Galvin, K. L., Sarant, J. Z., Alcántara, J. I., & Clark, G. M. (1990). Perception of sentences, words, and speech features by profoundly hearing-impaired children using a multichannel electrotactile speech processor. *The Journal of the Acoustical Society of America*, *88*(3), 1374–1384. <https://doi.org/10.1121/1.399715>
- Cristobal, R., & Oghalai, J. S. (2008). Hearing loss in children with very low birth weight: Current review of epidemiology and pathophysiology. *Archives of Disease in Childhood. Fetal and Neonatal Edition*, *93*(6), F462–F468. <https://doi.org/10.1136/adc.2007.124214>
- Daliri, A., Wieland, E. A., Cai, S., Guenther, F. H., & Chang, S.-E. (2018). Auditory-motor adaptation is reduced in adults who stutter but not in children who stutter. *Developmental Science*, *21*(2). <https://doi.org/10.1111/desc.12521>
- Dalton, D. S., Cruickshanks, K. J., Klein, B. E. K., Klein, R., Wiley, T. L., & Nondahl, D. M. (2003). The Impact of Hearing Loss on Quality of Life in Older Adults. *The*

- Gerontologist*, 43(5), 661–668. <https://doi.org/10.1093/geront/43.5.661>
- Daniel, E. (2007). Noise and Hearing Loss: A Review. *Journal of School Health*, 77(5), 225–231. <https://doi.org/10.1111/j.1746-1561.2007.00197.x>
- DePaolis, R. A., Vihman, M. M., & Keren-Portnoy, T. (2011). Do production patterns influence the processing of speech in prelinguistic infants? *Infant Behavior and Development*, 34(4), 590–601. <https://doi.org/10.1016/j.infbeh.2011.06.005>
- Desai, S., Stickney, G., & Zeng, F.-G. (2008). Auditory-visual speech perception in normal-hearing and cochlear-implant listeners. *The Journal of the Acoustical Society of America*, 123(1), 428–440. <https://doi.org/10.1121/1.2816573>
- di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, 91(1), 176–180. <https://doi.org/10.1007/BF00230027>
- Duck, S. W., Prazma, J., Bennett, P. S., & Pillsbury, H. C. (1997). Interaction Between Hypertension and Diabetes Mellitus in the Pathogenesis of Sensorineural Hearing Loss. *The Laryngoscope*, 107(12), 1596–1605. <https://doi.org/10.1097/00005537-199712000-00004>
- Duquesnoy, A. J., & Plomp, R. (1983). The effect of a hearing aid on the speech-reception threshold of hearing-impaired listeners in quiet and in noise. *The Journal of the Acoustical Society of America*, 73(6), 2166–2173. <https://doi.org/10.1121/1.389540>
- Dwyer-Hemmings, L., Manjaly, J. G., Nash, R., Mukherjee, A., & Lavy, J. A. (2019). Stapes Surgery for Profound Hearing Loss Secondary to Otosclerosis. *Ear, Nose & Throat Journal*, 98(5), 273–278. <https://doi.org/10.1177/0145561319834540>
- Edin, B. B. (2001). Cutaneous afferents provide information about knee joint movements in humans. *The Journal of Physiology*, 531(1), 289–297. <https://doi.org/10.1111/j.1469-7793.2001.0289j.x>
- Edin, B. B., & Abbs, J. H. (1991). Finger movement responses of cutaneous mechanoreceptors in the dorsal skin of the human hand. *Journal of Neurophysiology*, 65(3), 657–670. <https://doi.org/10.1152/jn.1991.65.3.657>
- Edin, B. B., & Johansson, N. (1995). Skin strain patterns provide kinaesthetic information to the human central nervous system. *The Journal of Physiology*, 487(Pt 1), 243–251.
- Emmett, S. D., & Keith P West, J. (2015). Nutrition and hearing loss: A neglected cause and global health burden. *The American Journal of Clinical Nutrition*, 102(5), 987. <https://doi.org/10.3945/ajcn.115.122598>
- Erber, N. P. (1969). Interaction of audition and vision in the recognition of oral speech stimuli. *Journal of Speech and Hearing Research*, 12(2), 423–425. <https://doi.org/10.1044/jshr.1202.423>
- Erber, N. P. (1971). Auditory and audiovisual reception of words in low-frequency noise by children with normal hearing and by children with impaired hearing. *Journal of Speech and Hearing Research*, 14(3), 496–512. <https://doi.org/10.1044/jshr.1403.496>
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870), 429–433. <https://doi.org/10.1038/415429a>
- Estivill, X., Fortina, P., Surrey, S., Rabionet, R., Melchionda, S., D'Agruma, L., Mansfield, E., Rappaport, E., Govea, N., Milà, M., Zelante, L., & Gasparini, P. (1998). Connexin-26

- mutations in sporadic and inherited sensorineural deafness. *The Lancet*, 351(9100), 394–398. [https://doi.org/10.1016/S0140-6736\(97\)11124-2](https://doi.org/10.1016/S0140-6736(97)11124-2)
- Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience*, 4.
- Finley, C. C., Holden, T. A., Holden, L. K., Whiting, B. R., Chole, R. A., Neely, G. J., Hullar, T. E., & Skinner, M. W. (2008). Role of Electrode Placement as a Contributor to Variability in Cochlear Implant Outcomes. *Otology & Neurotology*, 29(7), 920. <https://doi.org/10.1097/MAO.0b013e318184f492>
- Fletcher, M. D., Hadeedi, A., Goehring, T., & Mills, S. R. (2019). Electro-haptic enhancement of speech-in-noise performance in cochlear implant users. *Scientific Reports*, 9(1), 11428. <https://doi.org/10.1038/s41598-019-47718-z>
- Fletcher, M. D., Mills, S. R., & Goehring, T. (2018). Vibro-Tactile Enhancement of Speech Intelligibility in Multi-talker Noise for Simulated Cochlear Implant Listening. *Trends in Hearing*, 22, 233121651879783. <https://doi.org/10.1177/2331216518797838>
- Foch, C., Araujo, M., Weckel, A., Damase-Michel, C., Montastruc, J.-L., Benevent, J., Durrieu, G., & Lacroix, I. (2018). In utero drug exposure and hearing impairment in 2-year-old children A case-control study using the EFEMERIS database. *International Journal of Pediatric Otorhinolaryngology*, 113, 192–197. <https://doi.org/10.1016/j.ijporl.2018.07.035>
- Folkins, J. W., & Larson, C. R. (1978). In search of a tonic vibration reflex in the human lip. *Brain Research*, 151(2), 409.
- Fort, M., Spinelli, E., Savariaux, C., & Kandel, S. (2010). The word superiority effect in audiovisual speech perception. *Speech Communication*, 52, 525–532. <https://doi.org/10.1016/j.specom.2010.02.005>
- Fowler, C. A. (1986). An event approach to the study of speech perception from a direct–realist perspective. *Journal of Phonetics*, 14(1), 3–28. [https://doi.org/10.1016/S0095-4470\(19\)30607-2](https://doi.org/10.1016/S0095-4470(19)30607-2)
- Fowler, C. A., & Dekle, D. J. (1991). *Listening With Eye and Hand: Cross-Modal Contributions to Speech Perception*.
- Franken, M. K., Acheson, D. J., McQueen, J. M., Eisner, F., & Hagoort, P. (2017). Individual variability as a window on production–perception interactions in speech motor control. *The Journal of the Acoustical Society of America*, 142(4), 2007–2018. <https://doi.org/10.1121/1.5006899>
- Garramiola-Bilbao, I., & Rodríguez-Álvarez, A. (2016). Linking hearing impairment, employment and education. *Public Health*, 141, 130–135. <https://doi.org/10.1016/j.puhe.2016.09.013>
- Geers, A., Brenner, C., & Davidson, L. (2003). Factors Associated with Development of Speech Perception Skills in Children Implanted by Age Five. *Ear and Hearing*, 24(1), 24S. <https://doi.org/10.1097/01.AUD.0000051687.99218.0F>
- Ghosh, S. S., Matthies, M. L., Maas, E., Hanson, A., Tiede, M., Ménard, L., Guenther, F. H., Lane, H., & Perkell, J. S. (2010). An investigation of the relation between sibilant production and somatosensory and auditory acuity. *The Journal of the Acoustical Society of America*, 128(5), 3079–3087. <https://doi.org/10.1121/1.3493430>



- Gick, B., & Derrick, D. (2009). Aero-tactile integration in speech perception. *Nature*, 462(7272), 502–504. <https://doi.org/10.1038/nature08572>
- Goodwin, G. M., McCloskey, D. I., & Matthews, P. B. C. (1972). The contribution of muscle afferents to kinaesthesia shown by vibration induced illusions of movement and by the effects of paralysing joint afferents. *Brain*, 95(4), 705–748. <https://doi.org/10.1093/brain/95.4.705>
- Gordon-Salant, S. (2005). Hearing loss and aging: New research findings and clinical implications. *The Journal of Rehabilitation Research and Development*, 42(4s), 9. <https://doi.org/10.1682/JRRD.2005.01.0006>
- Gori, M. (2015). Multisensory Integration and Calibration in Children and Adults with and without Sensory and Motor Disabilities. *Multisensory Research*, 28(1–2), 71–99. <https://doi.org/10.1163/22134808-00002478>
- Gorman, B. M. (2018). *A Framework for Speechreading Acquisition Tools* (arXiv:1806.00812). arXiv. <http://arxiv.org/abs/1806.00812>
- Graham, A. S., Ben-Azu, B., Tremblay, M.-È., Torre, P., Senekal, M., Laughton, B., van der Kouwe, A., Jankiewicz, M., Kaba, M., & Holmes, M. J. (2023). A review of the auditory-gut-brain axis. *Frontiers in Neuroscience*, 17. <https://www.frontiersin.org/articles/10.3389/fnins.2023.1183694>
- Grandon, B., & Vilain, A. (2020). Development of fricative production in French-speaking school-aged children using cochlear implants and children with normal hearing. *Journal of Communication Disorders*, 86, 105996. <https://doi.org/10.1016/j.jcomdis.2020.105996>
- Grant, K. W., & Seitz, P. F. (2000). The use of visible speech cues for improving auditory detection of spoken sentences. *The Journal of the Acoustical Society of America*, 108(3 Pt 1), 1197–1208. <https://doi.org/10.1121/1.1288668>
- Grant, K. W., Walden, B. E., & Seitz, P. F. (1998). Auditory-visual speech recognition by hearing-impaired subjects: Consonant recognition, sentence recognition, and auditory-visual integration. *The Journal of the Acoustical Society of America*, 103(5 Pt 1), 2677–2690. <https://doi.org/10.1121/1.422788>
- Grosse, S. D., Ross, D. S., & Dollard, S. C. (2008). Congenital cytomegalovirus (CMV) infection as a cause of permanent bilateral hearing loss: A quantitative assessment. *Journal of Clinical Virology*, 41(2), 57–62. <https://doi.org/10.1016/j.jcv.2007.09.004>
- Gupta, S., Singh, S. kumar, & Dubey, P. K. (2012). Cochlear Implant Using Neural Prosthetics. *International Journal of Advancements in Research & Technology*, 1, 266–272.
- Hagbarth, K. E., Hellsing, G., & Löfstedt, L. (1976). TVR and vibration-induced timing of motor impulses in the human jaw elevator muscles. *Journal of Neurology, Neurosurgery, and Psychiatry*, 39(8), 719–728.
- Haile, L. M., Kamenov, K., Briant, P. S., Orji, A. U., Steinmetz, J. D., Abdoli, A., Abdollahi, M., Abu-Gharbieh, E., Afshin, A., Ahmed, H., Rashid, T. A., Akalu, Y., Alahdab, F., Alanezi, F. M., Alanzi, T. M., Hamad, H. A., Ali, L., Alipour, V., Al-Raddadi, R. M., ... Chadha, S. (2021). Hearing loss prevalence and years lived with disability, 1990–2019: Findings from the Global Burden of Disease Study 2019. *The Lancet*, 397(10278), 996–1009. [https://doi.org/10.1016/S0140-6736\(21\)00516-X](https://doi.org/10.1016/S0140-6736(21)00516-X)
- Håkansson, B., Reinfeldt, S., Persson, A.-C., Jansson, K.-J. F., Rigato, C., Hultcrantz, M., &

- Eeg-Olofsson, M. (2019). The bone conduction implant – a review and 1-year follow-up. *International Journal of Audiology*, 58(12), 945–955. <https://doi.org/10.1080/14992027.2019.1657243>
- Hall, S. J. (2006). The development of a new english sentence in noise test and an english number recognition test.
- Hellsing, G. (1977). A tonic vibration reflex evoked in the jaw opening muscles in man. *Archives of Oral Biology*, 22(3), 175–180. [https://doi.org/10.1016/0003-9969\(77\)90151-0](https://doi.org/10.1016/0003-9969(77)90151-0)
- Hellsing, G., & Klineberg, I. (1982). Reflex responses to taps in various directions from the human digastric muscle. *Archives of Oral Biology*, 27(11), 909–913. [https://doi.org/10.1016/0003-9969\(82\)90096-6](https://doi.org/10.1016/0003-9969(82)90096-6)
- Hermann, R., Lescanne, E., Loundon, N., Barone, P., Belmin, J., Blanchet, C., Borel, S., Charpiot, A., Coez, A., Deguine, O., Farinetti, A., Godey, B., Lazard, D., Marx, M., Mosnier, I., Nguyen, Y., Teissier, N., Virole, M. B., Roman, S., & Truy, E. (2019). French Society of ENT (SFORL) guidelines. Indications for cochlear implantation in adults. *European Annals of Otorhinolaryngology, Head and Neck Diseases*, 136(3), 193–197. <https://doi.org/10.1016/j.anorl.2019.04.006>
- Horga, D., & Liker, M. (2006). Voice and pronunciation of cochlear implant speakers. *Clinical Linguistics & Phonetics*, 20(2–3), 211–217. <https://doi.org/10.1080/02699200400027015>
- Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal. Biometrische Zeitschrift*, 50(3), 346–363. <https://doi.org/10.1002/bimj.200810425>
- Houde, J. F., & Jordan, M. I. (1998). Sensorimotor Adaptation in Speech Production. *Science*, 279(5354), 1213–1216. <https://doi.org/10.1126/science.279.5354.1213>
- Howard, J. (2019). *Temporomandibular Joint Disorders, Facial Pain, and Dental Problems in Performing Artists* (pp. 151–196).
- Ito, T., & Gomi, H. (2007). Cutaneous mechanoreceptors contribute to the generation of a cortical reflex in speech. *Neuroreport*, 18(9), 907–910. <https://doi.org/10.1097/WNR.0b013e32810f2dfb>
- Ito, T., Gracco, V. L., & Ostry, D. J. (2014). Temporal factors affecting somatosensory auditory interactions in speech processing. *Frontiers in Psychology*, 5. <https://doi.org/10.3389/fpsyg.2014.01198>
- Ito, T., & Ogane, R. (2022). Repetitive Exposure to Orofacial Somatosensory Inputs in Speech Perceptual Training Modulates Vowel Categorization in Speech Perception. *Frontiers in Psychology*, 13, 839087. <https://doi.org/10.3389/fpsyg.2022.839087>
- Ito, T., & Ostry, D. J. (2010). Somatosensory Contribution to Motor Learning Due to Facial Skin Deformation. *Journal of Neurophysiology*, 104(3), 1230–1238. <https://doi.org/10.1152/jn.00199.2010>
- Ito, T., Tiede, M., & Ostry, D. J. (2009). Somatosensory function in speech perception. *Proceedings of the National Academy of Sciences*, 106(4), 1245–1248. <https://doi.org/10.1073/pnas.0810063106>
- Jamieson, D. G., & Cheesman, M. F. (1987). The adaptation of produced voice-onset time. *Journal of Phonetics*, 15(1), 15–27. [https://doi.org/10.1016/S0095-4470\(19\)30534-0](https://doi.org/10.1016/S0095-4470(19)30534-0)

- Jansen, S., Luts, H., Dejonckere, P., van Wieringen, A., & Wouters, J. (2013). Efficient Hearing Screening in Noise-Exposed Listeners Using the Digit Triplet Test. *Ear & Hearing, 34*(6), 773–778. <https://doi.org/10.1097/AUD.0b013e318297920b>
- Jansen, S., Luts, H., Wagener, K. C., Frachet, B., & Wouters, J. (2010). The French digit triplet test: A hearing screening tool for speech intelligibility in noise. *International Journal of Audiology, 49*(5), 378–387. <https://doi.org/10.3109/14992020903431272>
- Johansson, R. S., Trulsson, M., Olsson, K. Å., & Abbs, J. H. (1988). Mechanoreceptive afferent activity in the infraorbital nerve in man during speech and chewing movements. *Experimental Brain Research, 72*(1), 209–214. <https://doi.org/10.1007/BF00248519>
- Kaganovich, N., & Schumaker, J. (2014). Audiovisual integration for speech during mid-childhood: Electrophysiological evidence. *Brain and Language, 139*, 36–48. <https://doi.org/10.1016/j.bandl.2014.09.011>
- Karltorp, E., Eklöf, M., Östlund, E., Asp, F., Tideholm, B., & Löfkvist, U. (2020). Cochlear implants before 9 months of age led to more natural spoken language development without increased surgical risks. *Acta Paediatrica, 109*(2), 332–341. <https://doi.org/10.1111/apa.14954>
- Kileny, P. R., Zwolan, T. A., & Ashbaugh, C. (2001). The Influence of Age at Implantation on Performance With a Cochlear Implant in Children. *Otology & Neurotology, 22*(1), 42.
- Kim, H. H., & Barrs, D. M. (2006). Hearing Aids: A Review of What's New. *Otolaryngology–Head and Neck Surgery, 134*(6), 1043–1050. <https://doi.org/10.1016/j.otohns.2006.03.010>
- Kral, A., & Eggermont, J. J. (2007). What's to lose and what's to learn: Development under auditory deprivation, cochlear implants and limits of cortical plasticity. *Brain Research Reviews, 56*(1), 259–269. <https://doi.org/10.1016/j.brainresrev.2007.07.021>
- Kral, A., & Sharma, A. (2012). Developmental neuroplasticity after cochlear implantation. *Trends in Neurosciences, 35*(2), 111–122. <https://doi.org/10.1016/j.tins.2011.09.004>
- Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience, 5*(11), 831–843. <https://doi.org/10.1038/nrn1533>
- Kuhl, P. K., Ramírez, R. R., Bosseler, A., Lin, J.-F. L., & Imada, T. (2014). Infants' brain responses to speech suggest Analysis by Synthesis. *Proceedings of the National Academy of Sciences, 111*(31), 11238–11245. <https://doi.org/10.1073/pnas.1410963111>
- Kühnert, B., & Pillot-Loiseau, C. (2022). Teaching pronunciation with direct visual articulatory feedback: Pedagogical considerations for the use of ultrasound in the classroom. *Recherches Anglaises et Nord Américaines, 55*, 9–24.
- Lachs, L., Pisoni, D. B., & Kirk, K. I. (2001). Use of Audiovisual Information in Speech Perception by Prelingually Deaf Children with Cochlear Implants: A First Report. *Ear and Hearing, 22*(3), 236–251.
- Lalonde, K., & Holt, R. F. (2016). Audiovisual speech perception development at varying levels of perceptual processing. *The Journal of the Acoustical Society of America, 139*(4), 1713–1723. <https://doi.org/10.1121/1.4945590>
- Lametti, D. R., Krol, S. A., Shiller, D. M., & Ostry, D. J. (2014). Brief Periods of Auditory Perceptual Training Can Determine the Sensory Targets of Speech Motor Learning. *Psychological Science, 25*(7), 1325–1336. <https://doi.org/10.1177/0956797614529978>
- Lametti, D. R., Nasir, S. M., & Ostry, D. J. (2012). Sensory Preference in Speech Production

- Revealed by Simultaneous Alteration of Auditory and Somatosensory Feedback. *Journal of Neuroscience*, 32(27), 9351–9358. <https://doi.org/10.1523/JNEUROSCI.0404-12.2012>
- Langereis, M. C., Bosman, A. J., van Olphen, A. F., & Smoorenburg, G. F. (1997). Changes in vowel quality in post-lingually deafened cochlear implant users. *Audiology: Official Organ of the International Society of Audiology*, 36(5), 279–297. <https://doi.org/10.3109/00206099709071980>
- Lassaletta, L., Sánchez-Cuadrado, I., Espinosa, J. M., Batuecas, Á., Cenjor, C., Lavilla, M. J., Cavallé, L., Huarte, A., Nuñez, F., Manrique, M., Ramos, Á., de Paula, C., & Gil-Carcedo, E. (2019). Active Middle Ear Implants. *Acta Otorrinolaringologica (English Edition)*, 70(2), 112–118. <https://doi.org/10.1016/j.otoeng.2019.03.006>
- Lazard, D. S., Vincent, C., Venail, F., Heyning, P. V. de, Truy, E., Sterkers, O., Skarzynski, P. H., Skarzynski, H., Schauwers, K., O’Leary, S., Mawman, D., Maat, B., Kleine-Punte, A., Huber, A. M., Green, K., Govaerts, P. J., Fraysse, B., Dowell, R., Dillier, N., ... Blamey, P. J. (2012). Pre-, Per- and Postoperative Factors Affecting Performance of Postlinguistically Deaf Adults Using Cochlear Implants: A New Conceptual Model over Time. *PLOS ONE*, 7(11), e48739. <https://doi.org/10.1371/journal.pone.0048739>
- Leite, J. N., Silva, V. S., & Buzo, B. C. (2016). Emissões otoacústicas em recém-nascidos com hipóxia perinatal leve e moderada. *CoDAS*, 28(2), 93–98. <https://doi.org/10.1590/2317-1782/20162015086>
- Lennartsson, B. (1979). Muscle spindles in the human anterior digastric muscle. *Acta Odontologica Scandinavica*, 37(6), 329–333. <https://doi.org/10.3109/00016357909004704>
- Leybaert, J., & LaSasso, C. J. (2010). Cued Speech for Enhancing Speech Perception and First Language Development of Children With Cochlear Implants. *Trends in Amplification*, 14(2), 96–112. <https://doi.org/10.1177/1084713810375567>
- Li, C.-M., Zhang, X., Hoffman, H. J., Cotch, M. F., Themann, C. L., & Wilson, M. R. (2014). Hearing Impairment Associated With Depression in US Adults, National Health and Nutrition Examination Survey 2005-2010. *JAMA Otolaryngology–Head & Neck Surgery*, 140(4), 293–302. <https://doi.org/10.1001/jamaoto.2014.42>
- Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74(6), 431–461. <https://doi.org/10.1037/h0020279>
- Liberman, A. M., Delattre, P. C., Cooper, F. S., & Gerstman, L. J. (1954). The role of consonant-vowel transitions in the perception of the stop and nasal consonants. *Psychological Monographs: General and Applied*, 68(8), 1–13. <https://doi.org/10.1037/h0093673>
- Liberman, A. M., Delattre, P., & Cooper, F. S. (1952). The Rôle of Selected Stimulus-Variables in the Perception of the Unvoiced Stop Consonants. *The American Journal of Psychology*, 65(4), 497–516. <https://doi.org/10.2307/1418032>
- Lin, B. M., Wang, M., Stankovic, K. M., Eavey, R., McKenna, M. J., Curhan, G. C., & Curhan, S. G. (2020). Cigarette Smoking, Smoking Cessation and Risk of Hearing Loss in Women. *The American Journal of Medicine*, 133(10), 1180–1186. <https://doi.org/10.1016/j.amjmed.2020.03.049>

- Lindblom, B. E., & Sundberg, J. E. (1971). Acoustical consequences of lip, tongue, jaw, and larynx movement. *The Journal of the Acoustical Society of America*, *50*(4), 1166–1179. <https://doi.org/10.1121/1.1912750>
- Maeda, S. (1990). Compensatory articulation during speech: Evidence from the analysis and synthesis of vocal tract shapes using an articulatory model. *W.J. Hardcastle and A. Marchal (Eds.), Speech production and speech modelling* (pp. 131-149). Boston: Kluwer Academic Publishers.
- Maillot, O., Attyé, A., Boyer, E., Heck, O., Kastler, A., Grand, S., Schmerber, S., & Krainik, A. (2016). Post traumatic deafness: A pictorial review of CT and MRI findings. *Insights into Imaging*, *7*(3), 341–350. <https://doi.org/10.1007/s13244-016-0490-9>
- Matkin, N. D., & Wilcox, A. M. (1999). Considerations in the education of children with hearing loss. *Pediatric Clinics of North America*, *46*(1), 143–152. [https://doi.org/10.1016/S0031-3955\(05\)70087-0](https://doi.org/10.1016/S0031-3955(05)70087-0)
- Mayberry, R. (2010). Early Language Acquisition and Adult Language Ability: What Sign Language Reveals About the Critical Period for Language. In *The Oxford Handbook of Deaf Studies, Language, and Education* (Vol. 2, pp. 281–291). <https://doi.org/10.1093/oxfordhb/9780195390032.013.0019>
- McGurk, H., & Macdonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*(5588), 746–748. <https://doi.org/10.1038/264746a0>
- Ménard, L., & Boe, L. (2004). *L'émergence du système phonologique chez l'enfant: L'apport de la modélisation articulatoire*. <https://doi.org/10.1353/CJL.2005.0003>
- Ménard, L., Schwartz, J.-L., & Aubin, J. (2008). Invariance and variability in the production of the height feature in French vowels. *Speech Communication*, *50*(1), 14–28. <https://doi.org/10.1016/j.specom.2007.06.004>
- Michaudet, C., & Malaty, J. (2018). Cerumen Impaction: Diagnosis and Management. *American Family Physician*, *98*(8), 525–529.
- Mick, P., Kawachi, I., & Lin, F. R. (2014). The Association between Hearing Loss and Social Isolation in Older Adults. *Otolaryngology–Head and Neck Surgery*, *150*(3), 378–384. <https://doi.org/10.1177/0194599813518021>
- Mildner, V., Šindija, B., & Horga, D. (2003). Evaluation of Croatian-speaking children's speech with cochlear implants and comparison with hearing-aid users. *Journal of Multilingual Communication Disorders*, *1*(1), 63–70. <https://doi.org/10.1080/1476967021000042119>
- Mildner, V., Sindija, B., & Zrinski, K. V. (2006). Speech perception of children with cochlear implants and children with traditional hearing aids. *Clinical Linguistics & Phonetics*, *20*(2–3), 219–229. <https://doi.org/10.1080/02699200400027031>
- Mollaei, F., Shiller, D. M., & Gracco, V. L. (2013). Sensorimotor adaptation of speech in Parkinson's disease. *Movement Disorders: Official Journal of the Movement Disorder Society*, *28*(12), 1668–1674. <https://doi.org/10.1002/mds.25588>
- Monasta, L., Ronfani, L., Marchetti, F., Montico, M., Brumatti, L. V., Bavcar, A., Grasso, D., Barbiero, C., & Tamburlini, G. (2012). Burden of Disease Caused by Otitis Media: Systematic Review and Global Estimates. *PLOS ONE*, *7*(4), e36226. <https://doi.org/10.1371/journal.pone.0036226>
- Murray, J. J., Hall, W. C., & Snoddon, K. (2019). Education and health of children with hearing

- loss: The necessity of signed languages. *Bulletin of the World Health Organization*, 97(10), 711–716. <https://doi.org/10.2471/BLT.19.229427>
- Musacchia, G., Arum, L., Nicol, T., Garstecki, D., & Kraus, N. (2009). Audiovisual Deficits in Older Adults with Hearing Loss: Biological Evidence. *Ear and Hearing*, 30(5), 505. <https://doi.org/10.1097/AUD.0b013e3181a7f5b7>
- Nasir, S. M., & Ostry, D. J. (2009). Auditory plasticity and speech motor learning. *Proceedings of the National Academy of Sciences*, 106(48), 20470–20475. <https://doi.org/10.1073/pnas.0907032106>
- Neilson, P. D., Andrews, G., Guitar, B. E., & Quinn, P. T. (1979). Tonic stretch reflexes in lip, tongue and jaw muscles. *Brain Research*, 178(2), 311–327. [https://doi.org/10.1016/0006-8993\(79\)90696-6](https://doi.org/10.1016/0006-8993(79)90696-6)
- Nikolopoulos, T. P., O’Donoghue, G. M., & Archbold, S. (1999). Age at Implantation: Its Importance in Pediatric Cochlear Implantation. *The Laryngoscope*, 109(4), 595–599. <https://doi.org/10.1097/00005537-199904000-00014>
- Nordin, M., & Thomander, L. (1989). Intrafascicular multi-unit recordings from the human infra-orbital nerve. *Acta Physiologica Scandinavica*, 135(2), 139–148. <https://doi.org/10.1111/j.1748-1716.1989.tb08561.x>
- Nordvik, Ø., Laugen Heggdal, P. O., Brännström, J., Vassbotn, F., Aarstad, A. K., & Aarstad, H. J. (2018). Generic quality of life in persons with hearing loss: A systematic literature review. *BMC Ear, Nose, and Throat Disorders*, 18, 1. <https://doi.org/10.1186/s12901-018-0051-6>
- O’Donoghue, G. M., Nikolopoulos, T. P., Archbold, S. M., & Tait, M. (1999). Cochlear implants in young children: The relationship between speech perception and speech intelligibility. *Ear and Hearing*, 20(5), 419–425. <https://doi.org/10.1097/00003446-199910000-00005>
- Ogane, R., Schwartz, J.-L., & Ito, T. (2020). Orofacial somatosensory inputs modulate word segmentation in lexical decision. *Cognition*, 197, 104163. <https://doi.org/10.1016/j.cognition.2019.104163>
- Ohashi, H., & Ito, T. (2019). Recalibration of auditory perception of speech due to orofacial somatosensory inputs during speech motor adaptation. *Journal of Neurophysiology*, 122(5), 2076–2084. <https://doi.org/10.1152/jn.00028.2019>
- Omstead, K. M., Williams, J., Weinberg, S. M., Marazita, M. L., & Burrows, A. M. (2023). Mammalian facial muscles contain muscle spindles. *The Anatomical Record*, 306(10), 2562–2571. <https://doi.org/10.1002/ar.25172>
- Ozimek, E., Kutzner, D., Sęk, A., & Wicher, A. (2009). Development and evaluation of Polish digit triplet test for auditory screening. *Speech Communication*, 51(4), 307–316. <https://doi.org/10.1016/j.specom.2008.09.007>
- Patri, J.-F., Perrier, P., Schwartz, J.-L., & Diard, J. (2018). What drives the perceptual change resulting from speech motor adaptation? Evaluation of hypotheses in a Bayesian modeling framework. *PLOS Computational Biology*, 14(1), e1005942. <https://doi.org/10.1371/journal.pcbi.1005942>
- Perkell, J. S., Guenther, F. H., Lane, H., Matthies, M. L., Stockmann, E., Tiede, M., & Zandipour, M. (2004). The distinctness of speakers’ productions of vowel contrasts is related to their discrimination of the contrasts. *The Journal of the Acoustical Society of*

- America*, 116(4), 2338–2344. <https://doi.org/10.1121/1.1787524>
- Peterson, D. C., Reddy, V., Launico, M. V., & Hamel, R. N. (2023). Neuroanatomy, Auditory Pathway. In *StatPearls*. StatPearls Publishing. <http://www.ncbi.nlm.nih.gov/books/NBK532311/>
- Pisoni, D. B. (2000). Cognitive Factors and Cochlear Implants: Some Thoughts on Perception, Learning, and Memory in Speech Perception. *Ear and Hearing*, 21(1), 70–78.
- Plomp, R. (1978). Auditory handicap of hearing impairment and the limited benefit of hearing aids. *The Journal of the Acoustical Society of America*, 63(2), 533–549. <https://doi.org/10.1121/1.381753>
- Potgieter, J.-M., Swanepoel, D. W., & Smits, C. (2018). Evaluating a smartphone digits-in-noise test as part of the audiometric test battery. *The South African Journal of Communication Disorders*, 65(1), 574. <https://doi.org/10.4102/sajcd.v65i1.574>
- Pulvermüller, F., Huss, M., Kherif, F., Martin, F. M. del P., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences*, 103(20), 7865–7870. <https://doi.org/10.1073/pnas.0509989103>
- Puschmann, S., Daeglau, M., Stropahl, M., Mirkovic, B., Rosemann, S., Thiel, C. M., & Debener, S. (2019). Hearing-impaired listeners show increased audiovisual benefit when listening to speech in noise. *NeuroImage*, 196, 261–268. <https://doi.org/10.1016/j.neuroimage.2019.04.017>
- Ramakrishnan, K., Sparks, R. A., & Berryhill, W. E. (2007). Diagnosis and Treatment of Otitis Media. *American Family Physician*, 76(11), 1650–1658.
- Reed, C. M., Rabinowitz, W. M., Durlach, N. I., Braida, L. D., Conway-Fithian, S., & Schultz, M. C. (1984). *Research on the Tadoma Method of Speech Communication*.
- Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G., & Matelli, M. (1988). Functional organization of inferior area 6 in the macaque monkey. *Experimental Brain Research*, 71(3), 491–507. <https://doi.org/10.1007/BF00248742>
- Rochet-Capellan, A., & Ostry, D. J. (2011). Simultaneous Acquisition of Multiple Auditory–Motor Transformations in Speech. *The Journal of Neuroscience*, 31(7), 2657–2662. <https://doi.org/10.1523/jneurosci.6020-10.2011>
- Rosemann, S., & Thiel, C. M. (2018). Audio-visual speech processing in age-related hearing loss: Stronger integration and increased frontal lobe recruitment. *NeuroImage*, 175, 425–437. <https://doi.org/10.1016/j.neuroimage.2018.04.023>
- Ross, L. A., Molholm, S., Blanco, D., Gomez-Ramirez, M., Saint-Amour, D., & Foxe, J. J. (2011). The development of multisensory speech perception continues into the late childhood years: Development of audiovisual speech perception. *European Journal of Neuroscience*, 33(12), 2329–2337. <https://doi.org/10.1111/j.1460-9568.2011.07685.x>
- Rouger, J., Fraysse, B., Deguine, O., & Barone, P. (2008). McGurk effects in cochlear-implanted deaf subjects. *Brain Research*, 1188, 87–99. <https://doi.org/10.1016/j.brainres.2007.10.049>
- Rouger, J., Lagleyre, S., Fraysse, B., Deneve, S., Deguine, O., & Barone, P. (2007). Evidence that cochlear-implanted deaf patients are better multisensory integrators. *Proceedings of the National Academy of Sciences*, 104(17), 7295–7300. <https://doi.org/10.1073/pnas.0609419104>
- Scarbel, L., Beautemps, D., Schwartz, J. L., & Sato, M. (2017). 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*, *101*, 39-46.
- Schorr, E. A., Fox, N. A., van Wassenhove, V., & Knudsen, E. I. (2005). Auditory-visual fusion in speech perception in children with cochlear implants. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(51), 18748–18750. <https://doi.org/10.1073/pnas.0508862102>
- Schulte, A., Marozeau, J., Ruhe, A., Büchner, A., Kral, A., & Innes-Brown, H. (2023). Improved speech intelligibility in the presence of congruent vibrotactile speech input. *Scientific Reports*, *13*(1), Article 1. <https://doi.org/10.1038/s41598-023-48893-w>
- Schwartz, J.-L., Basirat, A., Ménard, L., & Sato, M. (2012). The Perception-for-Action-Control Theory (PACT): A perceptuo-motor theory of speech perception. *Journal of Neurolinguistics*, *25*(5), 336–354. <https://doi.org/10.1016/j.jneuroling.2009.12.004>
- Schwartz, J.-L., Beutemps, D., Abry, C., & Escudier, P. (1993). Inter-individual and cross-linguistic strategies for the production of the [i] vs. [Y] contrast. *Journal of Phonetics*, *21*(4), 411–425.
- Schwartz, S. R., Magit, A. E., Rosenfeld, R. M., Ballachanda, B. B., Hackell, J. M., Krouse, H. J., Lawlor, C. M., Lin, K., Parham, K., Stutz, D. R., Walsh, S., Woodson, E. A., Yanagisawa, K., & Cunningham Jr, E. R. (2017). Clinical Practice Guideline (Update): Earwax (Cerumen Impaction). *Otolaryngology–Head and Neck Surgery*, *156*(S1), S1–S29. <https://doi.org/10.1177/0194599816671491>
- Sekiyama, K., Soshi, T., & Sakamoto, S. (2014). Enhanced audiovisual integration with aging in speech perception: A heightened McGurk effect in older adults. *Frontiers in Psychology*, *5*. <https://doi.org/10.3389/fpsyg.2014.00323>
- Shiller, D. M., Sato, M., Gracco, V. L., & Baum, S. R. (2009). Perceptual recalibration of speech sounds following speech motor learning. *The Journal of the Acoustical Society of America*, *125*(2), 1103–1113. <https://doi.org/10.1121/1.3058638>
- Skipper, J. I., Nusbaum, H. C., & Small, S. L. (2005). Listening to talking faces: Motor cortical activation during speech perception. *NeuroImage*, *25*(1), 76–89. <https://doi.org/10.1016/j.neuroimage.2004.11.006>
- Smits, C., Kapteyn, T. S., & Houtgast, T. (2004). Development and validation of an automatic speech-in-noise screening test by telephone. *International Journal of Audiology*, *43*(1), 15–28. <https://doi.org/10.1080/14992020400050004>
- Sprinzel, G. M., & Riechelmann, H. (2010). Current Trends in Treating Hearing Loss in Elderly People: A Review of the Technology and Treatment Options – A Mini-Review. *Gerontology*, *56*(3), 351–358. <https://doi.org/10.1159/000275062>
- Stål, P., Eriksson, P. O., Eriksson, A., & Thornell, L. E. (1987). Enzyme-histochemical differences in fibre-type between the human major and minor zygomatic and the first dorsal interosseus muscles. *Archives of Oral Biology*, *32*(11), 833–841. [https://doi.org/10.1016/0003-9969\(87\)90011-2](https://doi.org/10.1016/0003-9969(87)90011-2)
- Stål, P., Eriksson, P.-O., Eriksson, A., & Thornell, L.-E. (1990). Enzyme-histochemical and morphological characteristics of muscle fibre types in the human buccinator and orbicularis oris. *Archives of Oral Biology*, *35*(6), 449–458. [https://doi.org/10.1016/0003-9969\(90\)90208-R](https://doi.org/10.1016/0003-9969(90)90208-R)



- Strelnikov, K., Rouger, J., Demonet, J.-F., Lagleyre, S., Fraysse, B., Deguine, O., & Barone, P. (2010). Does Brain Activity at Rest Reflect Adaptive Strategies? Evidence from Speech Processing after Cochlear Implantation. *Cerebral Cortex*, 20(5), 1217–1222. <https://doi.org/10.1093/cercor/bhp183>
- Stucky, S. R., Wolf, K. E., & Kuo, T. (2010). The Economic Effect of Age-Related Hearing Loss: National, State, and Local Estimates, 2002 and 2030. *Journal of the American Geriatrics Society*, 58(3), 618–619. <https://doi.org/10.1111/j.1532-5415.2010.02746.x>
- Suemitsu, A., Dang, J., Ito, T., & Tiede, M. (2015). A real-time articulatory visual feedback approach with target presentation for second language pronunciation learning. *The Journal of the Acoustical Society of America*, 138(4), EL382–EL387. <https://doi.org/10.1121/1.4931827>
- Sumbly, W. H., & Pollack, I. (1954). Visual Contribution to Speech Intelligibility in Noise. *The Journal of the Acoustical Society of America*, 26(2), 212–215. <https://doi.org/10.1121/1.1907309>
- Taitelbaum-Swead, R., Avivi, M., Gueta, B., & Fostick, L. (2019). The effect of delayed auditory feedback (DAF) and frequency altered feedback (FAF) on speech production: Cochlear implanted versus normal hearing individuals. *Clinical Linguistics & Phonetics*, 33(7), 628–640. <https://doi.org/10.1080/02699206.2019.1574313>
- Tremblay, S., Shiller, D. M., & Ostry, D. J. (2003). Somatosensory basis of speech production. *Nature*, 423(6942), 866–869. <https://doi.org/10.1038/nature01710>
- Trigeminal nerve. (2024). In *Wikipedia*. [https://en.wikipedia.org/w/index.php?title=Trigeminal\\_nerve&oldid=1193385510](https://en.wikipedia.org/w/index.php?title=Trigeminal_nerve&oldid=1193385510)
- Trudeau-Fisette, P., Ito, T., & Ménard, L. (2019). Auditory and Somatosensory Interaction in Speech Perception in Children and Adults. *Frontiers in Human Neuroscience*, 13, 344. <https://doi.org/10.3389/fnhum.2019.00344>
- Trudeau-Fisette, P., Tiede, M., & Ménard, L. (2017). Compensations to auditory feedback perturbations in congenitally blind and sighted speakers: Acoustic and articulatory data. *PLOS ONE*, 12(7), e0180300. <https://doi.org/10.1371/journal.pone.0180300>
- Trulsson, M., & Johansson, R. S. (2002). Orofacial mechanoreceptors in humans: Encoding characteristics and responses during natural orofacial behaviors. *Behavioural Brain Research*, 135(1), 27–33. [https://doi.org/10.1016/S0166-4328\(02\)00151-1](https://doi.org/10.1016/S0166-4328(02)00151-1)
- Tye-Murray, N., Sommers, M. S., & Spehar, B. (2007). Audiovisual Integration and Lipreading Abilities of Older Adults with Normal and Impaired Hearing. *Ear and Hearing*, 28(5), 656. <https://doi.org/10.1097/AUD.0b013e31812f7185>
- Tye-Murray, N., Spencer, L., & Woodworth, G. G. (1995). Acquisition of Speech by Children Who Have Prolonged Cochlear Implant Experience. *Journal of Speech, Language, and Hearing Research*, 38(2), 327–337. <https://doi.org/10.1044/jshr.3802.327>
- van der Straaten, T. F. K., Briaire, J. J., Vickers, D., Boermans, P. P. B. M., & Frijns, J. H. M. (2021). Selection Criteria for Cochlear Implantation in the United Kingdom and Flanders: Toward a Less Restrictive Standard. *Ear and Hearing*, 42(1), 68. <https://doi.org/10.1097/AUD.0000000000000901>
- Vilain, A., Dole, M., Løevenbruck, H., Pascalis, O., & Schwartz, J.-L. (2019). The role of production abilities in the perception of consonant category in infants. *Developmental Science*, 22(6), e12830. <https://doi.org/10.1111/desc.12830>

- Villacorta, V. M., Perkell, J. S., & Guenther, F. H. (2007). Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception. *The Journal of the Acoustical Society of America*, *122*(4), 2306–2319. <https://doi.org/10.1121/1.2773966>
- Waltzman, S. B., & Cohen, N. L. (1998). Cochlear implantation in children younger than 2 years old. *The American Journal of Otolaryngology*, *19*(2), 158–162.
- Watkins, K. E., Strafella, A. P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia*, *41*(8), 989–994. [https://doi.org/10.1016/S0028-3932\(02\)00316-0](https://doi.org/10.1016/S0028-3932(02)00316-0)
- Weber, P. C., & Klein, A. J. (1999). Hearing loss. *Medical Clinics of North America*, *83*(1), 125–137. [https://doi.org/10.1016/S0025-7125\(05\)70092-4](https://doi.org/10.1016/S0025-7125(05)70092-4)
- Weisenberger, J. M., & Kozma-Spytek, L. (1991). Evaluating tactile aids for speech perception and production by hearing-impaired adults and children. *The American Journal of Otolaryngology*, *12 Suppl*, 188–200.
- Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, *7*(7), 701–702. <https://doi.org/10.1038/nn1263>
- Wong, J. D., Kistemaker, D. A., Chin, A., & Gribble, P. L. (2012). Can proprioceptive training improve motor learning? *Journal of Neurophysiology*, *108*(12), 3313–3321. <https://doi.org/10.1152/jn.00122.2012>
- World Report on Hearing. (2021). <https://www.who.int/teams/noncommunicable-diseases/sensory-functions-disability-and-rehabilitation/highlighting-priorities-for-ear-and-hearing-care>
- Wu, P., O'Malley, J. T., Gruttola, V. de, & Liberman, M. C. (2020). Age-Related Hearing Loss Is Dominated by Damage to Inner Ear Sensory Cells, Not the Cellular Battery That Powers Them. *Journal of Neuroscience*, *40*(33), 6357–6366. <https://doi.org/10.1523/JNEUROSCI.0937-20.2020>
- Zeeshan, F., Bari, A., Dugal, M. N., & Saeed, F. (2018). Hearing impairment after acute bacterial meningitis in children. *Pakistan Journal of Medical Sciences*, *34*(3), 655–659. <https://doi.org/10.12669/pjms.343.14373>
- Zokoll, M. A., Wagener, K. C., Brand, T., Buschermöhle, M., & Kollmeier, B. (2012). Internationally comparable screening tests for listening in noise in several European languages: The German digit triplet test as an optimization prototype. *International Journal of Audiology*, *51*(9), 697–707. <https://doi.org/10.3109/14992027.2012.690078>