



HAL
open science

Neural Changes and Cognitive Processes Underlying Novel Skill Acquisition

Snežana Todorović

► **To cite this version:**

Snežana Todorović. Neural Changes and Cognitive Processes Underlying Novel Skill Acquisition: Evidence from Motor Learning and Speech Production. Neuroscience. Aix-Marseille Université, France, 2023. English. NNT: . tel-04593424

HAL Id: tel-04593424

<https://hal.science/tel-04593424>

Submitted on 30 May 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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0
International License

THÈSE DE DOCTORAT

Soutenue à Aix-Marseille Université
le 11 juillet 2023 par

Snežana TODOROVIĆ

Neural Changes and Cognitive Processes
Underlying Novel Skill Acquisition:
Evidence from Motor Learning and Speech Production

Discipline

Sciences du langage

École doctorale

Cognition, Langage, Éducation (ED356)

Laboratoire/Partenaires de recherche

Laboratoire Parole et Langage
Centre IRM-INT @ Cerimed

Composition du jury

Hélène LOEVENBRUCK

Directrice de Recherche, CNRS
Université de Grenoble

Rapporteure

Robert HARTSUIKER

Professor
University of Ghent

Rapporteur

François-Xavier ALARIO

Directeur de Recherche, CNRS
Aix-Marseille Université

Président du jury

Cheryl FRENCK-MESTRE

Directrice de Recherche, CNRS
Aix-Marseille Université

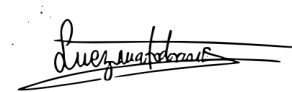
Directrice de thèse

Affidavit

I, undersigned, Snežana Todorović, hereby declare that the work presented in this manuscript is my own work, carried out under the scientific supervision of Dr. Cheryl Frenck-Mestre, following the original supervision of Dr. Elin Runnqvist and Dr. Sonja Kotz, in accordance with the principles of honesty, integrity and responsibility inherent to the research mission. The research work and the writing of this manuscript have been carried out in compliance with both the french national charter for Research Integrity and the Aix-Marseille University charter on the fight against plagiarism.

This work has not been submitted previously either in this country or in another country in the same or in a similar version to any other examination body.

Aix-en-Provence, July 11th 2023



This work is licensed under [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Public License](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Liste de publications et participation aux conférences

Liste des publications et/ou brevets réalisées dans le cadre du projet de thèse:

1. Todorović, S., Anton, J.L., Sein, N., Nazarian, B., Chanoine, V., Rauchbauer, B., Kotz, S. A. and Runnqvist, E. (2023). Cortico-cerebellar monitoring of speech sequence production. *Neurobiology of Language*, 1-21.

Participation aux conférences et écoles d'été au cours de la période de thèse:

Conférences:

1. Todorović, S., Anton, J.L., Sein, N., Nazarian, B., Kotz, S. A. and Runnqvist, E. *Learning to Pronounce Novel Sound Sequences Induces Changes in Resting State Functional Connectivity between Right Posterior Cerebellum and Left Posterior Parietal Cortex*. Présentation orale à: The International Max Planck Research School Conference for Language Sciences at the Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands, juin 2022.
2. Todorović, S., Chanoine, V., Brovelli, A., Badier, J.M., Kotz, S. A. and Runnqvist, E. *The Role of Cortico-Subcortical Loops in the Learning of Speech Motor Sequences*. Poster présenté à Eleventh Annual Meeting of the Society for the Neurobiology of Language, Helsinki, Finland, août 2019.
3. Todorović, S., Kotz, S. A. and Runnqvist, E. *Learning of Speech Motor Sequences: the Role of the Basal Ganglia, Cerebellum and Medial Frontal Cortex*. Poster présenté au Colloque des Jeunes Chercheurs en Sciences Cognitives, École Normale Supérieure, Paris, France, mars 2019.

Écoles d'été :

1. 5th Summer School of the Institute for Language, Communication and the Brain. Centre international des rencontres mathématiques, Marseille, France, August 29th–September 2nd, 2022.
2. 4th Summer School of the Institute for Language, Communication and the Brain. Centre international des rencontres mathématiques, Marseille, France, August 30th–September 3rd, 2021.
3. Learning the CONN Toolbox for Functional Connectivity Analysis, Martinos Center for Biomedical Imaging, Charlestown, Massachusetts, The United States. April 16th-May 14th, 2021 (online).
4. 3rd Summer School of the Institute for Language, Communication and the Brain. Centre international des rencontres mathématiques, Marseille, France, August 31st–September 4th, 2020.
5. 6th Barcelona Summer School on Bilingualism and Multilingualism, Universitat Pompeu Fabra, Barcelona, Spain, September 17th-20th, 2019.
6. 1st Summer School of the Institute for Language, Communication and the Brain, Marseille, France, September 3-7th, 2018.

Résumé

Compétence essentielle en communication, à l'interface entre l'apprentissage moteur et la production de la parole, la prononciation de nouvelles séquences phonétiques est aussi une forme de comportement adaptatif. À travers une série de méta-analyses, nous avons examiné les processus cognitifs impliqués dans l'apprentissage de séquences motrices et leurs corrélats neuronaux en fonction de la tâche et des propriétés de la séquence en cours d'apprentissage. Similaire au réseau cérébral que l'on retrouve dans la production de la parole, ce réseau variait en fonction des processus cognitifs pouvant être détectés à travers différentes conceptions expérimentales, permettant d'affiner certains modèles neurocognitifs existants sur l'apprentissage moteur. En nous focalisant davantage sur deux processus cognitifs sous-jacents à l'acquisition de compétences, le suivi de performance et le raffinement du modèle interne, nous avons réalisé une expérience en imagerie par résonance magnétique fonctionnelle, dans laquelle les participants ont appris à prononcer de nouvelles séquences vocales de différents niveaux de difficulté de prononciation en fonction de leur légalité phonotactique. Deux mécanismes de suivi de performance ont été observés : un sensorimoteur dans les régions motrices du cervelet et le cortex frontal médial, et un transmodal dans les régions cognitives du cervelet, notamment dans le crus I, et des zones temporo-pariétales. Finalement, l'amélioration de la motricité, soutenue par les ganglions de la base, et le raffinement du modèle interne, soutenu par le cervelet, ont été examinés à travers les changements de connectivité fonctionnelle entre ces régions et le reste du cerveau. La connectivité des deux régions avec les zones corticales frontales et médio-frontales a changé après l'amélioration motrice induite par la répétition de nouvelles syllabes. Les changements liés à l'apprentissage, qui reflèteraient la mise à jour du modèle interne du cervelet, ont entraîné une augmentation de la connectivité du crus I droit avec le précuneus gauche dans le cortex pariétal. Ce changement de connectivité pourrait s'expliquer par la capacité croissante du cervelet à prédire l'état futur de l'activité corticale pariétale, augmentant leur synchronisation. L'activité pariétale pourrait alors soutenir l'imagerie motrice précédant l'articulation du mouvement, pour améliorer davantage la prédiction calculée dans le cervelet. L'ensemble des travaux rapportés dans cette thèse représente une étape importante vers la compréhension de la production de nouvelles séquences de la parole en tant que compétence basée sur des processus cognitifs partagés avec d'autres comportements adaptatifs. Combinant les approches traditionnelles avec la connectivité fonctionnelle à l'état de repos, nous avons observé ces processus non seulement dans des zones cérébrales distinctes, mais nous avons également pu identifier leur couplage dans des réseaux fonctionnels.

Mots clés : acquisition de compétences, apprentissage moteur, production de la parole, suivi de la performance, modèles direct et inverse, imagerie par résonance magnétique fonctionnelle

Abstract

At the boundary of motor learning and speech production, pronouncing novel speech sequences is an essential skill for communication. A form of adaptive behavior, it relies both on domain-general and task-specific cognitive mechanisms. Using a series of activation likelihood estimation coordinate-based meta-analyses, we examined cognitive processes involved in non-speech motor sequence learning and their neural correlates. These processes, observed as a function of task as well as of the properties of the sequences being learned, revealed a global network similar to the one found in speech production. This network varied considerably between individual tasks and baselines, which indicates that the choice of the experimental design can favor isolating certain cognitive processes and related brain regions over others. As a consequence, we were able to propose a refinement of existing neurocognitive models. We went onto focus on two cognitive processes at the core of skill acquisition, performance monitoring and internal model refinement, and performed a functional magnetic resonance imaging study in which participants learned to pronounce novel speech sequences of different difficulties due to their phonotactic legality. Sensorimotor-driven, representation-specific performance monitoring was observed in motor regions of the cerebellum as well as in the medial frontal cortex. Transmodal monitoring was observed in cognitive regions of the cerebellum, and the crus I in particular, in concert with temporo-parietal areas. These results highlight the presence of multiple - and possibly hierarchically interdependent - mechanisms at the service of optimizing performance during speech production. Finally, motor improvement, sustained by the basal ganglia, and internal model refinement, sustained by the cerebellum, were examined through the changes in resting state functional connectivity between these regions and the rest of the brain. The connectivity of both the cerebellum and the basal ganglia with frontal and medial frontal cortical areas changed with participants' repetition-induced motor behavior improvement. Changes related to learning progress, arguably reflecting internal model updating, resulted in an increase in the connectivity between the right crus I in the cerebellum and the left precuneus in the parietal cortex. This change in connectivity could be explained by the cerebellum's increasing capacity to predict the future state of the parietal cortical activity and synchronize with it. The parietal activity could then sustain the motor imagery that precedes movement articulation, further improving cerebellar prediction. Drawing from both motor learning and speech production literatures, the body of the work reported in this dissertation represents a significant step towards understanding novel speech sequence production as a skill relying on cognitive processes shared with other adaptive behaviors. Combining traditional approaches with resting state functional connectivity, we were able not only to observe these processes in distinct cerebral areas but also to pinpoint their coupling in functional networks.

Keywords: skill acquisition, motor learning, speech production, performance monitoring, internal models, functional magnetic resonance imaging

À Sneža en 2010

Bravo pour cette ligne de
pour votre parcours de vie
exceptionnel,

Bia accident,

Bonne nuit

“A baby that speaks in its mother’s
womb should be able to give birth by
itself”

Contents

Affidavit	i
Liste de publications et participation aux conférences	ii
Résumé	iv
Abstract	v
Dedication	vi
Epigraph	vii
Contents	viii
List of Figures	xii
List of Tables	xiv
Abbreviation list	xv
1. Introduction	2
1.1. Cognitive processes underlying motor skill learning	2
1.1.1. Internal modeling	3
1.1.2. Performance monitoring	5
1.2. Motor sequence learning	7
1.2.1. Serial reaction time task	10
1.2.2. A neurobiological model of motor sequence learning	11
1.3. Speech sequence learning	11
1.3.1. Neurocognitive models related to syllable learning	13
1.4. Principles of functional magnetic resonance imaging	20
1.4.1. Functional connectivity	21
1.4.2. Coordinate based meta-analysis	21
1.5. The objectives of this dissertation	22

2. Domain-General Cognitive Subcomponents of Motor Sequence Learning: a Coordinate-Based Meta-Analysis	25
2.1. Introduction	26
2.2. Materials and methods	28
2.2.1. Rationale of the meta-analytic approach	28
2.2.2. Literature search and study selection	29
2.2.3. Activation likelihood estimation	34
2.3. Results	35
2.4. Discussion	36
2.4.1. Hippocampus	40
2.4.2. Thalamus	40
2.4.3. Motor cortical regions	41
2.4.4. Frontal regions	41
2.4.5. Parietal cortex	42
2.4.6. Basal ganglia	43
2.4.7. Cerebellum	43
2.5. Conclusion	44
3. Motor and Cognitive Aspects of Performance Monitoring During Novel Speech Sequence Production: an fMRI Study	47
3.1. Introduction	48
3.2. Methods	54
3.2.1. Participants	54
3.2.2. Materials	54
3.2.3. Procedure	55
3.2.4. MRI data acquisition	55
3.2.5. Behavioral data analyses	56
3.2.6. Neuroimaging data preprocessing	59
3.2.7. Whole brain analysis	61
3.2.8. Region of interest analysis	61
3.3. Results	63
3.3.1. Behavioral results	63
3.3.2. Neuroimaging results	65
3.4. Discussion	65
3.4.1. Transmodal and sensorimotor-driven monitoring in distinct cerebellar regions	70

3.4.2. Sensorimotor-driven feedback monitoring through the medial frontal cortex	71
3.4.3. Transmodal predictive monitoring through temporo-parietal cortices	72
3.5. Conclusion	73
4. Motor Tuning and Internal Model Refinement During Novel Skill Acquisition: a Resting-state Functional Connectivity Analysis	75
4.1. Introduction	76
4.1.1. Cerebellum and internal modeling	76
4.1.2. Basal ganglia and motor learning	76
4.1.3. Objectives	77
4.2. Methods	77
4.2.1. Participants	77
4.2.2. Experimental design and behavioral data analysis	77
4.2.3. Imaging data acquisition, anatomical and functional data pre-processing	78
4.2.4. Functional connectivity analyses	78
4.3. Results	80
4.3.1. Behavioral results	80
4.3.2. Neuroimaging results	80
4.4. Discussion	82
4.4.1. Syllable production accuracy improvement as motor tuning	84
4.4.2. Syllable learning as an internally modeled mental activity	85
4.5. Conclusion	86
5. General discussion	87
5.1. Overview	87
5.2. Main results	88
5.3. Speech motor sequence learning as skill acquisition at the interface of language and movement	92
5.4. Limitations	94
5.5. Conclusion	95
5.6. Perspectives	95
References	97

Notes	124
APPENDICES	126
A. Motor sequence learning and motor adaptation in table-top games . .	126
B. Foci and converging clusters for individual meta-analyses	127
C. Learning-related functional connectivity changes between the right crus I and the left precuneus	133

List of Figures

1.1. A neurobiological model of motor sequence learning (Doyon et al., 2018)	12
1.2. The Hierarchical State Feedback Control (HSFC) model of speech production (Hickok, 2012)	15
1.3. The Directions into Velocities of Articulators (DIVA) model of speech motor control (Guenther, Ghosh & Tourville, 2006)	17
1.4. A subcortico-cortical framework of speech production (Kotz & Schwartz, 2010)	19
2.1. The PRISMA flowchart	31
2.2. The results of the five conducted meta-analyses	39
2.3. Tentative specification of the cognitive processes in the neurobiological model of brain plasticity associated with motor sequence learning developed by Doyon and colleagues	46
3.1. The monitoring hypotheses	52
3.2. Depiction of an experimental trial	55
3.3. Experimental procedure	57
3.4. Functional data preprocessing flowchart	62
3.5. Average production accuracy over five experimental runs	64
3.6. Results of the ROI analysis	66
3.7. Results of the whole brain analysis	68
3.8. Task-related fMRI analysis summary	69
4.1. Seed ROI used in seed-to-voxel connectivity analysis	79
4.2. Task-related functional connectivity changes	81
4.3. Learning-related functional connectivity changes	83
1. Motor sequence learning and motor adaptation in table-top games	126
2. Foci and converging clusters for the meta-analysis of all studies regardless of the task and the baseline	128

3.	Foci and converging clusters for the meta-analysis focusing on the serial reaction time task	129
4.	Foci and converging clusters for the meta-analysis focusing on studies that accounted for motor execution in the baseline	130
5.	Foci and converging clusters for the meta-analysis focusing on studies that used known or previously practiced sequences as the baseline	131
6.	Foci and converging clusters for the meta-analysis focusing on studies that used random sequences as the baseline	132
7.	Learning-related functional connectivity changes between the right crus I and the left precuneus	133

List of Tables

2.1. Studies included in the meta-analyses	34
2.2. Results of global and task-specific meta-analyses	37
2.3. Results of baseline-specific meta-analyses	38
3.1. MNI coordinates and references of the ROI classified by anatomical regions	63
3.2. Results of the whole-brain analyses of the external and internal monitoring contrasts	67
4.1. Task-induced changes in the resting state connectivity	82

Abbreviation list

ACC	anterior cingulate cortex
aCompCor	anatomical CompCor
AFNI	Analysis of Functional NeuroImages software
ALE	activation likelihood estimation
AMU	Aix-Marseille University
ANTs	Advanced Normalization Tools software
AP	additional practice
AS	associative striatum
aSMG	anterior supramarginal gyrus
BA	Brodmann area
Bi	bilateral
BOLD	blood-oxygen-level dependent
Calc	calcarine fissure
CB	cerebellum
CC	cerebellar cortex
Cerimed	Centre Européen de Recherche en Imagerie Médicale (European Center for Research in Medical Imaging)
CI	confidence interval
CN	cerebellar nuclei
CNRS	Centre national de la recherche scientifique (French National Centre for Scientific Research)
CompCor	component based noise correction method
CSF	cerebrospinal fluid
DAQ	data acquisition system
DIVA	Directions Into Velocities of Articulators
DSP	discrete sequence production task
EPI	echo planar imaging
FAR	frontal associative regions

FD	framewise displacement
FDR	false discovery rate
FLIRT	FMRIB's Linear Image Registration Tool
fMRI	functional magnetic resonance imaging
FMRIB	fMRI of the brain
fMRIPrep	fMRI preprocessing pipeline
FOS	finger-to-thumb sequence opposition task
FOV	field of view
FSL	FMRIB software library
FTT	finger tapping task
FWHM	full-width at half-maximum
GLM	general linear model
GLMM	generalized linear mixed model
GM	grey matter
GODIVA	Gradient Order DIVA
GRec	gyrus rectus
Hipp	hippocampus
HSFC	hierarchical state feedback control
IFG	inferior frontal gyrus
INT	Institut de Neurosciences de la Timone (Timone Neuroscience Institute)
IOG	inferior occipital gyrus
IPG	inferior parietal gyrus
IPL	inferior parietal lobule
IRM	<i>see</i> MRI
k	cluster extent
L	left
L-VI	cerebellar lobule VI
L-VIII	cerebellar lobule VIII
LabVIEW	Laboratory Virtual Instrument Engineering Workbench
LG	lingual gyrus
LTM	long-term memory
M1	primary motor cortex
MarsBaR	MARSeille Boîte À Régions d'Intérêt (Marseille regions of interest toolbox)

MCR	motor cortical regions
MFG	medial frontal gyrus
MidFG	middle frontal gyrus
MidOG	middle occipital gyrus
MidTG	middle temporal gyrus
MNI	Montreal Neurological Institute
MoSFG	medial orbital superior frontal gyrus
MPRAGE	Magnetization Prepared - RApid Gradient Echo
MRI	magnetic resonance imaging
MSE	mean standard error
NI	National Instruments
PC	parietal cortex
PCG	precentral gyrus
PCL	paracentral lobule
PET	positron emission tomography
PoCG	postcentral gyrus
PoCiG	posterior cingulate gyrus
preSMA	pre-supplementary motor area
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
pSTG	posterior superior temporal gyrus
pSTS	posterior superior temporal sulcus
Put	putamen
R ²	coefficient of determination
R	right
ROI	region of interest
S1	primary somatosensory cortex
SBref	single-band reference
SC	spinal cord
SD	standard deviation
SFG	superior frontal gyrus
SI	special issue
SISL	serial interception sequence learning task
SMA	supplementary motor area
SMC	superior medial cerebellum
SOG	superior occipital gyrus

List of Tables

SPG	superior parietal gyrus
SPL	superior parietal lobule
SPM	Statistical Parametric Mapping software
Spt	Sylvian fissure at the parietal-temporal boundary
SRTT	serial reaction time task
SS	sensorimotor striatum
ST	striatum
STG	superior temporal gyrus
STS	superior temporal sulcus
t	t-statistic
T1w	T1-weighted image
T2w	T2-weighted image
TE	echo time
TI	inversion time
TMST	timed motor sequence task
TR	repetition time
UMR	Unité mixte de recherche (Joint Research Unit)
v	ventral
WAW	waveform audio file format
WM	white matter

1. Introduction

1.1. Cognitive processes underlying motor skill learning

In order to master a new motor skill, we need to acquire new, sequential, motor programs. This type of adaptive behavior, necessary for survival, communication or having social interactions (Gharbawie, Karl & Whishaw, 2007; Ninokura, Mushiake & Tanji, 2004; Shima et al., 1996; Hermer-Vazquez & Moshtagh, 2009), relies on various domain-general cognitive processes. We may begin to learn by acquiring the correct stimulus-response mapping, a process sustained by the basal ganglia (Graybiel et al., 1994). In a controlled experimental design, stimulus-response mapping usually means that the participant should associate seeing, for example, the numbers 1, 2, 3 and 4 on a computer screen with pressing a button with their index, middle finger, ring finger and little finger respectively. This is the underlying principle of commonly used motor sequence learning paradigms. In addition to having a correct stimulus-response association, one needs to be able to select and perform the appropriate motor action. In order to select the most appropriate action, we may need to suppress competing responses that could cause conflict, such as syllables that are similar but easier to pronounce, a similar-sounding stored lexical entry or the elevator button corresponding to the floor of our second residence. Response inhibition is also sustained by the basal ganglia (van den Wildenberg et al., 2006). After selecting the appropriate action, one proceeds to planning the execution itself. This planning is sustained both by motor cortical regions, in cingulate areas and in the pre-supplementary motor area (preSMA), but also the basal ganglia (Doyon et al., 2009). The latter two regions may be critically involved in predicting when the following event of the sequence should occur (Kotz, Schwartz & Schmidt-Kassow, 2009). Maintaining the sequence in its correct temporal order could be mediated by the prefrontal cortex (Doyon et al., 2002). Immediately preceding movement onset and during its execution, the premotor cortex would engage in anticipatory control (Grafton et al., 2008). This includes predicting

1. Introduction – 1.1. Cognitive processes underlying motor skill learning

the consequences of actions before actually performing them, and is increasingly used as learning progresses.

When the motor sequence is novel, such as a syllable to which a listener has had little or no exposure, these operations are performed for each segment of the syllable. During learning, bits of information that recurrently appear together in a particular order can be stored as chunks, or unique representations in memory (Miller, 1956; Gobet et al., 2001). Formation of chunks in a given sequence can be identified based on the similarities in response times between neighboring elements of the sequence (Boyd et al., 2009). Chunking of sensorimotor sequential actions, such as syllable production or more general motor sequence execution, is mediated by the basal ganglia (Graybiel, 1998; Boyd et al., 2009; Orban et al., 2011). For this reason, after a stroke leading to basal ganglia damage, patients are able to perform motor execution of the individual elements of a new sequence, but their motor performance is slowed down compared to healthy individuals, suggesting a difference in chunking (Boyd et al., 2009). In healthy populations, once the novel sequences are organized as chunks, motor programs specifying how they should be executed are created and stored cortically in long-term memory (Doyon et al., 2002; Lehericy et al., 2005; Grafton et al., 2008), and sustained by the preSMA in the medial frontal cortex (Kennerley, Sakai, & Rushworth, 2004), for faster retrieval (Cholin et al., 2008).

1.1.1. Internal modeling

When we chose to move our limbs to change gears or pick up a cup of coffee, or our vocal tract to pronounce a syllable, our central nervous system sends an efferent, exiting motor command. A copy of that signal, called an efference copy, is also created. It is converted into a corollary discharge (Sperry, 1950; Blakemore, Wolpert & Frith, 1998), i.e. sensory input corresponding to the consequences we predict our movement will produce (Wolpert, Ghahramani & Jordan, 1995). This prediction can then be compared to the actual sensory input we receive while performing the movement (Poon & Merfeld, 2005). When the prediction matches the actual sensory feedback, as is the case for well-mastered self-generated actions, conscious perception of the stimulation is reduced, the somatosensory cortex is not activated, and the cerebellar activity is decreased (Blakemore, Wolpert & Frith, 1998). Errors in sensory predictions, on the other hand, allow us to improve future movements (Sokolov, Miall & Ivry, 2017).

The process of predicting the sensory outcome of our actions is called a forward model. This is a type of internal model, a neural representation of how the elements

1. Introduction – 1.1. Cognitive processes underlying motor skill learning

of the external world would behave when manipulated (Imamizu et al., 2000; Kawato, Furukawa & Suzuki, 1987; Jordan & Rumelhart, 1992). As we learn how to perform an action, future sensory predictions are refined and prediction error signal is reduced (Sokolov, Miall & Ivry, 2017). Once the model is learned, the action can be executed precisely without the need to rely on external feedback (Ito, 2008). Forward models explain why we cannot tickle ourselves, and more importantly, allow us to differentiate the sensory input coming from our actions from other, unpredictable input (Blakemore, Wolpert & Frith, 1998). Originally conceptualized for examining motor actions, forward models are thought to guide such operations as eye movements, respiration and posture (Poon & Merfeld, 2005). They could also be applied to more abstract operations, notably those where real-time predictions are crucial. In language, a forward model would imply predicting upcoming linguistic information based on the current context (Sokolov, Miall & Ivry, 2017), or how our pronunciation of a syllable will sound. Forward models also find an application in social behavior, when we predict the social consequences of our acts or words, or when we create a model of the mental state of another person and predict how this state, or our actions, would influence their behavior (Sokolov, Miall & Ivry, 2017). Forward model alterations may be involved in thought disorder in schizophrenia (Sokolov, Miall & Ivry, 2017; Ford & Mathalon, 2012). A complementary form of internal models are inverse models, used to calculate the motor commands needed in order to arrive at a desired state (Blakemore & Sirigu, 2003). For instance, inverse models are used to determine how one should move their arm to pick up a cup of coffee or their vocal tract in order to pronounce the desired syllable. On a more abstract level, they could be used to understand what caused another person's specific behavior (Sokolov, Miall & Ivry, 2017).

Internal models are mainly sustained by the cerebellum (Ito, 2008). With its high level of connections to large portions of the cerebral cortex, the cerebellum can base its predictions regarding the future states of both the outside world and of neural activity elsewhere in the brain on various inputs (Sokolov, Miall & Ivry, 2017; Imamizu, 2000; Ito, 2000). There are several manners in which the cerebellum could contribute to internal modeling. As it responds to incongruent stimuli, the cerebellum could have a role in error prediction. The same sensitivity to errors is also observed in cortical regions, as widely as in temporal, parietal, and frontal areas (Sokolov, Miall & Ivry, 2017), raising the question of possible cerebellar specialization in an aspect of internal modeling and its relation with the rest of the brain. One plausible account is that the cerebellum generates predictions of a future state in an automatic fashion,

1. Introduction – 1.1. Cognitive processes underlying motor skill learning

independently of conscious awareness (Blakemore & Sirigu, 2003). It could also be that these predictions are compared with the actual action outcomes elsewhere, in the cerebral cortex or in the basal ganglia, and that the calculated error signal is communicated back to the cerebellum, so as to improve future predictions. In addition, the cerebellum likely contributes to ensuring smooth, coordinated processing within broader, fronto-parietal and fronto-temporal networks (Sokolov, Miall & Ivry, 2017). In speech articulation, internal models, particularly under demanding circumstances such as when sequences are phonotactically illegal in one's native language, also rely on brain regions involved with speech production planning and auditory-based processing. These regions include Broca's area, anterior insula, anterior superior temporal sulcus/gyrus (STS/STG), planum temporale, superior temporal parietal area and supramarginal gyrus (Callan et al., 2004).

1.1.2. Performance monitoring

In order to ensure that our actions result in a desired outcome and determine when our behavior needs adjustment, we must constantly compare our actual performance and its consequences to that intended. This monitoring process is particularly challenging in situations that are not optimal, such as in a noisy environment, while experiencing higher cognitive load, speaking in a recently learned foreign language or driving a new car (Oomen & Postma, 2002; Albarqi & Tavakoli, 2022).

Performance monitoring operates externally, after the action has been performed, as well as internally, during the action planning (Peterburs & Desmond, 2016). External monitoring relies on the sensorimotor and auditory feedback from the external world, as well as on our bodily representations and proprioceptive feedback (Faivre et al., 2020). We compare the information about the actual outcome of our action to the predicted outcome and, when there is a mismatch, a conflict is detected (Reason, 1978). Mismatches between the predicted and the actual outcome of the action are used to adjust the current performance and to improve the future one. In an experimental design, this could mean the participant receiving feedback from the experimenter on whether she pressed the correct button or not, or realizing herself she committed an error upon seeing the wrong light on the screen. In speech production, while hearing our own speech, we compare the auditory feedback of our voice to the utterance we thought we would pronounce.

In addition to external monitoring, it is important to monitor our actions internally while planning them, so that we can intercept errors before they occur. When

1. Introduction – 1.1. Cognitive processes underlying motor skill learning

errors in performance are detected, they are intercepted by modifying motor commands. The update in motor commands modulates the prediction of the sensory consequences of our action as well as our neural response to it, refining the forward model (Blakemore, Wolpert & Frith, 1998). Internal monitoring allows us to correct our utterances before finishing them, should there be an important mismatch with what was intended (Levelt, 1983). It also enables us to report errors in inner, non-articulated speech, and even to avoid lapsus resulting in a taboo word (Motley, Camden & Baars, 1982)¹.

Performance monitoring is generally thought to rely on frontal, striatal and cerebellar regions of the brain (Peterburs & Desmond, 2016), with increased parietal activity in conjunction with greater use of feedback (Grafton et al., 2008). Various accounts for error monitoring in speech production have been proposed (see Postma, 2000 for a review), focusing on the cerebellum, medial frontal and temporo-parietal cortices (Runnqvist et al., 2021). These accounts converge when it comes to monitoring skills that are both motor and cognitive in nature, such as producing novel syllables. For instance, cerebellar activation in performance monitoring has been widely reported, but its precise role is still unclear (Peterburs & Desmond, 2016). As an interface between motor behavior and cognition (Stoodley & Schmahmann, 2010; Peterburs & Desmond, 2016), the cerebellum is specifically presumed to be involved in monitoring speech production. Cerebellar activity has indeed been detected in both internal and external speech error monitoring, when errors occur on a lexical, non-articulatory level (Runnqvist et al., 2016; 2021). When monitoring articulatory speech errors, one straightforward expectation is that the motor regions of the cerebellum monitor the motor aspects of speech production, and those sustaining cognitive functions its non-articulatory aspects. More interestingly, the cognitive areas could also maintain a hierarchical role and regulate speech monitoring even on the articulatory level.

Another broad area implicated in performance monitoring is the medial frontal cortex, which could be involved in different ways. It may increase the monitoring by signaling the presence of a conflict between several competing responses (Nozari, Dell & Schwartz, 2011), such as the syllables beginning with the same phoneme but that differ as regards their structure or degree of complexity. It is noteworthy that selecting one response or another would not eliminate conflict, although it could resolve it. Another way the medial frontal cortex could contribute is via external monitoring, through verbal auditory feedback (Loh et al., 2020; Runnqvist et al., 2021).

Finally, we detect errors while understanding speech, and, according to some accounts, it is unimportant whether the speech is ours or someone else's (e.g., Levelt, Roelofs & Meyer, 1999; Hartsuiker & Kolk, 2001; Roelofs, 2020). As a region known for its role in speech comprehension (Indefrey & Levelt, 2004), the temporal cortex has been proposed as the neural basis for monitoring in these theories, as well as in some more recent models (e.g., Tourville & Guenther, 2011; Guenther, Ghosh & Tourville, 2006; Hickok, 2012; Hickok, 2014; Pickering & Garrod, 2013). Although its role may be limited to external monitoring (Hansen, McMahon & de Zubicaray, 2019; Okada, Matchin & Hickok, 2018; Runnqvist et al., 2021), some views expand it to internal monitoring as well (Hirano et al., 1997; Shergill et al., 2000; Indefrey & Levelt, 2004).

1.2. Motor sequence learning

Motor skill learning relies on multiple cognitive subcomponents. Studying these subcomponents in laboratory conditions requires some simplifications and easily implemented setups that aim to exclude the involvement of less specialized, but required cognitive skills, such as attention or cognitive control (Doyon et al., 2018). In order to select the most appropriate setup, one must first distinguish between motor sequence learning and motor adaptation, which are often modeled separately as they show strong evidence of striatal and cerebellar specialization respectively (e.g. Doyon, Penhune & Ungerleider, 2003). While both result in effortless performance achieved through repeated practice (Willingham, 1998), motor adaptation usually involves manipulating the relationship between the movements one performs and their outcomes (Doyon, Penhune & Ungerleider, 2003). Typical motor adaptation paradigms include reaching a target, either with one's limbs, a joystick or a robotic arm. For example, a participant might be required to manipulate a joystick in a computer game, while the cursor has been programmed to move at a speed that does not correspond to the force she applies (Flament et al., 1996; Shadmehr, Mussa-Ivaldi & Bizzi, 1993). In speech production, motor adaptation is usually studied through modifications in the auditory or somatosensory feedback that is presented to the participant. For instance, drawing a parallel between speech production and hand reaching movements, Houde and Jordan (1998) changed the vowel formants produced by their participants, modifying the relation between the initial position and target configuration. Lowering the first formant and increasing the second, the participants heard "i" they pronounced as "a", and compensated for this change by

1. Introduction – 1.2. Motor sequence learning

modifying their pronunciation. Auditory feedback can also be administered with a delay, leading the speakers to adopt compensatory strategies such as increased speech rates (Katz & Lackner, 1977). Specially designed dental appliances can be programmed to apply forces to the lower jaw during speech production, resulting in modified jaw movements (Darainy, Vahdat & Ostry, 2019). In motor sequence learning, on the other hand, we incrementally acquire movements of a novel sequence, performing novel combinations of movements, as is the case with a novel syllable type in a foreign language. Learning is typically indexed by a decrease in the number of errors, or by the reduction of time between the execution of one element of a sequence and the next (Doyon, Penhune & Ungerleider, 2003). A real-life example of motor sequence learning and motor adaptation required in table-top games is reported in Appendix A.

Acquiring expertise in complex motor skills requires both explicit instruction and implicit practice (Sanchez & Reber, 2013). Learning some motor skills, such as playing a new melody on a musical instrument, can be facilitated if we first become familiar with the melody and then rely on this explicit knowledge to learn how to perform it (Doyon, Penhune & Ungerleider, 2003). This is also often the case when studying a foreign language initially, when we are first presented with some basic vocabulary and a model of how to pronounce those novel sound sequences. However, most of the motor skills we master are learned implicitly, without prior explicit knowledge. Even the motor skills for which we first relied on our explicit knowledge are performed implicitly once they are overlearned, without conscious recollection (Doyon, Penhune & Ungerleider, 2003). For instance, we do not rehearse the pronunciation of a word in a foreign language in which we have become fluent, nor do we think about the different components of the action of brushing our teeth, or changing gears, once we have internalized this knowledge. Implicitly learned skills tend to be more robust (Jiménez, Vaquero & Lupiáñez, 2006). In sports-related motor tasks, slightly greater effectiveness of implicit as compared to explicit learning becomes apparent after consolidation (Kal et al., 2018). Implicitly learned skills are also more context-dependent than explicitly learned ones (Berry et al., 1993). For instance, the procedural knowledge necessary for riding a bicycle may not be forgotten, but it will be of limited help for riding a monocycle. On the other hand, explicit knowledge enables the learner to adopt specific learning strategies, so that the learning effects are maintained despite contextual changes (Jiménez, Vaquero & Lupiáñez, 2006). These differences are also reflected at the neuropsychological level. Indeed, explicit and implicit memory systems are clearly separated in the brain. The explicit memory system is located in

1. Introduction – 1.2. Motor sequence learning

the medial temporal lobe, and the implicit memory system depends on a number of different brain regions (Reber, 2008). During learning, implicit and explicit memory systems operate independently (see Loonis et al., 2017 for a meta-analysis that found different neural oscillation patterns for implicit and explicit learning in non-human primates). Explicit skill learning relies on the explicit memory system, and is therefore sustained by the medial temporal lobe (Sanchez & Reber, 2013), while implicit skill learning recruits the basal ganglia (Doyon et al., 2009).

Another important factor in motor sequence learning studies is that of the practice schedule: massed or distributed. In massed practice, there are no intermittent pauses (Singer, 1980), or the pauses are shorter than the practice time in each session. In distributed practice, rest intervals are relatively longer compared to massed practice (Lee & Genovese, 1988). Substantial experimental evidence has shown that distributed practice seems to favor improvement and retention in motor learning in general (for some early studies, see Lee & Genovese, 1988 for a meta-analysis, but also Singer, 1965; Stelmach, 1969). This is also true of playing video games (Metalis, 1985), tennis (Fuentes-García et al., 2022), or speech motor sequence learning (Kaipa et al., 2020). Effects of distributed practice were demonstrated both when all the practice takes place in a single day (Kwon, Kwon & Lee, 2015), and over several years (Bahrack et al., 1993).

Allocating attentional resources to another task could be detrimental to learning, even when learning is implicit (Nissen & Bullemer, 1987). In particular, a secondary task negatively impacts learning in older adult populations, as shown in studies of motor learning during rehabilitation (Nejati et al., 2008) and speech motor sequence learning in healthy older adults and those suffering from Parkinson's disease (Whitfield & Goberman, 2017). However, performing another task while learning may also be beneficial for learning (Goh et al., 2012; Goh, Lee & Fisher, 2013; Roche et al., 2007; Hemond, Brown & Robertson, 2010). Interestingly, these opposing results may be explained by the choice of the task, and the apparent learning impairment could be explained as a disruption to task performance rather than learning (Schumacher & Schwarb, 2009). Another source of discrepancy in dual task experiments may come from the task difficulty. As Polskaia and her colleagues (Polskaia et al., 2023) recently suggested, the prefrontal cortical activation needed to maintain performance during a dual task learning paradigm depends on the difficulty of the sequence being learned. Indeed, the task difficulty is an important factor when designing motor sequence learning studies (see Witt, Laird & Meyerand, 2008 for a meta-analysis of neuroimaging

motor sequence learning studies with different task complexities), as increasing the complexity of novel sequences may cause deficits in learning rate and magnitude, as seen in older adults (King et al., 2013) or stroke survivors (dos Santos Palma, Correa & Torriani-Pasin, 2020). Interestingly, complexity seems less important in speech motor learning compared to non-speech motor learning literature (Kaipa, 2016).

1.2.1. Serial reaction time task

In motor sequence learning studies, participants are often required to tap their fingers at a pace given by an auditory or visual stimulus, or at their own pace, and perform sequences of different complexity or repeating patterns. These tasks reveal a broad network of motor cortical regions, namely the primary sensorimotor cortex, the supplementary motor area (SMA), the basal ganglia and the cerebellum (Witt, Laird & Meyerand, 2008). However, as Witt and her colleagues argue in their meta-analysis of finger tapping task (FTT) studies, variations in the experimental paradigms make the results of those studies difficult to interpret. It is therefore interesting to examine the most commonly used finger-tapping tasks.

The prototypical example of a motor sequence learning paradigm is the serial reaction time task (SRTT). In the standard version of the task (Nissen & Bullemer, 1987), the participant responds to a stimulus, such as a light or an auditory cue, by pressing buttons associated to those cues as quickly as possible. Learning is then indexed by a decrease in reaction times (Verwey, Shea, & Wright, 2015; for a neurocognitive model of sequence representation based on the SRTT, see Keele et al., 2003). This task, although simple to implement and adapt to a specific research question (see Du & Clark, 2018 for a study using SRTT with foot stepping), still recruits various cognitive processes, such as stimulus-response mapping, response selection, inhibition of competing responses, motor execution, error detection and performance monitoring; as learning occurs, the internal model of the task is refined. Neuroimaging studies that use the SRTT revealed neural circuits involving premotor, medial frontal and parietal cortices, as well as the cerebellum and the basal ganglia, together with some other areas such as the hippocampus and the spinal cord (Hardwick et al., 2013; Janacsek et al., 2020; Doyon et al., 2018).

1.2.2. A neurobiological model of motor sequence learning

In the motor sequence learning model developed by Doyon and his colleagues (Doyon et al., 2018; Doyon et al., 2002; Doyon & Benali, 2005; Doyon et al., 2009), the authors argue that motor sequence learning in its early phase depends on a large network of brain structures. Those areas range from motor cortical areas, which include the premotor cortex, the SMA, the preSMA and the anterior cingulate cortex (ACC), frontal areas, temporal areas including the hippocampus, parietal areas, the basal ganglia, the cerebellum and spinal the cord (see Figure 1.1, adapted from Doyon et al., 2018). Which of these structures will actually be involved depends on the cognitive processes required during the initial practice of the task at hand (Doyon et al., 2018). The specificities of the sequence that is being learned can also determine which cognitive subcomponents of learning are involved. In other words, it is likely that the neural correlates of motor sequence learning revealed in functional neuroimaging studies may reflect cognitive processes involved during the initial practice of a particular task. The differences in the experimental design, related to cognitive processes on which they focus, may explain the lack of consensus on the neural correlates of motor sequence learning despite this being the focus of a number of studies (Berlot, Popp & Diedrichsen, 2020).

1.3. Speech sequence learning

The process of speech motor sequence learning, in which a speaker achieves, through practice, an increasingly accurate production of a novel combination of speech sounds (Segawa et al., 2015) has been studied less extensively than motor sequence learning in general. One of the first neuroimaging studies of speech motor sequence learning was conducted by Rauschecker, Pringle and Watkins (2008) in a paradigm using covert repetition. The authors examined the neural correlates of learning novel phoneme combinations that varied in length and complexity and were presented either once or four times. Learning was indexed by error rates in post scanning behavioral testing, and by participants' self-reported time of the end of production during the fMRI acquisition through button presses. After only a few expositions to the novel sequences, the activation was reduced in a broad network that included the premotor cortex, the SMA, the inferior frontal gyrus, the superior temporal cortex, the basal ganglia and the cerebellum. Segawa and her colleagues (Segawa et al., 2015) had participants overtly pronounce syllables of different phono-

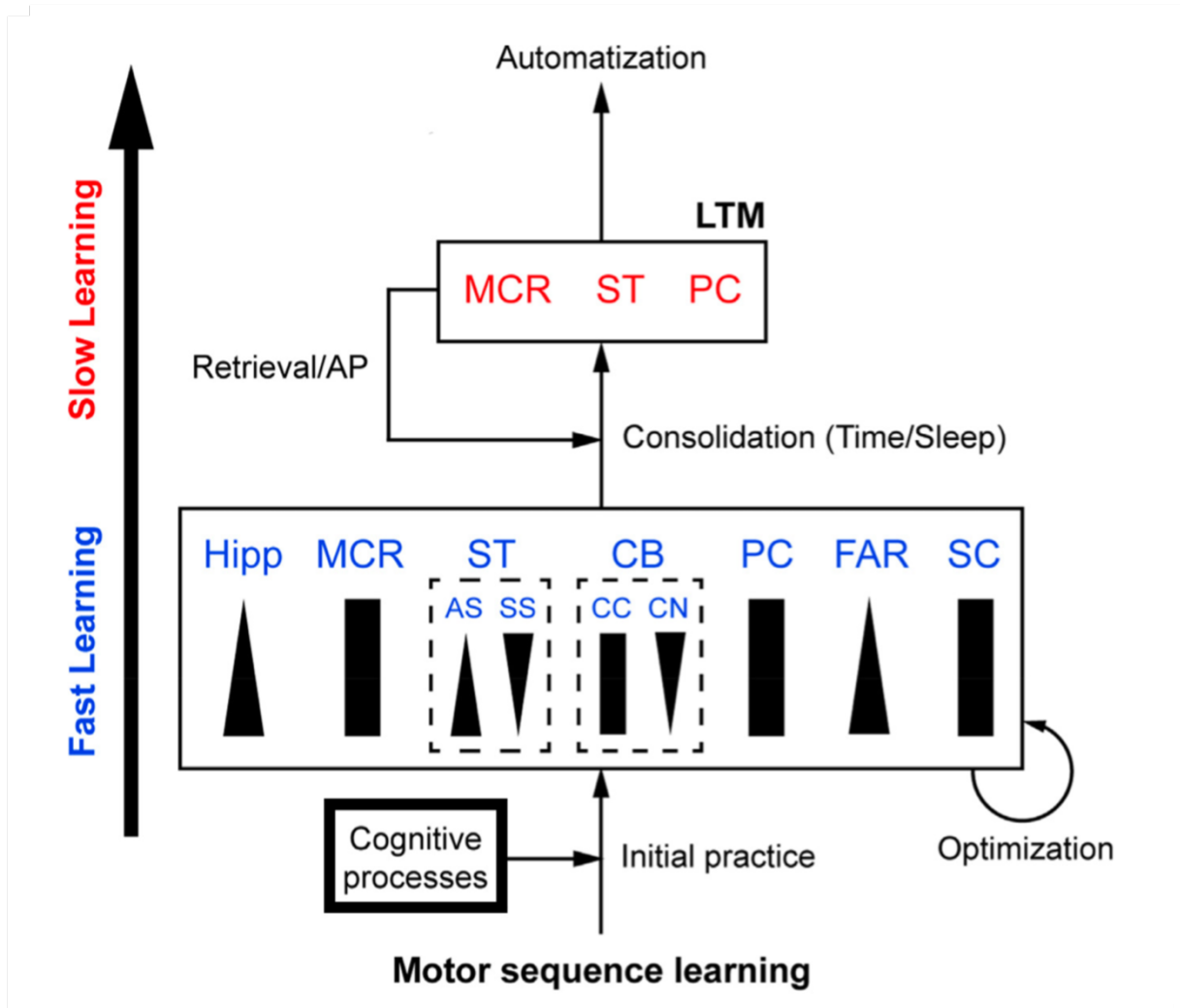


Figure 1.1. – *Doyon's neurobiological model of motor sequence learning*. According to the model, recruitment of specific brain areas during initial practice depends on the cognitive processes involved. Hipp = hippocampus, MCR = motor cortical regions, ST = striatum, CB = cerebellum, PC = parietal cortex, FAR = frontal associative regions, SC = spinal cord, AS = associative striatum, SS = sensorimotor striatum, CC = cerebellar cortex, CN = cerebellar nuclei, LTM = long-term memory, AP = additional practice, ▲ = the region is decreasingly involved, ▼ = the region is increasingly involved. Adapted from Doyon et al., 2018.

tactic probabilities, i.e., probabilities that two consonants occur together in a given language. In addition, the syllables were either completely novel or already practiced the previous day. Learning was indexed through error rates and utterance duration, and only correct trials were included in the analyses. The detected activations were fairly similar to those found in the covert repetition study (Rauschecker, Pringle & Watkins, 2008). Namely, novel syllables elicited more activation in the premotor cortex, the preSMA, the superior temporal and parietal cortices and the basal ganglia. Production of phonotactically illegal syllables, in which the consonant clusters have low phonotactic probability, was compared to the legal baseline, in which the consonant clusters have higher phonotactic probability. Activations were detected in the premotor cortex, the SMA, the superior parietal lobule (SPL), the posterior superior temporal gyrus (pSTG) and the cerebellum. Overt production and planning of novel speech sequences was directly compared by Bohland and Guenther (2006), indicating a broad initiation network that included primary motor, somatosensory, auditory and insular cortices as well the thalamus, the basal ganglia and the cerebellum.

While these studies allow us to discern the brain regions at work during an articulation learning task, they also raise important questions about the precise contribution of these regions to speech production. In an fMRI study, Basilakos and her colleagues (Basilakos et al., 2018) found evidence that both articulation and non-speech oral-motor movement production showed a similarly strong response in the articulation-responsive regions within the precentral gyrus, with stronger responses to conditions that required more complex articulatory movements. The present dissertation aims to provide novel information pertaining to this issue.

1.3.1. Neurocognitive models related to syllable learning

1.3.1.1. The Hierarchical State Feedback Control (HSFC) model of speech production

The HSFC model is the first speech production model to integrate psycholinguistic and motor control theories (Hickok, 2012). In this model, illustrated in Figure 1.2, production errors can be detected and corrected both internally, prior to the articulation, and externally, based on the auditory and somatosensory feedback available once the articulation has taken place. The two types of feedback are parts of parallel, communicating, hierarchically organized loops. The auditory feedback loop is considered as higher level, corresponding to syllable articulation, and the somatosensory

1. Introduction – 1.3. Speech sequence learning

feedback loop as lower level, corresponding to phoneme articulation. Production begins with the selection of the motor programs indicating how the articulators should be moved to achieve correct pronunciation. In the case of syllables, motor programs are located in the frontal cortex, in Brodmann area 44. The auditory targets, or the way the syllable should ideally sound, are located in temporal structures (STG/STS) (Hickok, 2012). These two components are connected through coordinate transform, which allows for sensorimotor integration. The transformation is required as the movement errors are not specified in motor terms, but in terms of sensory outcomes, such as proprioceptive or auditory signals, and hence they must be translated into motor terms in order for the feedback to be used for future motor improvement (Wolpert, Miall & Kawato, 1998). According to the HSFC model, the coordinate transform for syllables is located in the Sylvian fissure at the parieto-temporal boundary (Spt). Motor selection, forward prediction, error detection and correction are part of the same mechanism. In this model, the emphasis is on the cortical contribution to speech coordination and comprehension. The basal ganglia have a supportive role via their general contributions to cognitive processes such as response selection. The role of the cerebellum is acknowledged but limited to a lower level coordinate transform of the phonemes.

1. Introduction – 1.3. Speech sequence learning

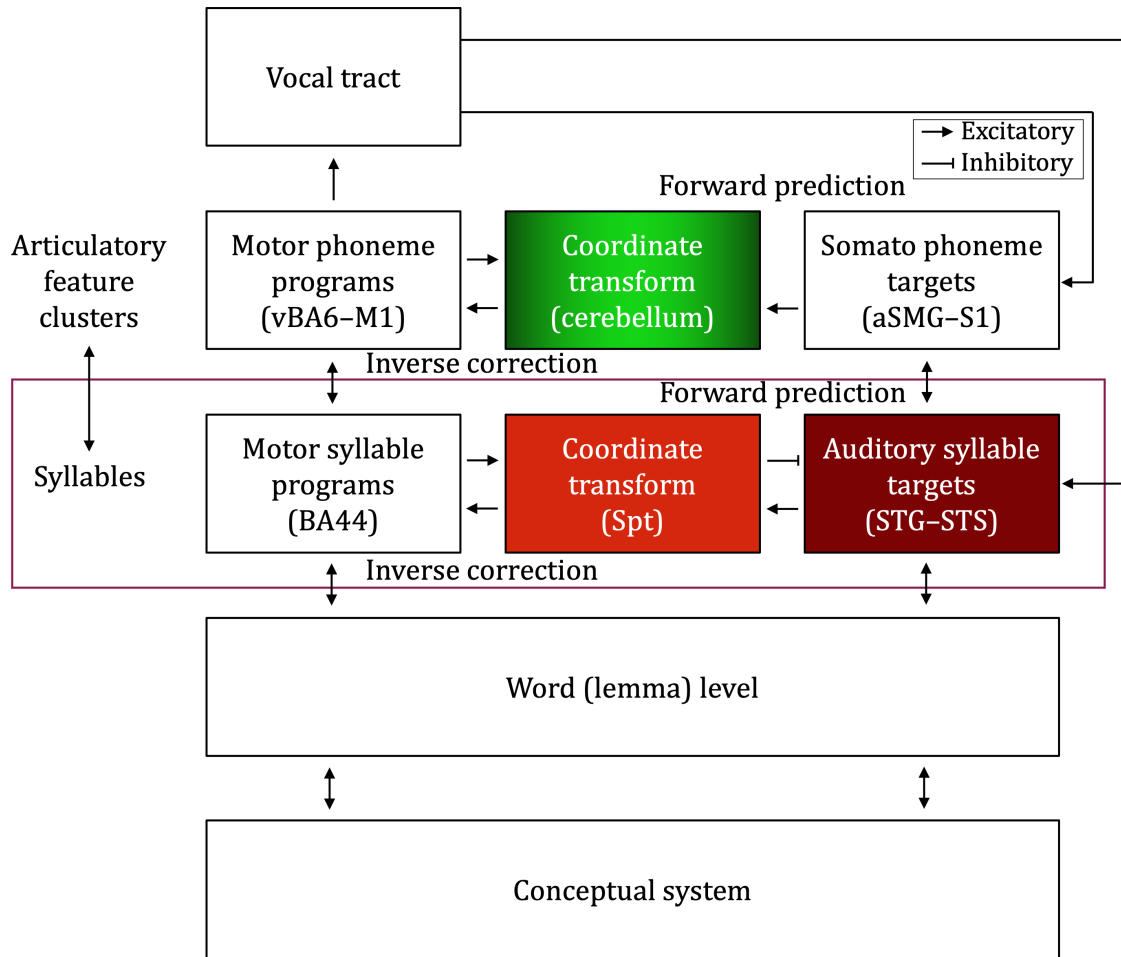


Figure 1.2. – *The Hierarchical State Feedback Control model of speech production.* v = ventral, BA = Brodmann area, M1 = primary motor cortex, aSMG = anterior supramarginal gyrus, S1 = primary somatosensory cortex, Spt = Sylvian fissure at the parietal-temporal boundary, STG = superior temporal gyrus, STS = superior temporal sulcus. Adapted from Hickok, 2012.

1.3.1.2. The Directions into Velocities of Articulators (DIVA) model of speech motor control

One of the most detailed and complete models of speech production and motor control, the DIVA model (Figure 1.3) accounts for computational, neural and cognitive aspects of speech motor control (Guenther, 1994; 2006). It is constantly updated (Guenther, Ghosh & Tourville, 2006; Golfinopoulos, Tourville & Guenther, 2010; Tourville & Guenther, 2011) and extended (Bohland, Bullock & Guenther, 2010; Kearney et al., 2020; Miller & Guenther, 2021; Weerathunge et al., 2022; Zuk, Loui & Guenther, 2022) and allows researchers to empirically test hypotheses from domains as diverse as pathology, development, adaptation and learning (Civier, Tasko & Guenther, 2010; Terband et al., 2014; Martin et al., 2018; Terband, Rodd & Maas, 2020). In this family of models, producing a commonly used articulatory sequence such as a syllable or a word requires choosing and executing the most appropriate available set of muscle commands to achieve the targeted phonological output (Miller & Guenther, 2021). In DIVA, similarly to the HSFC model, accurately producing a syllable means using motor commands located in the ventral primary motor cortex in the frontal lobe. The motor commands serve to attain both an auditory speech sound target, i.e. the auditory representation of the syllable, and a somatosensory target corresponding to its somatosensory representation. Akin to the HSFC model, the auditory targets lie in the temporal cortex, in the pSTG, and the somatosensory targets are located in the parietal cortex. Motor commands in the DIVA model are fine-tuned based on the auditory and somatosensory external feedback. A mismatch between the intended and the produced syllable, based on the external feedback, creates an error signal that is used to update motor commands. The model accounts for the role of subcortical areas, which are integrated into larger loops. The basal ganglia form part of the initiation circuit, and the cerebellum of the articulation circuit, where it contributes to precisely timed activation. An extension of the model, Gradient Order DIVA (GODIVA), accounts for the sequencing process in which transformations of the phonological representations in working memory that become properly timed motor sequences (Bohland, Bullock, & Guenther, 2010). Both models share the motor loop, responsible for generating the highly coordinated articulatory sequence, sustained by the SMA, thalamus, putamen and the premotor cortex. The planning loop, added in the GODIVA model, temporarily stores the elements of the upcoming phonological sequence. It involves the inferior frontal sulcus, the caudate nucleus, and the preSMA (Miller & Guenther, 2021). Of particular interest to syllable learning studies, the models

1. Introduction – 1.3. Speech sequence learning

postulate that infrequent syllables do not have stored motor programs, so they would have to be created during learning of novel consonant clusters (Guenther, Ghosh & Tourville, 2006).

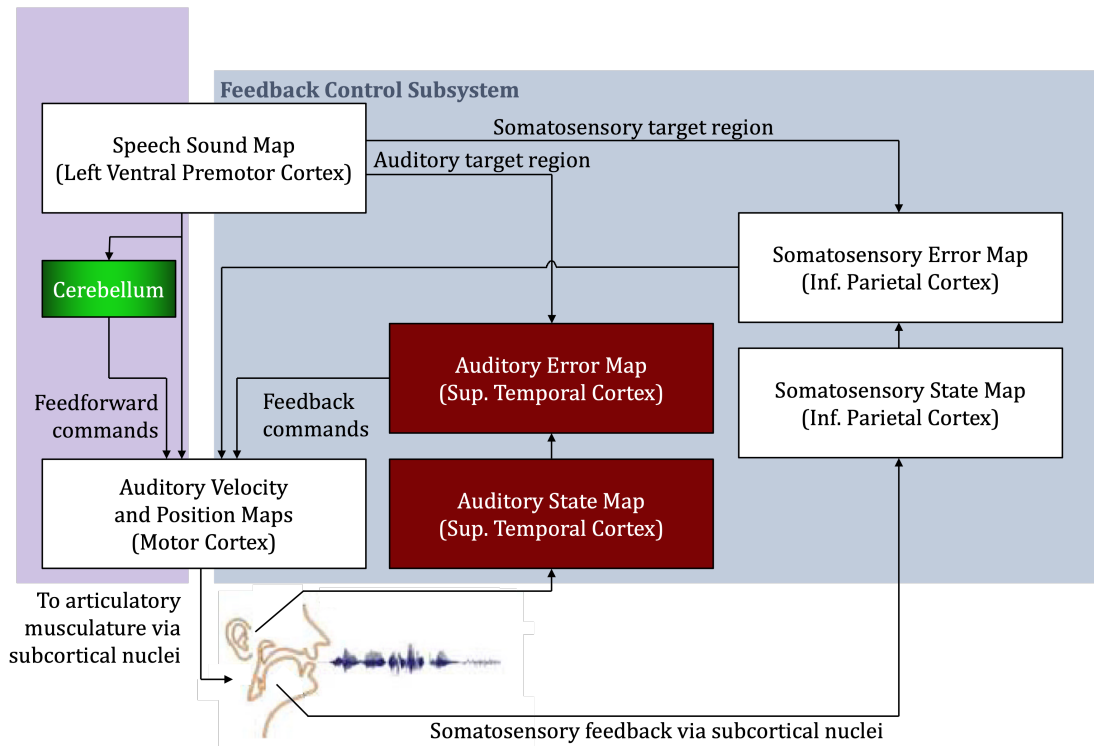


Figure 1.3. – *The Directions into Velocities of Articulators model of speech motor control.*
Adapted from Guenther, Ghosh & Tourville, 2006

1.3.1.3. A subcortico-cortical framework of speech perception and production

The framework proposed by Kotz and Schwartz (2010) (Figure 1.4) takes into account the temporal aspect of speech, as speaking necessarily involves arranging speech constituents, such as phonemes in a syllable, in the temporal dimension. It also highlights the roles of the basal ganglia and the cerebellum, in addition to the cortex, which generates internal representations of the syllables' temporal structures. In this framework, pronouncing a syllable, or any other speech event, begins by activating its memory representation, which is transmitted from the temporal cortex to the frontal cortex. There, the syllable takes its form of a sequence organized in time, which is provided by the preSMA together with the cerebellum and the basal ganglia. These three regions serve as a pacemaker in the planning of speech production. Next, the premotor cortex, the primary motor cortex, and the SMA in particular, use this information for motor control during the articulation (Kotz & Schwartz, 2010). Although this framework does not explicitly account for novel syllable acquisition, it highlights the importance of recurrence for facilitating speech processing and, as a consequence, production. Events that are less likely to occur together, such as phonemes ordered in an unfamiliar manner within a syllable, would be predicted less accurately and processed with more difficulty, which may be reflected in the cortico-subcortical loops involved in speech perception and production.

1. Introduction – 1.3. Speech sequence learning

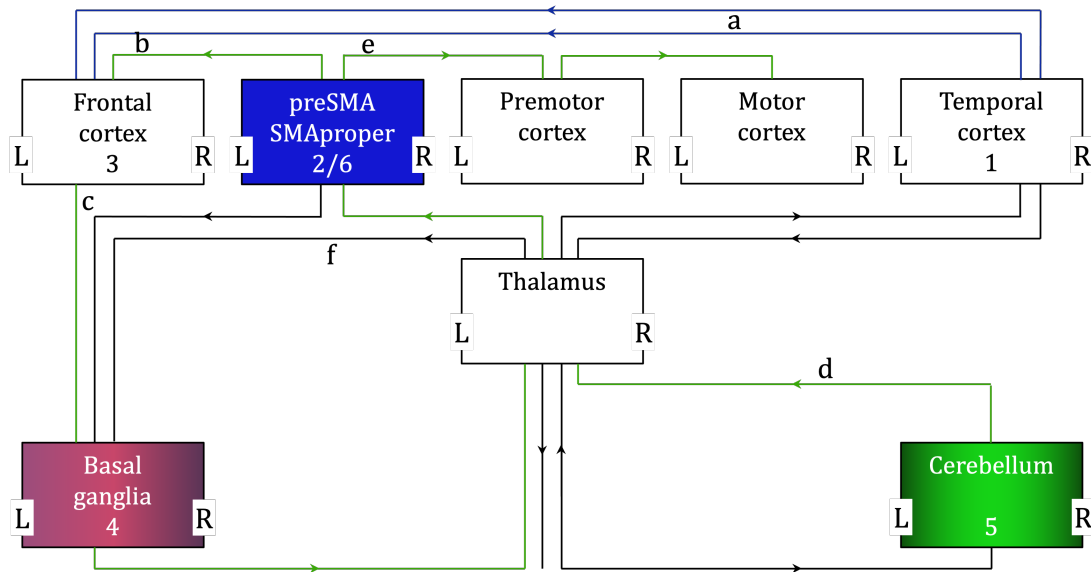


Figure 1.4. – A subcortico-cortical framework of speech production. Adapted from Kotz & Schwartze, 2010.

1.4. Principles of functional magnetic resonance imaging

Functional magnetic resonance imaging (fMRI) is based on the fact that neural processing requires oxygen, provided by hemoglobin in the blood cells surrounding neurons. Providing oxygen changes the ratio between deoxygenated and oxygenated hemoglobin. Due to their different magnetic properties, blood carrying one or the other type of hemoglobin will behave differently in the presence of a magnetic field. This difference creates blood oxygen level dependent (BOLD) signal detected by fMRI (Ogawa et al., 1990).

BOLD signal can be used to study cognition in various ways. A typical experimental paradigm consists in contrasting, on one hand, the BOLD signal measured while a participant performs a task supposed to isolate the phenomenon we want to study, and on the other hand, the BOLD signal acquired during a control condition (Amaro & Barker, 2006). For instance, to detect the brain areas that are active when the subject learns to pronounce a novel type of syllables, one approach would be to compare the BOLD images during repeated production of novel syllables with those acquired during repeated production of more familiar ones. Alternatively, we could opt for an equally controlled baseline condition, but tailored to isolate a different cognitive subcomponent of learning, such as internal and external self-monitoring. Another scenario common in non-speech motor sequence learning studies is contrasting the BOLD signal during the production of a sequence that is being learned with that of an equally novel, but random sequence, whose non-repetition prevents learning.

Whichever two conditions we choose to compare, the contrast will likely take place inside a general linear model (GLM) regression analysis where the observed BOLD signal of a given area, usually whole brain divided in voxels, or an anatomically or functionally defined region of interest (ROI), is tentatively explained by a set of parameters (Weber et al., 2015; Friston et al., 1999). Concretely, using a linear regression of the usual form $y = X\beta + \epsilon$, the observed BOLD time-course intensity for each given voxel (y) can be modeled as a sum of predetermined variables, predictors (X), and an error estimate (ϵ), which is the unexplained part of the BOLD signal. This is done for each time point available. The way the predictors are summed is such that the error is kept minimal, so different weights (β) are attributed to each predictor (Monti, 2011). Contrasting two conditions then means contrasting the β values of these conditions,

1. Introduction – 1.4. Principles of functional magnetic resonance imaging

usually through a statistical test such as a t-test. First, the average values in each condition, are compared in a within-subject analysis. This is followed by a second level or group analysis. Once we have determined whether the difference between the two conditions is significant in any given voxel, a correction for multiple comparisons is applied to avoid false positives due to the large number of tests. The results are usually represented in form of an activation map where each voxel is colored according to the intensity of the signal, represented via the value of the t-statistic (Chanoine & Badier, 2016).

1.4.1. Functional connectivity

Functional magnetic resonance imaging can be used to study cognition not only while the subject is performing a task, but also at rest, as resting state data reveals correlations between low-frequency signal fluctuations of anatomically distant, but functionally related brain areas (Biswal et al., 1995). Those functional networks are experience-dependent and can reorganize with practice (Jolles et al., 2013). Contrasting resting state data before and after a learning task can shed a different, complementary light on the brain networks that sustain learning. The resting state functional connectivity can even predict task-induced BOLD activity (Mennes et al., 2010), similarly to the activity of one brain region at a given time predicting another region's future activity (Sokolov, Miall & Ivry, 2017).

1.4.2. Coordinate based meta-analysis

While a single-neuroimaging study can address specific research questions, multiple studies can be compared to investigate how consistent activations are across studies treating a similar topic (Turkeltaub et al., 2002). This approach is called a meta-analysis. The meta-analysis we used in this dissertation was based on activation likelihood estimation (ALE). In this analysis, precise foci are transformed into probability distributions centered at the peak coordinates, summarizing the reported findings. A value equal to the probability that at least one of the points in the dataset actually falls within a voxel is assigned to each voxel in the brain for each data point. The union of the probabilities for all the data points gives the ALE score (Turkeltaub et al., 2002).

1.5. The objectives of this dissertation

At the boundary of motor and speech learning, pronouncing novel speech sequences relies on cognitive mechanisms shared with other adaptive sequential behaviors. A handful of neuroimaging studies that examined this phenomenon, as well as neurocognitive models of motor sequence learning, on one hand, and of speech perception, production and motor control, on the other, suggests the implication of various cortical areas, as well as the cerebellum and the basal ganglia. Taking into account both motor learning and speech production frameworks, this dissertation aims to address the domain-specific and domain-general role of these areas and the ways in which they connect during the acquisition of sequential (speech) motor behavior.

In motor sequence learning, the actual involvement of frontal and parietal cortical areas, as well as subcortical areas including the cerebellum and the basal ganglia, and even the spinal cord depends on the cognitive processes that are required during the initial practice of the task at hand (Doyon et al., 2009; 2018). The cognitive processes that enable this learning can be accounted for by using specific experimental designs. The lack of agreement on the specific neural correlates of motor sequence learning despite the numerous studies on this question (Berlot, Popp & Diedrichsen, 2020) could be explained by task-related effects, and contradictory results of otherwise identical functional neuroimaging studies could be explained by the choice of the baseline condition (Brandt, 2006). In Chapter 2, we argue that the different choices in the experimental design could be turned into a useful tool for better understanding the cognitive subcomponents of motor sequence learning and their neural correlates. Using a meta-analytic approach, we examined the influence of using a particular task or baseline to assess motor sequence learning on the converging pattern of the activated brain regions. We argue that the detected differences and similarities reflect the different cognitive processes involved and we propose a tentative refinement of Doyon's model of motor sequence learning (2018) through linking those processes to their neural correlates.

In line with the consideration that certain cognitive mechanisms are shared between the acquisition of mainly motor skills, such as moving one's fingers, and more cognitive ones, such as language production, the second goal of this thesis was to examine whether one of those mechanisms, namely performance monitoring, operates in a similar manner and recruits similar brain structures when monitoring more or less motor related action performance. Indeed, two scenarios regarding the cerebellar

1. Introduction – 1.5. The objectives of this dissertation

monitoring of speech production could be expected. The parts of the cerebellum that are usually involved in motor functions could monitor motor aspects of speech production, such as articulation and syllabification, and those that are usually involved in cognitive functions could monitor pre-articulatory levels of language production, such as selecting the correct lexical candidate. Alternatively, areas of the cerebellum underlying cognitive function could perform a hierarchically superior form of monitoring across all types of linguistic representations, including motor ones (Sokolov, Miall & Ivry, 2017). In Chapter 3, we used a speech production paradigm where a post-lexical, articulatory variable was manipulated while participants pronounced syllables of different difficulty that therefore required different monitoring loads during functional magnetic resonance imaging (fMRI). The cerebellum is traditionally considered to be involved in both motor and cognitive functions, and its activation has been observed in speech production while monitoring the words for their lexical content, on a non-motor, pre-articulatory level (Runnqvist et al., 2021). Using an articulatory task, we examined motor and cognitive aspects of performance monitoring on a motor, articulatory level. Additionally, internal and external monitoring contrasts were used to disentangle the medial frontal cortex conflict-based (Nozari et al., 2011; Gauvin & Hartsuiker, 2020) and error-based (Loh et al., 2020) monitoring role. Finally, we aimed to inform an ongoing debate on the role of the temporal cortex in speech monitoring, as there is evidence that the pSTG sustains external monitoring, but its role in internal monitoring remains unclear (Meekings & Scott, 2021).

In the same study, we addressed internal model refinement during skill acquisition. Motor learning literature indicates that refining internal models involves the cerebellum, whose activity decreases with learning progress. The cerebellar activity also reflects the acquired internal model (Imamizu et al., 2000). Various cortical regions are also involved, depending on the type of action whose representations we are modeling. Once the model is acquired, motor tuning leads to smooth, automatized performance. This achievement involves both the cerebellum (Lisberger, 2010) and the basal ganglia (Hikosaka, 1991). These two regions are functionally interconnected, forming loops with the cerebral cortex. While these loops are known to exist, there is little evidence on their creation. Indeed, it remains unclear how they emerge when an internal model is first created, at the beginning of learning, and how they evolve when the model becomes more refined as learning progresses. In Chapter 4, we report the analysis of resting state fMRI data before and after participants performed the syllable learning task, aiming to shed light on the evolution of the functional brain networks

1. Introduction – 1.5. The objectives of this dissertation

supporting internal model refinement over the course of learning. Specifically, we examined changes in functional connectivity between a set of striatal and cerebellar regions that could be related to learning progress and, therefore, to the refinement of the internal models of syllables as their pronunciation is being learned.

2. Domain-General Cognitive Subcomponents of Motor Sequence Learning: a Coordinate-Based Meta-Analysis

Extensive studies of motor learning have revealed the engagement of the premotor cortex, the basal ganglia and the cerebellum in motor learning, and hinted at the importance of the experimental task and baseline in identifying specific brain regions. A neurobiological model of motor sequence learning [Doyon et al., 2018] suggested that the regions involved in the early stages depend on the cognitive processes required during the initial practice. Aiming to specify the cognitive processes involved in the initial learning that are isolated by focusing on specific tasks and baseline conditions in motor sequence learning studies, we performed a series of activation likelihood estimation (ALE) coordinate-based meta-analyses of 48 neuroimaging experiments on motor sequence learning testing 640 healthy non-elderly adult human participants, grouping them by individual tasks and baselines. We observed a global network consisting of frontal and parietal lobes, the thalamus, and the cerebellum. Frontal cortical regions were linked with individual movement planning, old and new sequence representation and attention. Activations in the parietal cortex were related to visuomotor integration and possibly internal model refinement. The cerebellar activity patterns suggest roles in error detection, with a further distinction between complexity processing and online motor tuning in its anterior and posterior portions respectively. The basal ganglia were observed only when a particular control condition was used, that consisted of randomly ordered sequences. The differences in the results when focusing on individual tasks or baselines gave clear indications that the choice of the experimental design favors isolating some cognitive processes and related brain regions over others. The proposed cognitive roles that specific brain regions have in motor sequence learning could be used to further refine the existing neurocognitive models of motor sequence learning.

Keywords: motor sequence learning, activation likelihood estimation, cognitive subcomponents of learning, experimental design

2.1. Introduction

Acknowledging that the learning strategies at place in the beginning of motor sequence learning can recruit different cognitive processes, Doyon and colleagues proposed a neurobiological model of motor sequence learning accounting for multiple brain circuits (Doyon et al., 2002; Doyon & Benali, 2005; Doyon et al., 2009; Doyon et al., 2018). Several other cognitive models of motor sequence learning and related processes, such as motor control, are based on commonly used tasks and acknowledge that differences observed in motor sequence learning studies can be task-related. For instance, in their account of motor selection and execution in motor learning, Diedrichsen and Kornysheva (2015) argue that in both serial reaction time task and discrete sequence production task learning starts at the selection level, when one learns to predict the next element of the sequence. In other tasks, such as visuomotor tracking or more complex movement execution, learning would mainly involve the execution level (Diedrichsen & Kornysheva, 2015). Hikosaka and colleagues (2002) propose that, in motor skill learning, encoding is double. On one hand, one encodes visuospatial coordinates of the movement, which would be sustained by frontoparietal cortical areas and associative parts of the basal ganglia and the cerebellum. On the other hand, one encodes motor programs, sustained by motor cortical areas and motor regions of basal ganglia and the cerebellum. While visuospatial coordinates are usually encoded before the motor program, the order is reversed in case of the SRTT (Luft & Buitrago, 2005), making task-dependent predictions about the neural pattern of motor sequence learning.

In addition to choosing a task that isolates the targeted cognitive components of motor sequence learning, a critical issue in functional neuroimaging studies in general is the choice of the control condition. Indeed, the baseline choice can explain seemingly contradictory results of otherwise identical functional neuroimaging studies (Brandt, 2006). Once the phenomenon we want to investigate has been clearly identified, an appropriate baseline should account for as many confounds as possible, therefore isolating only the neural activity related to the targeted phenomenon (Soares et al., 2016). It is recommended to make control conditions as similar as possible to the task condition (Brandt, 2006). Using multiple baseline conditions is another solution (Soares et al., 2016), but this is not always possible due to time constraints. When selecting a minimally different baseline condition for a motor sequence learning study, one can opt for sequences that have already been practiced or learned. This could involve highly automatized tapping movements such as 1-1-1-1 or 1-2-3-4, or

2. Domain-General Cognitive Subcomponents of Motor Sequence Learning: a Coordinate-Based Meta-Analysis – 2.1. Introduction

sequences learned in a previous training session, as well as contrasting the execution of the same sequence in the beginning and at a later stage of learning. Another common choice is using sequences that are novel, but cannot be learned because they are unpredictable and random. Depending on this choice, the control condition could account for cognitive processes such as response selection or motor execution, but not for the creation of motor chunks, that in the case of a known sequence have already been created, and fail to be created in the case of random orders. The baseline can also exclude motor execution, if what we are interested in is motor imagery. Indeed, contrasting motor execution with motor imagery, it has been shown that the cerebellum likely monitors cortical output and provides corrective information only when the movements actually take place, and not during motor imagery (Nair et al., 2003). Manipulating sequence difficulty across the conditions, the same study (Nair et al., 2003) suggested that the superior parietal lobule may have a more important role in tasks that need greater coordination, perhaps through attention and memory resources that are solicited for more difficult sequences. In sum, it is plausible to think that choosing one or another baseline will isolate different cognitive processes during the initial practice (Doyon et al., 2018).

The use of different learning strategies as well as the recruitment of related overlapping cognitive processes renders the results of motor sequence learning studies hard to interpret (Doyon et al., 2018). The choice of a specific experimental design leads to results that may be driven by the specificities of the task and control condition that reflect different cognitive subcomponents of learning. While this could be seen as an inconvenience, we argue that the different experimental designs can be informative as concerns the particular cognitive subcomponents of motor sequence learning and their neural correlates. In contrast to the results from a single study, which do not allow one to examine all aspects of motor sequence learning, a meta-analytic approach could yield a more complete image of this complex process. Indeed, previous meta-analyses of motor sequence learning have identified some task-specific (Hardwick et al., 2013; Witt, Laird & Meyerand, 2008) patterns of activation. Hardwick and colleagues (Hardwick et al., 2013) compared consistent activations across task-specific analyses focusing on sensorimotor tasks and SRTT variants. The two tasks showed stronger recruitment of subcortical regions in the basal ganglia and the cerebellum, and of cortical and thalamic structures respectively. Focusing on finger tapping tasks, another meta-analysis (Witt, Laird & Meyerand, 2008) demonstrated that the use of a visual or auditory pacing stimulus resulted in a more diverse network in comparison to

2. Domain-General Cognitive Subcomponents of Motor Sequence Learning: a Coordinate-Based Meta-Analysis – 2.2. Materials and methods

varying the complexity of the tapping task. Focusing on baseline-specific activations within SRTT, Janacsek and her colleagues showed that choosing random sequences for the baseline recruits exclusively the basal ganglia, suggesting a contribution of cerebellar and premotor regions to aspects of task not directly related to sequence learning (Janacsek et al., 2020). Focusing on the cerebellar activations linked to formation of internal representations of new motor tasks, another meta-analysis found evidence of modularity, observing differential engagement of anterior cerebellar regions for implicit and explicit learning and visuomotor adaptation (Bernard & Seidler, 2013). Other meta-analyses focusing on behavioral data revealed a role of the basal ganglia as indicated by impairment of implicit learning in Parkinson's disease (Siegert et al., 2006; Hayes, Hunsaker & Dibble, 2015).

Here, we conducted a series of quantitative meta-analyses of the existing neuroimaging studies of motor sequence learning in humans, grouped by task and baseline conditions found in these studies. Using the observed commonalities and differences that arise from these meta-analyses, we can specify the cognitive processes involved in the initial learning stages that are isolated by focusing on specific tasks and baseline conditions and their links to the neural structures in motor sequence learning models (Doyon et al., 2002; Doyon & Benali, 2005; Doyon et al., 2009; Doyon et al., 2018). Furthermore, the activations resulting from learning a non-linguistic motor skill could be informative of the domain-general or task-specific nature of motor sequence learning linked to speech.

2.2. Materials and methods

2.2.1. Rationale of the meta-analytic approach

A coordinate-based meta-analysis approach was chosen to overcome the limitations of neuroimaging studies, namely small sample sizes and the influence of the particular experimental design (Carp, 2012). This approach allows one to identify which brain areas are consistently activated across studies and is therefore relevant to test the existing models of neural correlates of motor sequence learning. Concretely, we used the activation likelihood estimation. In addition to a global meta-analysis, we divided the selected studies by the tasks and baseline conditions used and ran an additional meta-analysis on each subset of the studies. The consistent activations that emerged from these analyses were then used to tentatively link the brain regions

2. *Domain-General Cognitive Subcomponents of Motor Sequence Learning: a Coordinate-Based Meta-Analysis – 2.2. Materials and methods*

involved in initial motor sequence learning according to the existing models (Doyon et al., 2002; Doyon & Benali, 2005; Doyon et al., 2009; Doyon et al., 2018) with their possible cognitive roles, and provide a further confirmation to the model's relevance. After defining the inclusion criteria, we double-checked that selected studies matched all of them, reducing the chance of a selection bias (Müller et al., 2018).

2.2.2. Literature search and study selection

The search was conducted on the 9th of April 2020. We looked for the relevant studies in the PubMed database, a free full-text archive of biomedical and life sciences journal literature (Canese & Weis, 2013), using the following query:

(motor[Title] AND (sequence[Title] OR sequential[Title])) AND (learning[Title] OR acquisition[Title] OR novel[Title] OR new[Title]).

In addition, we used Google Scholar, which was assessed via the version 19.4.0 of Harzing's Publish or Perish software that enables indexing of results and advanced search criteria, using the following query:

("learning" OR "acquisition" OR "acquired" OR "learned" OR "novel") AND "motor" AND ("sequence" OR "sequences" OR "sequential")[Title] AND ("fMRI" OR "PET")[Text] [Google Scholar].

After adding records from the existing meta-analyses on motor sequence learning (Hardwick et al., 2013; Janacsek et al., 2020; Bernard & Seidler, 2013; Siegert et al., 2008) and removing duplicates, we obtained an initial pool of 215 results. From this initial pool, we screened the title and the abstract of each study, and then the full text of the matching studies against the following inclusion criteria (See Figure 2.1 adapted from (Moher et al., 2009) for the detailed study selection process):

1. Studies written in English;
2. Empirical neuroimaging (fMRI or positron emission tomography, PET) studies, excluding preprints and other non peer-reviewed publications, posters, doctoral dissertations reporting data available in published articles, reviews and meta-analyses, which were inspected to identify any additional studies;
3. Studies reporting whole-brain activation coordinates; we excluded studies reporting small volume corrected analyses or analyses based ROI or not reporting any coordinates after contacting the corresponding author to request

2. *Domain-General Cognitive Subcomponents of Motor Sequence Learning: a Coordinate-Based Meta-Analysis – 2.2. Materials and methods*

them;

4. Studies covering the whole brain; we excluded studies in which the cerebellum or any other area was not fully covered in all the participants;
5. Studies focusing on early motor sequence learning; we excluded studies in which initial training and/or consolidation took place before the scanning session;
6. Studies on healthy non-elderly right handed adult humans; we excluded studies on other species, children, adolescents, elderly and pathological populations;
7. Studies involving sequence production; we excluded experiments that used observational paradigms to study motor sequence learning.

The studies that were removed from the initial pool after screening the title and the abstract included studies not using PET or fMRI (46), studies on non-human species (6), meta-analyses (4), not peer reviewed preprints (1), conference posters (1) and doctoral dissertations whose authors did not respond to our request for a copy of the manuscript (1). After assessing the full text of the remaining studies (156), we excluded the articles that did not report whole-brain analysis (42), did not include a non-elderly adult subject group (21), did not cover full brain (12), did not focus on sequence learning (13), did not report stereotaxic coordinates and authors did not provide them upon request (9), did not report within-subject control group results in case of a study on a pathology (8), did not study production in sequence learning (2), or reported data available in other papers we screened (1).

This screening left us with 48 studies reporting 55 experiments on 48 distinct subject groups, for a total of 640 subjects and 1074 foci (See Table 2.1 for details of the studies included). We next screened the selected articles for the relevant tasks and contrasts that could isolate different cognitive subcomponents of motor sequence learning, that could not be defined a priori. Unsurprisingly, the most commonly used task was the SRTT (29 subject groups), enabling a subsequent meta-analysis focusing on this task. The second most common task was the timed motor sequence task (TMST). In this task, the participant must press the button that corresponds to sequences of stimuli on the screen. However, instead of responding as quickly as possible, the participant is required to respond in synchrony with the stimuli as they appear (Steele, 2008). Other tasks included discrete sequence production (DSP), finger-to-thumb opposition sequence task (FOS) and serial interception sequence

2. Domain-General Cognitive Subcomponents of Motor Sequence Learning: a Coordinate-Based Meta-Analysis – 2.2. Materials and methods

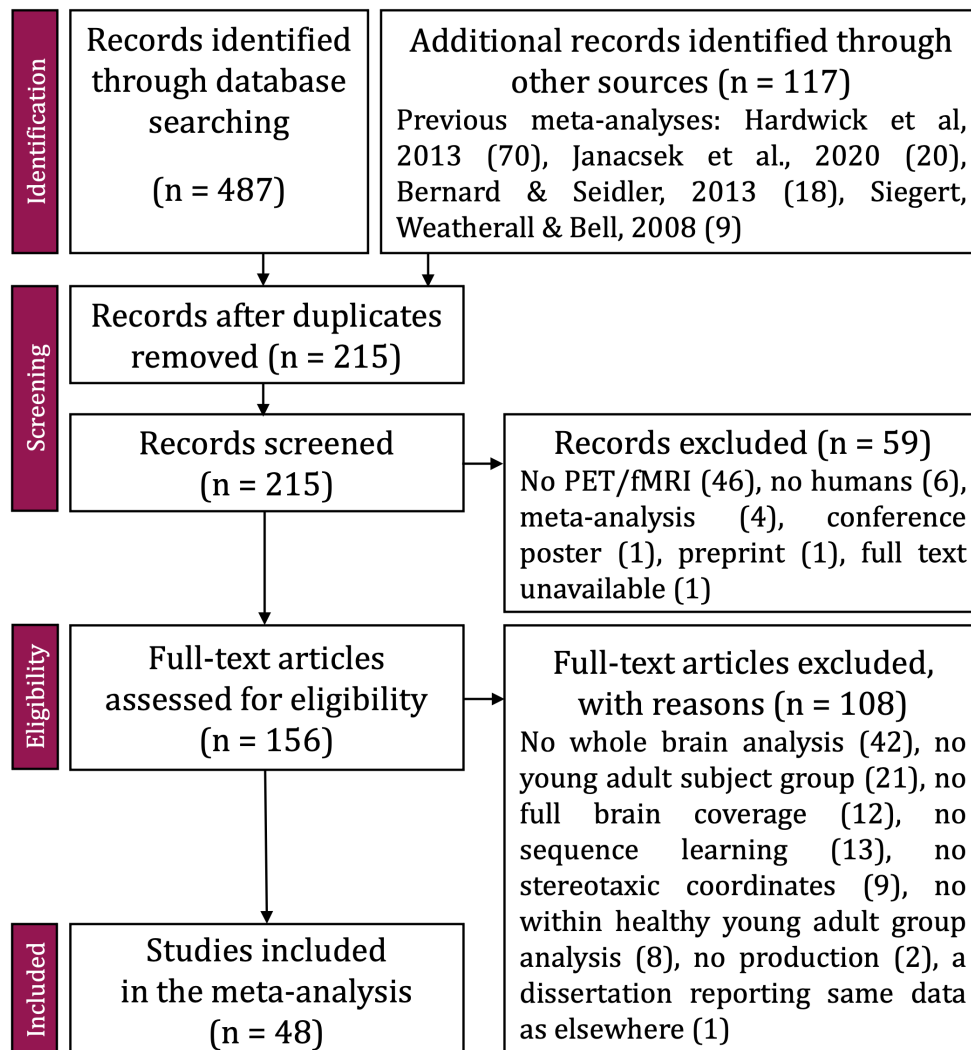


Figure 2.1. – The PRISMA flowchart describing the study selection process. The PRISMA flowchart is adapted from Moher et al., 2009.

2. Domain-General Cognitive Subcomponents of Motor Sequence Learning: a Coordinate-Based Meta-Analysis – 2.2. Materials and methods

learning task (SISL). In DSP, the sequences are clearly separated in time and practiced through a larger number of repetitions: 500 per sequence or more as opposed to 80-100 repetitions in the SRTT (Verwey & Wright, 2004). FOS is based on attributing numbers from 1 to 5 to each finger. If 1 represents the thumb, then the remaining four digits correspond to the remaining fingers and the participant is asked to, when presented with sequences of numbers, execute those sequences by touching the corresponding fingers with their thumb. SISL is similar to the SRTT, but the cues are moving on the screen and the participants must time their motor response to intercept them (Gobel, Parrish & Reber, 2011). These tasks, sufficiently different from the SRTT not to be merged together in a subsequent analysis, included less than 7 subject groups. Therefore, separate subsequent task-specific meta-analyses were not possible either, as it has been suggested that at least 17 subject groups are necessary to detect effects present in about a third of the population with a power of .8 (Eickhoff et al., 2016).

Examination of different baselines showed that a motor baseline was used in 25 subject groups. Specifically, a random sequence was used as the control condition in 12 subject groups; in 11 groups, the control condition was either a known sequence or a previously practiced sequence, or the initial learning was contrasted with a later stage of learning of the same sequence². These sample sizes allow to detect an effect size of .45 and .42 respectively (Eickhoff et al., 2016). As our aim was to tentatively refine the existing neurocognitive models of motor sequence learning and generate hypotheses that will guide future research, we decided to perform separate meta-analyses on these two baseline groups despite the smaller effect detectability.

2. Domain-General Cognitive Subcomponents of Motor Sequence Learning: a
Coordinate-Based Meta-Analysis – 2.2. Materials and methods

Study	Method	N	Task	Motor baseline	Foci
Albouy, 2012	fMRI 3T	30	FTT	/	3
Bischoff-Grethe, 2004	fMRI 1.5T	16	SRTT	/	54
Buccino, 2004	fMRI 1.5T	12	Imitation	/	13
Cross, 2007	fMRI 1.5T	27	SRTT	Practiced	34
Daselaar, 2003	fMRI 1.5T	26	SRTT	Random	21
Debaere, 2004	fMRI 1.5T	20	Coordination	/	27
Dennis, 2011	fMRI 4T	12	SRTT	Random	16
Doyon, 1996	PET	14	SRTT	/	9
Ettinger, 2013	fMRI 1.5T	26	SRTT	Random	5
Fischer, 2005	fMRI 1.5T	8	FOS	/	21
Floyer-Lea, 2005	fMRI 3T	22	Target tracking	Practiced	9
Gabitov, 2016	fMRI 3T	15	FOS	Practiced	13
Gobel, 2011	fMRI 3T	18	SISL	Practiced	9
Gonzalez, 2018	fMRI 3T	12	Target tracking	Random	11
Grafton, 1995	PET	12	SRTT	/	26
Grafton, 1998	PET	10	SRTT	/	8
Grafton, 2001	PET	6	Target tracking	/	10
Grafton, 2002	PET	8	SRTT	/	31
Haaland, 2004	fMRI 1.5T	14	SRTT	Known	9
Heun, 2004	fMRI 1.5T	10	SRTT	Random FTT	19
Jouen, 2013	fMRI 1.5T	18	DSP	Known	20
Karim, 2017	fMRI 3T	13	FTT	Random	2
Kawashima 1998b	PET	8	Ball rotation	/	18
Kawashima, 1998a	PET	9	Target tracking	/	35
Kincses, 2008	fMRI 3T	15	SRTT	/	20
Konovalov, 2018	fMRI 3T	24	SRTT	Previous trial	26
Kumari, 2002	fMRI 1.5T	6	SRTT	Random	14
Lin, 2011	fMRI 3T	16	SRTT	/	10
Muller, 2002	fMRI 1.5T	7	SRTT	Known	19
	fMRI 1.5T	7	SRTT	Known	31
Muller, 2003*	fMRI 1.5T	8	SRTT	Known	17
	fMRI 1.5T	8	SRTT	Known	25
Müller, 2004*	fMRI 1.5T	8	SRTT	Known	29
Naismith, 2010	fMRI 1.5T	20	SRTT	Random	21

2. Domain-General Cognitive Subcomponents of Motor Sequence Learning: a Coordinate-Based Meta-Analysis – 2.2. Materials and methods

Nyberg, 2006	fMRI 1.5T	8	SRTT	/	4
Nyberg, 2007	fMRI 3T	12	SRTT	/	3
Olson, 2006	fMRI 3T	10	SRTT	/	14
Parsons, 2005	fMRI 1.5T	12	SRTT	/	24
Rauch, 1997	PET	9	SRTT	Random	6
Rieckmann, 2010	fMRI 1.5T	14	SRTT	Random	11
Sakai, 2002**	PET	12	SRTT	Random	8
	PET	8	SRTT	Random	11
Sakreida, 2018	fMRI 3T	13	Imitation	Practiced	41
Schubotz, 2001	fMRI 3T	12	Imitation	/	20
Seidler, 2002	fMRI 4T	6	SRTT	/	9
Steele, 2008	fMRI 3T	15	TMST	/	28
	fMRI 3T	15	TMST	Known	63
Steele, 2010	fMRI 3T	15	TMST	Known	16
	fMRI 3T	15	TMST	Practiced	12
	fMRI 3T	15	TMST	/	44
Strangman, 2005	fMRI 3T	9	SRTT	/	36
Toni, 1998	fMRI 2T	3	FTT	/	23
van der Graaf, 2004	fMRI 1.5T	12	SRTT	/	35
	fMRI 1.5T	12	SRTT	NA	20
van der Graaf, 2006	fMRI 1.5T	8	SRTT	Random	11

Table 2.1. – *Studies included in the meta-analyses.* N = number of subjects in the study, FTT = finger tapping task, SRTT = serial reaction time task, FOS = finger to thumb sequence opposition task, SISL = serial interception sequence learning, DSP = discrete sequence production, / = no particular sequence served as the baseline: either the signal change was reported, with or without correlation with performance, or an independent component analysis was used, * = studies using the same dataset, therefore included into further analyses as a single study, ** = two experiments on separate subjects, therefore considered as separate experiments.

2.2.3. Activation likelihood estimation

We performed a series of the activation likelihood estimation (ALE) meta-analysis using the GingerALE software (Laird et al., 2009, 2011; Fox et al., 2013; <http://brainmap.org/ale>).

2. Domain-General Cognitive Subcomponents of Motor Sequence Learning: a Coordinate-Based Meta-Analysis – 2.3. Results

This software takes as the input files lists of foci from each subject group. For each analysis we prepared an input text file with the foci from the included subject groups and the number of participants in each group, after converting all the reported coordinates to the same reference space using the Lancaster's icbm2tal transform (Lancaster et al., 2007). For the experiments that performed a between-subject analysis, if the number of participants differed across the groups we entered the lower number. Foci groups in the input files consisted of subject groups rather than of individual contrasts from the same subjects as GingerALE's algorithm focuses on finding agreement across subject groups in order to minimize the within-group effects (Fox et al., 2013). In other words, if several contrasts from the same subject group were included in an analysis, they were considered as a single group³. After loading the input files to the GingerALE software, we assessed the overlap among the spatial probability distributions of the reported activations through permutation testing with 10000 permutations, allowing to differentiate between true convergence and noise (Nichols & Hayasaka, 2003). Correction for family-wise error rate was applied to the resulting ALE statistics at $p < .05$ on cluster level (Eickhoff et al., 2012).

2.3. Results

As Table 2.2 shows, the global meta-analysis of all of the selected studies and contrasts revealed a wide cortico-subcortical activation, which covered the frontal (precentral gyrus, superior, medial and middle frontal gyri) and the parietal cortex (superior and inferior parietal lobules and precuneus), as well as the cerebellum. The subsequent meta-analysis that focused on the SRTT revealed activations in the frontal cortex (precentral, middle, medial and inferior frontal gyrus) and in the parietal cortex (superior parietal lobule and precuneus).

Regardless of the task used, accounting for the movement execution, i.e. focusing only on the contrasts in which the control condition included movement execution, the meta-analysis revealed consistent activation in the frontal cortex (superior, medial, middle frontal and precentral gyri and the insula) and in the parietal cortex (superior and inferior lobules and precuneus). In addition, activation was found in the anterior cerebellum. Further focusing on the type of the motor sequences used in the control condition, the 12 experiments that accounted for movement execution by using random sequences converged in putamen in the basal ganglia. The remaining 11 experiments that used either a previously practiced or known sequence or compared the

activations during the initial learning with those in its later stages revealed a broader network consisting of the frontal (medial, middle frontal and precentral gyri and the insula) and parietal cortex (superior parietal lobule), thalamus, and the anterior cerebellum. Detailed results of the three meta-analyses concerning the baselines are presented in Table 2.3, while Figure 2.2 shows the results of all the conducted meta-analyses. In addition, foci and converging clusters for each meta-analysis separately are shown in Appendix B.

2.4. Discussion

Motor sequence learning has been extensively studied in neuroimaging experiments through different paradigms and has been accounted for in neurocognitive models. Notably, a neurocognitive model of Doyon and colleagues (Doyon et al., 2002; Doyon & Benali, 2005; Doyon et al., 2009; Doyon et al., 2018) proposed that various brain structures in frontal, medial frontal and parietal cortices, as well as the hippocampus, the striatum, the cerebellum and the spinal cord are involved in the early phase of motor sequence learning. Which particular areas are involved would depend on the cognitive processes required for the initial practice of the task at hand.

In this study, we used a meta-analytic approach to look into how different control conditions and tasks commonly used in motor sequence learning studies that likely isolate different cognitive subcomponents of learning can further refine this model. After selecting the relevant studies that met our inclusion criteria, we focused on the studies in which the control condition accounted for motor execution. These studies were divided in two groups. One group consisted of baselines where sequences of unknown, randomly ordered movements were used. The other group encompassed sequences that were not new to the participant, as they were either already known or they had been practiced prior to or after the fMRI data acquisition. Given the sample sizes and the heterogeneity of the experimental paradigms, another meaningful analysis was also possible by focusing on the most commonly used task in our sample, the SRTT.

Most of the regions that Doyon and colleagues (Doyon et al., 2002; Doyon & Benali, 2005; Doyon et al., 2009; Doyon et al., 2018) account for in their model were detected in our meta-analyses. Namely, we detected frontal and motor cortical areas, the parietal cortex, the basal ganglia and the cerebellum. Two structures that the model accounts for were not detected, namely, the spinal cord and the hippocampus,

2. Domain-General Cognitive Subcomponents of Motor Sequence Learning: a
Coordinate-Based Meta-Analysis – 2.4. Discussion

Anatomical location	Cluster size (mm ³)	x	y	z	ALE max.
Learning > All baselines, all tasks (48 experiments, N = 640, foci = 1074)					
Frontal cortex					
SFG	5048	2	15	48	.041
MFG		-3	-2	54	.035
		-2	-7	63	.023
MidFG	2904	-26	0	52	.033
	2728	30	1	55	.031
PCG	2904	-37	-14	60	.032
	2728	34	-8	58	.031
	43	-18	59	.027	
	2000	-53	5	34	.033
Parietal cortex					
Precuneus	4648	-22	-62	56	.037
		-39	-65	41	.025
SPL		-28	-55	55	.034
IPL	2104	43	-34	50	.032
		43	-45	44	.029
Thalamus	1784	-12	-19	9	.039
Cerebellum					
Anterior	2408	25	-52	-27	.034
		14	-60	-21	.029
Posterior		3	-64	-18	.030
Anterior	1368	-25	-58	-21	.028
		-14	-52	-22	.021
Posterior		-25	-70	-15	.023
Learning > All baselines, SRTT (29 experiments, N = 382, foci = 600)					
Frontal cortex					
MFG	3464	-7	7	55	.024
MidFG		-22	2	49	.023
PCG		-33	-10	57	.022
IFG	912	-53	9	29	.021
		-51	17	22	.019
Parietal cortex					
Precuneus	3184	-18	-64	49	.024
		-37	-66	38	.017
SPL		-33	-58	51	.025

Table 2.2. – Results of global and task-specific meta-analyses. All the results are significant at cluster level of .05. x, y, z = peak coordinates in the MNI space, ALE = activation likelihood estimation, N = number of participants in each subject group, SFG = superior frontal gyrus, MFG = medial frontal gyrus, MidFG = middle frontal gyrus, PCG = precentral gyrus, SPL = superior parietal lobule, IPL = inferior parietal lobule, SRTT = serial reaction time task.

2. Domain-General Cognitive Subcomponents of Motor Sequence Learning: a
Coordinate-Based Meta-Analysis – 2.4. Discussion

Anatomical location	Cluster size (mm ³)	x	y	z	ALE max.
Learning > Movement execution (25 experiments, N = 372, foci = 541)					
Frontal cortex					
SFG	1408	4	13	48	.028
MFG		-7	7	55	.018
MidFG	1016	-26	-2	57	.020
	776	34	1	60	.021
PCG	1016	-26	-8	59	.022
Insula	848	-31	21	-4	.026
Parietal cortex					
Precuneus	2200	-15	-65	47	.019
		-26	-53	62	.016
SPL		-24	-62	58	.023
		-33	-58	51	.022
IPL	936	45	-47	53	.021
		41	-45	48	.020
Cerebellum					
Anterior	769	3	-54	-4	.027
Learning > Known or practiced sequence (11 experiments, N = 172, foci = 359)					
Frontal cortex					
MFG	1432	3	79	19	.021
MidFG	1120	-26	3	-1	.019
PCG		-28	38	-7	.019
Insula	1240	-31	42	20	.025
Parietal cortex					
SPL	1016	-30	5	-53	.020
Thalamus	784	-13	86	-12	.016
		-15	91	-26	.015
		-11	7	-20	.014
		-13	78	-18	.013
Cerebellum					
Anterior	904	3	34	-56	.025
Learning > Random sequence (12 experiments, N = 166, foci = 137)					
Basal ganglia					
Putamen	704	-14	12	-5	.015

Table 2.3. – *Results of baseline-specific meta-analyses.* All the results are significant at cluster level of .05. x, y, z = peak coordinates in the MNI space, ALE = activation likelihood estimation, N = number of participants in each subject group, SFG = superior frontal gyrus, MFG = medial frontal gyrus, MidFG = middle frontal gyrus, PCG = precentral gyrus, SPL = superior parietal lobule, IPL = inferior parietal lobule, SRTT = serial reaction time task.

2. Domain-General Cognitive Subcomponents of Motor Sequence Learning: a
Coordinate-Based Meta-Analysis – 2.4. Discussion

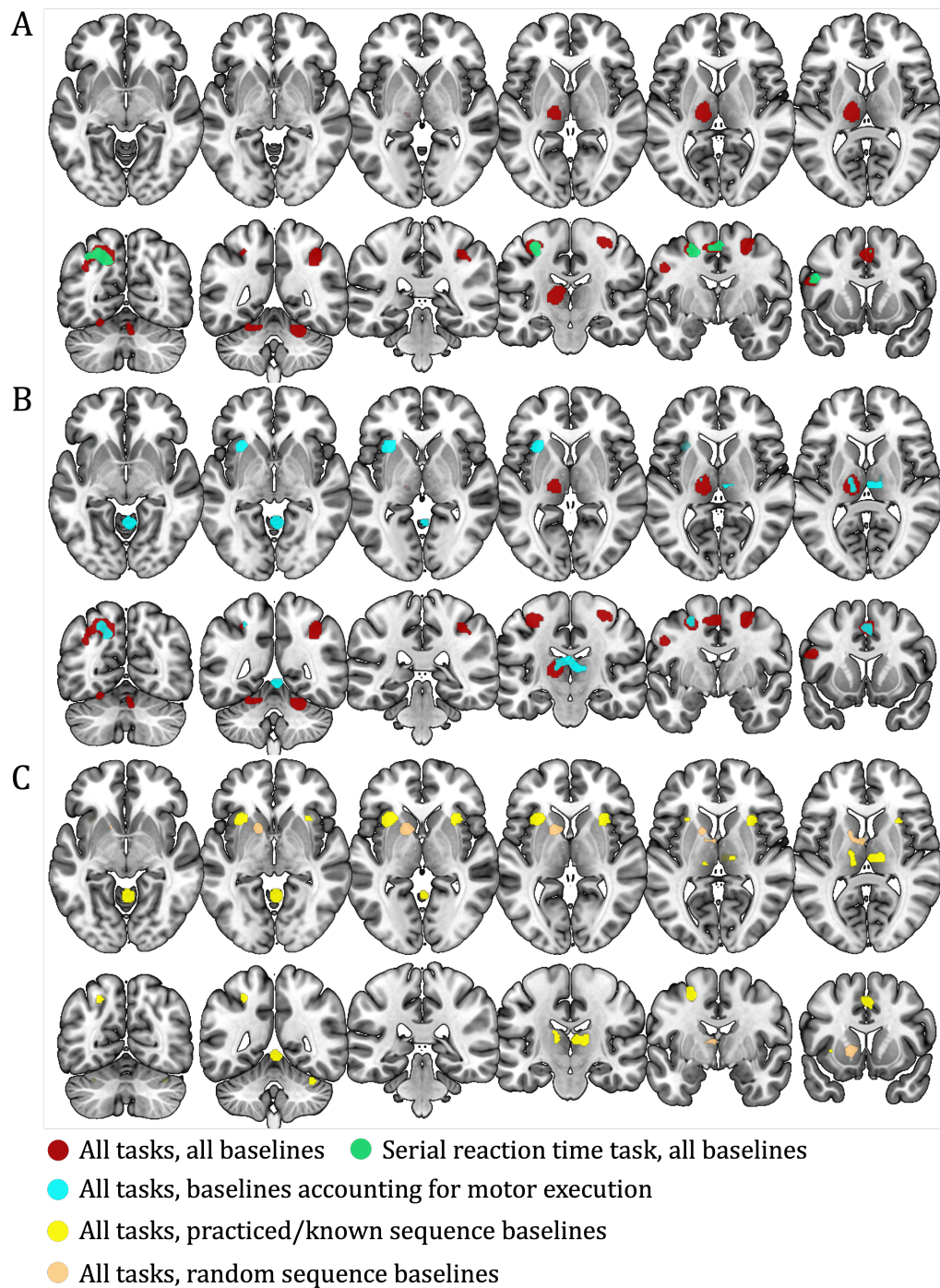


Figure 2.2. – *The results of the five conducted meta-analyses.* A: Results of the global meta-analysis covering all tasks and baselines, and of the SRTT meta-analysis. B: Results of the global meta-analysis, and of the meta-analysis focusing on studies that used baselines accounting for motor execution. C: Results of the subsequent meta-analyses focusing on studies that accounted for motor execution via random and previously known or practiced sequences.

2. Domain-General Cognitive Subcomponents of Motor Sequence Learning: a Coordinate-Based Meta-Analysis – 2.4. Discussion

while the thalamus was detected even though it is not accounted for in the model. The explanation for failing to detect the spinal cord is banal: it was beyond the scope of our inclusion criteria.

2.4.1. Hippocampus

The hippocampus did not converge in any of our meta-analysis, even though it has been reported in previous motor sequence learning studies (Albouy et al., 2015; 2012; 2008; Doyon et al., 2011; Lungu et al., 2014; Schendan et al., 2003; Fernandez-Seara et al., 2009; Gheysen et al., 2010; Ramnani et al., 2000; Rose et al., 2011; Steele & Penhune 2010; see King et al., 2017 for a review) and has its place in the model, according to which it is assumed to show decreasing activation as learning progresses. Hippocampal involvement in motor sequence learning becomes more prominent during the stabilization of the motor performance (Albouy et al., 2012; see Albouy et al., 2013 for a review), and its role in early, initial motor learning is linked to memory trace consolidation that takes place with subsequent sleep (Albouy et al., 2015; 2012; 2008). A plausible explanation of why we have not detected it in our meta-analyses comes, somewhat ironically, from the way we chose our contrasts. To detect the hippocampus, we would need to compare the cerebral activations during initial learning to those during a later learning stage. Our meta-analysis of experiments comparing the cerebral activations when a sequence is being learned to those when a known or practiced sequence is executed did encompass this type of contrast, but was not limited to it. This heterogeneity, combined with the small sample size of this particular meta-analysis, likely resulted in failing to detect the hippocampus.

2.4.2. Thalamus

Detected in our global meta-analysis as well as in the one focusing on the studies with known and previously practiced sequences as the baseline condition, the thalamus is absent from the model of Doyon and colleagues (Doyon et al., 2002; Doyon & Benali, 2005; Doyon et al., 2009; Doyon et al., 2018). It is also absent from our meta-analysis that focused only on the SRTT, while other studies have found that its stimulation improves motor sequence learning through the SRTT in participants with tremor (Terzic et al., 2022). An interesting possibility would be that the thalamus, through its implication in the initial learning, contributes to the consolidation process of motor skills, in a similar fashion to that of the hippocampus. Indeed, changes in

2. Domain-General Cognitive Subcomponents of Motor Sequence Learning: a Coordinate-Based Meta-Analysis – 2.4. Discussion

functional connectivity between the thalamus and the hippocampus show positive correlations with overnight performance improvements that follow initial learning (Boutin et al., 2018). These changes possibly come from spindle oscillations (Boutin et al., 2018), burstlike signals reflecting neuronal oscillations in the thalamus during sleep, known to be modified by experience and to contribute to memory consolidation (Fernandez & Lüthi, 2020).

2.4.3. Motor cortical regions

According to the model, motor cortical regions are constantly involved in the early phases of motor sequence learning. Our meta-analyses are in line with this point of the model, as we detected motor cortical regions, and more specifically, the primary motor cortex / precentral gyrus and the medial frontal gyrus, in all of our meta-analyses except for the one focusing only on the studies using random sequences as the baseline condition. Precentral gyrus could be involved in improvements in temporal precision during motor sequence learning (Wadden et al., 2013). Learning new dance sequences, in which temporal precision is important, induces more plasticity in the precentral gyrus than performing repetitive exercises (Müller et al., 2017). A possible role for medial frontal cortex could be in error monitoring, as suggested by previous studies of performance monitoring in human and non-human primates (see Fu et al., 2023 for a review) and speech error monitoring (Runqvist et al., 2021). Superior frontal gyrus, detected in the global meta-analysis and in the subsequent meta-analysis focusing on the baselines accounting for movement execution, could also support error monitoring (Fu et al., 2023). Another possibility is that it contributes to planning individual movements (Doyon et al., 2009). Both cognitive processes are accounted for when comparing the condition where a sequence is being learned to the one with random sequences, so not detecting superior frontal gyrus in this subsequent meta-analysis is not surprising, but it does not allow us to distinguish between these two potential roles.

2.4.4. Frontal regions

An interesting finding in the frontal cortex was the detection of the inferior frontal gyrus (IFG) exclusively in the meta-analysis focusing on the SRTT. In SRTT, the sequences are often created with second-order conditional transitions between the elements of the sequence. In other words, the combination of two consecutive ele-

2. Domain-General Cognitive Subcomponents of Motor Sequence Learning: a Coordinate-Based Meta-Analysis – 2.4. Discussion

ments predicts the next element. The IFG has been linked precisely to the retention of this type of sequences in a transcranial direct current stimulation experiment (Lum et al., 2018). Therefore, it seems plausible that the IFG would not have a substantial role in motor sequence learning in real life, outside of a (too) well controlled experimental paradigm. If that is the case, this result shows the impact the choice of the experimental task may have on the observed activation and a potential risk of overinterpretation of artefacts as genuine learning-related activity.

A possible role of the middle frontal gyrus in motor sequence learning, given that we have detected it in every baseline except for the random one, could be the one proposed by Dandolo & Schwabe (2019). These authors found that the middle frontal gyrus is particularly important in distinguishing between the old and new sequence representations, after learning had already taken place. With random sequences as a baseline, this distinction is not possible, but it is when the sequence had been already known for a long time, as is the case in a part of the studies included in our meta-analysis of previously known or practiced sequences.

Finally, we detected the insula when focusing on the control conditions accounting for movement execution, and in particular through previously known or practiced sequences, but not through random sequences. This may be because of the role the insula has in attention (MacIntosh et al., 2007). Indeed, using a novel, random sequence as a control condition does not allow to separate learning from attention (MacIntosh et al., 2007). Executing a sequence being learned and a novel, random sequence both require attention more than executing a known sequence. Comparing a condition that requires more attention with the the one where less attention is needed would then isolate insula; comparing two conditions that both require attention, as is the case with the experiments that used randomly ordered sequences as the baseline, would not.

2.4.5. Parietal cortex

In line with the model of Doyon and colleagues (Doyon et al., 2002; Doyon & Benali, 2005; Doyon et al., 2009; Doyon et al., 2018) that attributes a constant role to the parietal cortex in early motor sequence learning, our global meta-analysis detected a general activation of the superior and inferior parietal lobules and the precuneus when all tasks and baselines are collapsed. A likely role of the parietal cortex in motor sequence learning is visual-sensorimotor integration (Doyon et al., 2002, Hardwick et al., 2013). The absence of the precuneus and the inferior parietal lobule in the

2. Domain-General Cognitive Subcomponents of Motor Sequence Learning: a Coordinate-Based Meta-Analysis – 2.4. Discussion

two subsequent meta-analyses, focusing on the studies using random and practiced sequences as the control condition, that also had less statistical power, does not allow us to further distinguish these regions' role. Nevertheless, the superior parietal lobule was still detected in the meta-analysis in which the experiments contained a previously known or practiced sequence as the control condition. The meta-analysis of Hardwick and colleagues (2013) indicated that the role of this region in motor learning may be the integration of visual and proprioceptive information. In line with this, Travers and her colleagues (2014) showed that superior parietal lobule is less engaged in motor sequence learning when the individual's reliance on proprioceptive information is atypical, i.e., in autism spectrum disorder (Haswell et al., 2009; Izawa et al., 2012). As performing a previously practiced or already learned motor sequence requires less proprioception than performing a novel one, detecting the superior parietal lobule in this particular meta-analysis goes in the direction of these findings.

2.4.6. Basal ganglia

Detecting putamen only when the learning condition was contrasted with random motor sequences was a nice sanity check, as this finding is in line with the meta-analysis of Janacsek and her colleagues (2020), in which they focused on this particular baseline. Other previous studies indicated that the putamen is involved in forming predictive associations between the chunks of the sequence (Tzvi et al., 2015; Penhune & Steele, 2012), and in stimulus-response association (Graybiel et al., 1994). Indeed, random sequences are the only baseline condition in which it is impossible to form predictive associations, so comparing new, learnable sequences to them could indeed be isolating this cognitive subcomponent of learning.

2.4.7. Cerebellum

Unsurprisingly, the cerebellum was detected in several of our meta-analyses, as expected from the Doyon and colleagues' model (Doyon et al., 2002; Doyon & Benali, 2005; Doyon et al., 2009; Doyon et al., 2018) and its established role in motor learning (Ito, 2000; Glickstein, 1992; Llinás & Welsh, 1993; Houk, Buckingham & Barto, 1996; de Zeeuw & Ten Brinke, 2015). It was found both in the global analysis, as well as when motor execution was accounted for, and in particular when using previously practiced or known sequences as the control condition. A more detailed look into the detected activations indicates a role of the cerebellar vermis in the meta-analysis on studies

2. Domain-General Cognitive Subcomponents of Motor Sequence Learning: a Coordinate-Based Meta-Analysis – 2.5. Conclusion

that accounted for movement execution and in particular when the activation during the execution of the sequence that is being learned is contrasted with the one of a previously practiced or known sequence. This could be explained by the suggested role of the vermis during motor execution errors (Ito, 2013). Previously practiced and known sequences are less likely to induce motor execution errors, so they would not account for this process when compared to the sequences that are just being learned. Another possible role of the vermis is in processing sequence complexity (Boecker et al., 2002). This account is in line with our finding, as the movements used in the control conditions of the experiments in this meta-analysis included less complex sequences, such as 1-1-1-1 or 1-2-3-4. This contrast could therefore reasonably isolate both error detection and processing of complex sequences. Interestingly, the lobule VI was part of the detected cerebellar activation in the global meta-analysis. It has been related to execution errors monitoring and online tuning of movements in motor sequence learning studies (Grafton et al., 2008; Bernard & Seidler, 2013; Orban et al., 2010).

2.5. Conclusion

According to existing neurocognitive models of motor sequence learning (Doyon et al., 2002; Doyon & Benali, 2005; Doyon et al., 2009; Doyon et al., 2018), neural underpinnings sustaining early phases of learning depend on the cognitive processes at play during the initial practice. In this study, we argued that the choice of a specific task and control condition could determine which learning aspects, and therefore brain regions sustaining them, could be detected in a particular study. We tested this hypothesis through a series of meta-analyses of 48 previous studies of motor sequence learning, focusing on the most commonly used task, SRTT, and control conditions – a randomly ordered sequence or a previously known or practiced sequence. Several implications can be drawn from our results, as shown in Figure 2.3, where we propose a tentative further specification of the model of Doyon and colleagues (Doyon et al., 2002; Doyon & Benali, 2005; Doyon et al., 2009; Doyon et al., 2018). First, four out of five of our meta-analyses confirmed that motor cortical regions are involved in the initial motor sequence learning. A possible role could be suggested for the precentral gyrus in temporal precision improvement; the medial frontal cortex, and maybe the superior frontal gyrus, could contribute to error monitoring. The latter may also support planning individual movements. Second, the roles we propose for

2. Domain-General Cognitive Subcomponents of Motor Sequence Learning: a Coordinate-Based Meta-Analysis – 2.5. Conclusion

the frontal regions in early motor sequence learning, namely, distinguishing old and new sequences (middle frontal gyrus) and attention (insula) are both likely to be less important as learning takes place, in line with the model of Doyon and colleagues (Doyon et al., 2002; Doyon & Benali, 2005; Doyon et al., 2009; Doyon et al., 2018), in which frontal regions are involved in the beginning of learning but have a decreasingly important role. Third, the parietal cortex is likely the place of visuomotor integration, and more precisely of visuo-proprioceptive integration in the superior parietal lobule. Fourth, the basal ganglia, and specifically putamen, could be linked to the formation of predictive associations between chunks. And fifth, while cerebellar contribution may be mainly through error monitoring, the cerebellar vermis could have a role in sequence complexity processing. Besides these specifications, we also recommend to further specify the role of the hippocampus, and acknowledge the one of the thalamus in the model of Doyon and colleagues (Doyon et al., 2002; Doyon & Benali, 2005; Doyon et al., 2009; Doyon et al., 2018). Although the meta-analyses we conducted could not confirm, given our choice of contrasts, that (potentially combined) involvement of these two areas in initial learning may predict subsequent consolidation, this seems to be the case based on the previous studies and should therefore find its place in the model. A final, general consideration is that of the importance of carefully considering the most appropriate experimental design for each research question, to ensure it would not bias the results towards a specific cognitive process while dissimulating the role of other, equally important ones.

2. Domain-General Cognitive Subcomponents of Motor Sequence Learning: a Coordinate-Based Meta-Analysis – 2.5. Conclusion

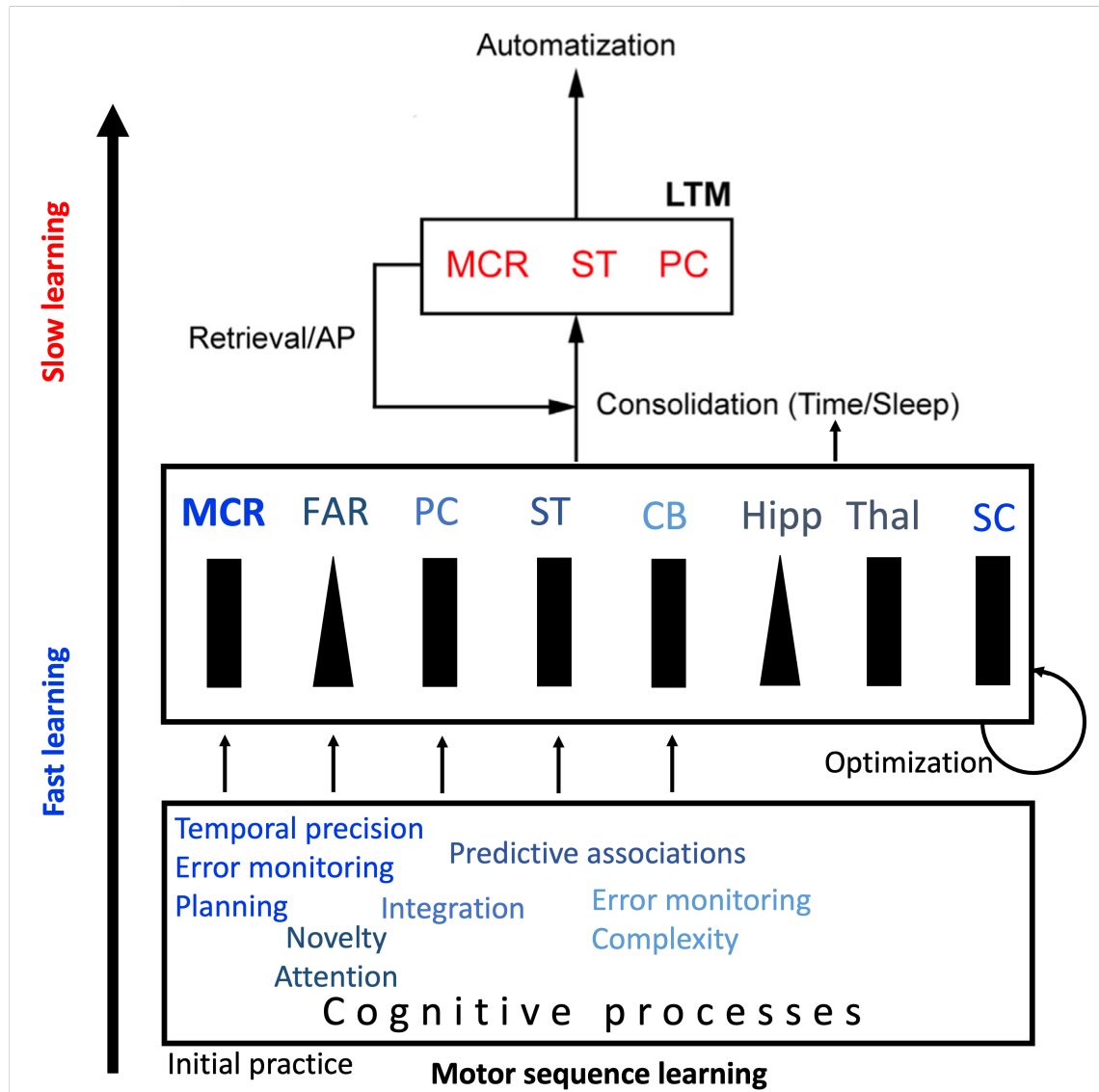


Figure 2.3. – Tentative specification of the cognitive processes in the neurobiological model of brain plasticity associated with motor sequence learning developed by Doyon and colleagues. Hipp = hippocampus, MCR = motor cortical regions, ST = striatum, CB = cerebellum, PC = parietal cortex, FAR = frontal associative regions, SC = spinal cord, AS = associative striatum, SS = sensorimotor striatum, CC = cerebellar cortex, CN = cerebellar nuclei, LTM = long-term memory, AP = additional practice, ▲ = the region is decreasingly involved, ▼ = the region is increasingly involved. Adapted from Doyon et al., 2018.

3. Motor and Cognitive Aspects of Performance Monitoring During Novel Speech Sequence Production: an fMRI Study

Based on: Todorović, S., Anton, J.L., Sein, N., Nazarian, B., Chanoine, V., Rauchbauer, B., Kotz, S. A. and Runnqvist, E. Cortico-cerebellar monitoring of speech sequence production. *Neurobiology of Language (SI: The role of the cerebellum in language comprehension and production)*. Under review.

In a functional magnetic resonance imaging study, we examined the transmodal or sensorimotor driven nature of internal and external speech error monitoring in the cerebellum, the medial frontal cortex, and the temporo-parietal cortex while manipulating an articulatory variable, phonotactic probability. Participants learned to pronounce novel speech sequences that required a high or low monitoring load dependent on their phonotactic legality. Engagement of a cognitive cerebellar region, the crus I, suggests that this region monitors performance in a domain-general, transmodal manner. The activation of cerebellar motor regions, in the superior medial cerebellum and lobules VI and VIII indicates additional presence of sensorimotor-driven, possibly hierarchically subordinate control. The combined pattern of the preSMA and the ACC activations suggests sensorimotor-driven feedback monitoring in the medial frontal cortex, via proprioceptive and auditory feedback through overt errors. Finally, differential temporal and parietal cortical activation regarding both internal and external monitoring indicates involvement of these regions beyond sensorimotor-driven feedback, in line with speech production models linking these regions with auditory targets and internal modeling-like mechanisms. These results highlight the presence of multiple, possibly hierarchically interdependent mechanisms at the service of optimizing performance during speech production.

Keywords: action monitoring; speech production; internal modeling; cerebellum; medial frontal cortex, temporo-parietal cortex

3.1. Introduction

Speakers monitor their utterances for errors during speech production. This cognitive skill can explain in part why speech is produced so accurately and why errors underlie our ability to learn when we produce speech. Such monitoring can take place internally, while detecting impending errors during speech planning, and externally, while detecting overt errors during articulation. Monitoring can concern both more sensorimotor-related representations, such as articulation, and less sensorimotor-related ones, such as semantics. An inherent assumption of existing speech monitoring models is that both internal and external monitoring rely on the same mechanisms, applied during different processing stages, before and after articulation (e.g., Levelt et al., 1999; Nozari et al., 2011; Pickering & Garrod, 2013). However, several studies have shown that inner speech and articulated speech differ in error patterns (e.g., Oppenheim & Dell, 2008) and velocity of repair after error detection (Nooteboom & Quené, 2017), and display partially different neural correlates of error monitoring (e.g., Okada, Matchin & Hickok, 2018; Hansen et al., 2019; Runnqvist et al., 2021), in areas as diverse as medial frontal, temporal and parietal cortices or the cerebellum. To what extent such patterns are actually resulting from dissociations between inner and overt speech, or rather from representations that are more or less sensorimotor-related, is as of yet unclear. Interestingly, one region that is implicated both in situations taxing internal monitoring, as indexed by different error probabilities, and external monitoring, as indexed by overt errors, is the crus I in the posterior cerebellum (e.g., Runnqvist et al., 2016; 2021). Hence, it is likely that the cerebellum is involved in broader and more continuous error monitoring compared to other brain structures. Consistent with this, the cerebellum is considered central in the internal modeling of self-produced actions (e.g., Blakemore, Frith & Wolpert, 2001; Knolle et al., 2013). Parallel to programming actions, we also predict their sensory consequences, by comparing a copy of the expected signal, an efference copy, with actual sensory consequences of our actions. If the prediction is correct, the efference copy inhibits the expected neural response in the relevant part of somatosensory cortex, which is known as reafference cancellation. Otherwise, an error signal will alert about a mismatch between expected and actual sensory consequences (e.g., McCloskey, 1981; Wolpert et al., 1995). Importantly, such efference copying can take place internally, before an action becomes overt (e.g., Desmurget & Grafton, 2000; Blakemore & Sirigu, 2003), and is found for motor and mental actions, including language (e.g., Ito, 2008; Moberget et al., 2014; Lesage et al., 2017; Argyropoulos, 2016). Some authors have hypothesized that this internal

3. *Motor and Cognitive Aspects of Performance Monitoring During Novel Speech Sequence Production: an fMRI Study* – 3.1. *Introduction*

modeling is being carried out across the whole cerebellum as its unique computation (e.g., Medina & Mauk, 2000; Moberget & Ivry, 2016). This hypothesis was generated indirectly based on the structural and synaptic properties of the cerebellum. In fact, the same type of basic neural circuit is repeated throughout all cerebellar subdivisions, constituting the fundamental functional module of the cerebellum (e.g., Purves et al., 2001). In turn, this suggests that a single, characteristic computation may be common to all the functions it is involved in. A strong candidate for such a unique computation is performance monitoring through internal modeling of self-generated actions (e.g., Ito, 2008; Peterburs & Desmond, 2016). This computation could then be applied to different types of functions or representations, located in different cerebellar compartments. Functional properties of such compartments could reflect the distinction between neural processing of motor and higher order cognitive domains (e.g., Fiez, 1996; Strick, Dum & Fiez, 2009), and functional cerebellar topography that separates motor and cognitive subcompartmentalization (e.g., Middleton & Strick, 1994; Stoodley & Schmahmann, 2010; Buckner, 2013). Furthermore, functional connectivity studies have allowed to discern a double motor, modality-specific cerebellar representation, with a first representation in lobules I-VI and a second one in the lobule VIII, and a triple cognitive representation of several modalities, with a first representation in the lobule VI and the crus I, a second one in the crus II and the lobule VIIB, and a third one in lobules IX and X. Functional differences are expected between these different representations (Guell et al., 2018). Yet another, slightly different division of motor and cognitive compartments emerges when considering cerebellar phylogeny. Namely, phylogenetically older spinocerebellum, comprising the superior medial cerebellum, the vermis and the intermediate parts of the hemispheres is thought to regulate proprioceptive input that might develop and calibrate internal models (e.g., Bosco & Poppele, 2001). More recent cerebrocerebellum, comprising the lateral parts of the hemispheres, crus I and lobule VIII, is thought to engage in predictive internal modeling proper, deployed not only for motor control but also for cognitive control (e.g., Tanaka et al., 2020). Given these divisions of the cerebellum to regions processing motor, modality-specific and higher cognitive, transmodal representations, an interesting question is to explore the relationship between cerebellar anatomy and motor and higher cognitive monitoring of a phenomenon that requires both; namely, the speech production. A first hypothesis (Figure 3.1, top panel) concerning cerebellar involvement in speech production monitoring is that different compartments, motor and non-motor, might be involved in monitoring motor and non-motor aspects of

3. *Motor and Cognitive Aspects of Performance Monitoring During Novel Speech Sequence Production: an fMRI Study* – 3.1. *Introduction*

speech respectively. In the case of such compartment specific monitoring, manipulating an articulatory motor variable such as phonotactic probability should only activate cerebellar motor compartments, such as the superior medial cerebellum and the lobule VIII, but not the crus I. Integrating phylogenetic considerations into this hypothesis, one might expect a dissociation between the superior medial cerebellum and the lobule VIII. A second hypothesis concerning cerebellar involvement in speech production monitoring is that supposed transmodal parts of the cerebellum, such as crus I, engage in the monitoring of all types of language-related actions, whether motor or cognitive in nature, reflecting a hierarchically superior transmodal type of control. In this case, the crus I should be implicated regardless of the monitoring type and level of linguistic representation.

A second goal of this study was to further elaborate on the role of the medial frontal cortex in speech monitoring. Several studies have indeed observed activation of the medial frontal cortex in external monitoring as indexed by overt errors. Notably, both the preSMA and the ACC show a differential activation in such circumstances (e.g., Gauvin et al., 2016; Runnqvist et al., 2021). One way to account for this pattern is the concept of conflict-based monitoring (e.g., Nozari et al., 2011; Gauvin & Hartsuiker, 2020). In this framework, conflict itself triggers monitoring in the medial frontal cortex, that, in turn, relays the information to the prefrontal cortex to regulate behavior. The detection of cognitive conflict (e.g., de Zubicaray et al., 1998; Gauvin et al., 2016; Nozari et al., 2011; Riès et al., 2010) would imply the involvement of the medial frontal cortex in the two monitoring types, as conflict is present in both. This would support the transmodal monitoring hypothesis. However, activation of the preSMA in speech production has also been linked to motor-related difficulties, suggesting its role in early stages of motor sequence learning related to the planning and execution of complex speech sequences (Alario et al., 2006). Linking this evidence to that of the preSMA in overt error processing, a plausible hypothesis is that this region is involved in monitoring articulatory, sensorimotor-driven aspects of production. When manipulating an articulatory, sensorimotor variable, this account also predicts the engagement of the preSMA for internal monitoring, in situations of high articulatory error probability, and external monitoring, for overt articulatory errors. Regarding the role of the ACC in speech monitoring, an alternative to conflict monitoring could be a sensorimotor-driven hypothesis, according to which monitoring of feedback provided via the overt errors (Loh et al., 2020; Runnqvist et al., 2021) would differentially engage the ACC in external monitoring only. The two hypotheses about the performance

3. *Motor and Cognitive Aspects of Performance Monitoring During Novel Speech Sequence Production: an fMRI Study* – 3.1. Introduction

monitoring in the medial frontal cortex are illustrated in the middle panel of Figure 3.1.

A third goal of the current study was to inform an ongoing debate about the involvement of temporal and temporo-parietal regions in speech monitoring (e.g., Meekings & Scott, 2021). Traditionally, the detection of speech production errors was considered to exclusively relate to sensory feedback (e.g., Levelt et al., 1999; Hartsuiker & Kolk, 2001; Roelofs, 2020). Either phonologically encoded inner speech or articulated and audible speech would feed into the same speech comprehension loop, leading to a conceptualizer where the appropriateness of an utterance would be monitored. That is, speakers would detect errors by listening to and understanding their own inner and overt speech, similarly to how they listen to and understand someone else's speech. Because of the central role of speech comprehension to this theory, its hypothesized neural basis is the pSTG. More recent models of speech production also integrate the temporal (pSTG) and temporoparietal (Spt) cortex as the location of auditory target and error maps. In these models, mechanisms akin to internal modeling have been implemented as an error signal generated through a mismatch between a (somato-)sensory target and a motor command (e.g., Tourville & Guenther, 2011; Guenther, Ghosh & Tourville, 2006; Hickok, 2012; Hickok, 2014; Pickering & Garrod, 2013). Furthermore, in these models the cerebellum is hypothesized to contribute to sensory predictions that form the auditory and somatosensory targets. Hence, these models would also predict pSTG or Spt activation during internal and external monitoring, arguably in combination with cerebellar activation. Indeed, there is converging evidence that the pSTG engages in external monitoring as indexed by articulated, though not necessarily audible speech, overt speech errors, and halted speech (e.g., Hansen et al., 2019; Okada et al., 2018; Runnqvist et al., 2021). Note that some studies manipulating auditory feedback (e.g., Hirano et al., 1997) or looking into auditory hallucinations (e.g., Shergill et al., 2000) also link pSTG to internal monitoring (e.g., Indefrey & Levelt, 2004). Nonetheless, these studies varied in their use of auditory variables, and it is unclear whether the pSTG activation is actually reflecting internal monitoring, especially as studies explicitly controlling for auditory variables do not seem to observe pSTG in conditions taxing internal monitoring (e.g., Hansen et al. 2019; Runnqvist et al., 2021). The two hypotheses are illustrated in the bottom panel of Figure 3.1.

To test our hypotheses, we targeted the postlexical process of syllabification in a speech production paradigm. Concretely, we manipulated an articulatory variable, namely phonotactic legality, while participants learned to pronounce novel sound

3. Motor and Cognitive Aspects of Performance Monitoring During Novel Speech Sequence Production: an fMRI Study – 3.1. Introduction

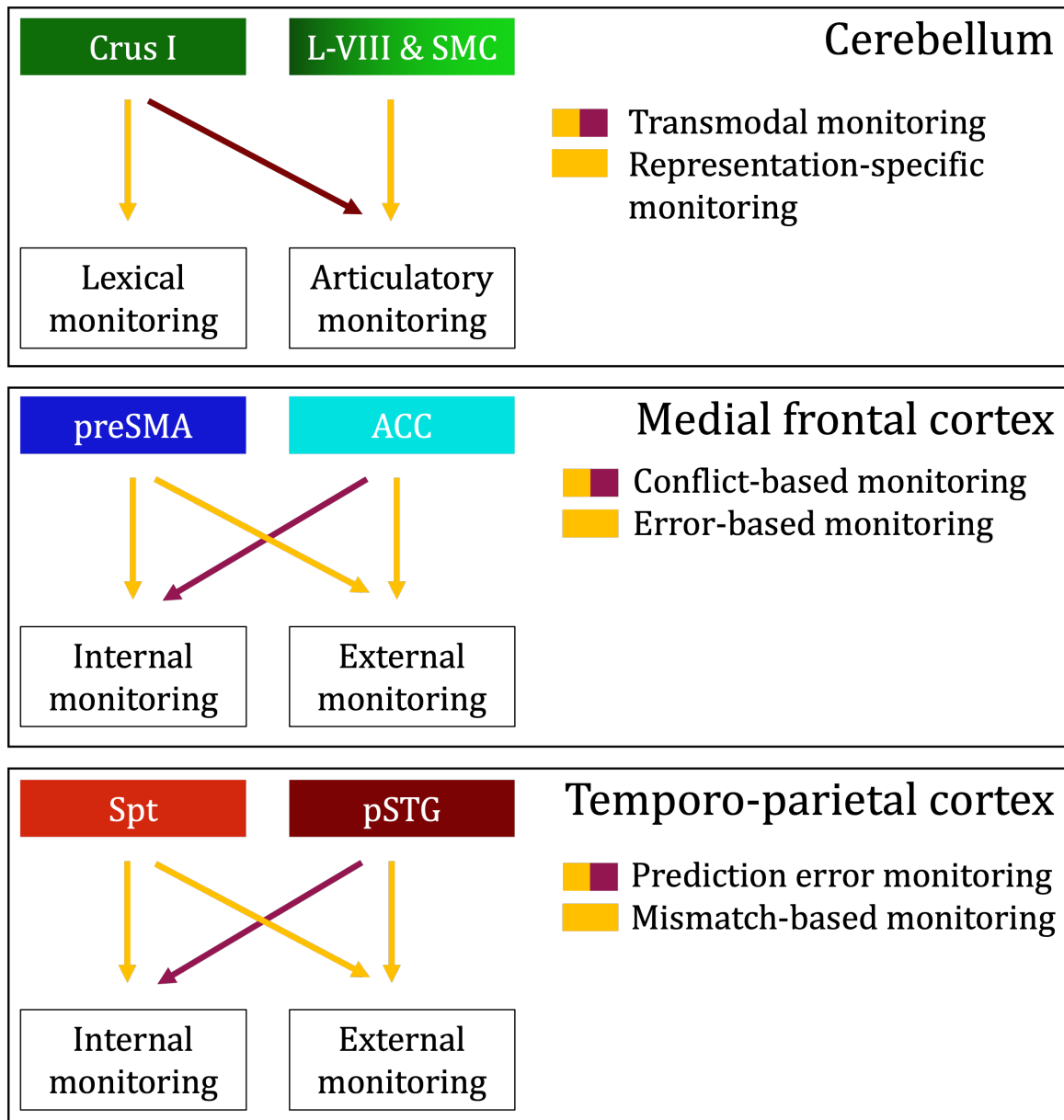


Figure 3.1. – *The monitoring hypotheses.* L-VIII = cerebellar lobule VIII, SMC = superior medial cerebellum, preSMA = pre-supplementary motor area, ACC = anterior cingulate cortex, Spt = Sylvian fissure at the parietal-temporal boundary, pSTG = posterior superior temporal gyrus.

3. *Motor and Cognitive Aspects of Performance Monitoring During Novel Speech Sequence Production: an fMRI Study* – 3.1. *Introduction*

combinations that were not part of their native French language repertoire. At baseline, participants also produced pseudowords consisting of sound combinations that exist in French. Postlexical internal monitoring was indexed by contrasting correct trials in two conditions of phonotactic legality: illegal, impossible phoneme sequences versus legal, possible phoneme sequences. That is, given that speech production in the illegal, learning condition likely leads to more errors and requires more monitoring than the legal, control condition, a difference between the two necessarily comprises a difference in internal monitoring. External monitoring was indexed as in prior studies (Runnqvist et al., 2021; Gauvin et al., 2016) by contrasting errors with correct production. Note that our internal monitoring contrast likely also involves some degree of external monitoring, and similarly the external monitoring contrast involves some degree of internal monitoring. Importantly, the ratio of each monitoring type should differ as in one case speakers manage to avoid errors despite a high error probability, presumably relying more on internal monitoring, while in the other case they do not, presumably relying more on external monitoring.

The monitoring processes of interest were examined using an ROI approach, focusing on 7 ROI within the cerebellum, medial frontal, temporal and temporo-parietal cortices (see Table 3.1). In the cerebellum, we chose three right hemispheric regions: reportedly transmodal crus I (Guell et al., 2018) and phylogenetically older superior medial cerebellum (SMC), for both of which a role in speech monitoring was suggested by previous studies on a higher cognitive level of linguistic representations (Runnqvist et al., 2021), and the motor, phylogenetically younger posterior lobule VIII, that has been linked to covert speech sequence learning (Rauschecker et al., 2008). In addition, we chose two regions in the medial frontal cortex (left ACC and preSMA, see Runnqvist et al., 2021) and two regions in temporal and temporo-parietal cortices (left pSTG and Spt, see Runnqvist et al., 2021). This selection allows shedding light on potential dissociations across internal and external monitoring for the articulatory variable of interest. Complementing the ROI analyses, we also examined the differential activations related to internal and external monitoring contrasts in a whole brain analysis.

3.2. Methods

3.2.1. Participants

The study received ethical approval (filed under id “2017-A03614-49” at the regional ethical committee “Comité de Protection des Personnes Sud Méditerranée I”). The participants (N = 26, 17 female) were right-handed (Oldfield, 1971) young adult ($26.2 \pm \text{SD } 3.5$) native speakers of French without known neurological or psychiatric medical history or ongoing medical treatment and they participated in the experiment in exchange for monetary compensation. Two participants (1 female, 1 male) failed to complete the experiment until the end and were excluded from further analyses. Participants’ knowledge or previous exposure to foreign languages was controlled for, in order to eliminate the possibility that they are familiar with the consonant clusters we used as phonotactically illegal (see “Materials”).

3.2.2. Materials

Target stimuli were 36 CCVCC syllables (C – consonant, V – vowel) composed of French phonemes. In one half of the stimuli, both consonant clusters were possible in French, and were thus phonotactically legal, so their production required less monitoring, even if the syllable was meaningless. The other half of the stimuli were syllables consisting of novel, phonotactically illegal consonant clusters, requiring increased monitoring as their production had to be learned. Two-consonant clusters were selected from the French language database lexique.org (New et al., 2004) and screened for their frequency of occurrence at the beginning and the end of a syllable (New & Spinelli, 2013). The clusters with high frequency (141.6 ± 251.6) were used to form legal, low monitoring load stimuli, and those with a frequency close to zero ($.3 \pm .8$) to form illegal, high monitoring load ones. The resulting syllables were checked for orthographic neighbors using WordGen software (Duyck et al., 2004). Only one stimulus had one orthographic neighbor (“spald”), the rest of stimuli had none.

All stimuli were recorded by a native Serbian speaker, as Serbian allows pronunciation of all used phoneme combinations. Stimuli were also visually presented according to French orthographic rules. As this experiment was designed as a part of a larger study that took place over two sessions, any given participant was only presented with half of the stimuli, in a counterbalanced manner. Each stimulus was presented 25 times over five experimental runs. The order of the stimuli was pseu-

3. Motor and Cognitive Aspects of Performance Monitoring During Novel Speech Sequence Production: an fMRI Study – 3.2. Methods

dorandomized in eight different lists in the way that the same category (high or low monitoring load) was not presented more than twice in a row and that each stimulus was presented N times before any other stimulus was presented N+1 times. Thus, each participant was presented with 225 (25 x 9) high monitoring load and 225 low monitoring load syllables.

3.2.3. Procedure

Syllables remained on the screen for 798 ms and were synchronously presented auditorily for 660 ms. They were preceded by a fixation cross for 532 ms and followed by a blank screen during which the participants were instructed to pronounce the presented syllable. The time for pronunciation was 1530 ms. The jittered interstimulus interval was 930 to 1464 ms. In sum, on average a single trial lasted 4.059 seconds, and the experimental runs lasted 6 minutes each. Figure 3.2 shows an exemplary trial. Stimulus presentation was controlled by a custom-made software compiled using the LabVIEW development environment (National Instruments) and OptoActive audio system (OptoAcoustics). This software also allowed the recording of vocal productions. Trials were recorded and labeled independently using a multifunction NI PCIe-6353 DAQ (National Instruments) and the FOMRI-III optical noise canceling microphone (OptoAcoustics). Each trial was saved in an independent 16-bit WAV audio file.

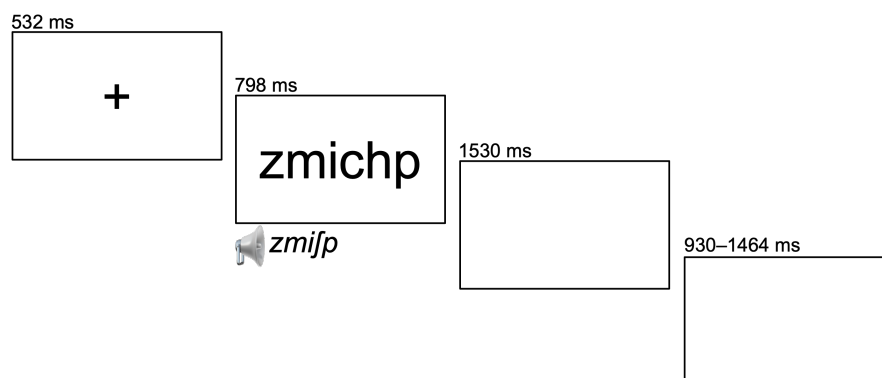


Figure 3.2. – *Depiction of an experimental trial.*

3.2.4. MRI data acquisition

Participants were scanned using a 3 Tesla Siemens Prisma Scanner (Siemens, Erlangen, Germany) at the Marseille MRI Center (Centre IRM-INT@CERIMED, UMR7289

3. *Motor and Cognitive Aspects of Performance Monitoring During Novel Speech Sequence Production: an fMRI Study* – 3.2. *Methods*

CNRS & AMU) using a 64-channel head coil. We first performed a T2-weighted structural scan during which participants were shown a video designed to minimize motion during the fMRI data acquisition (Vanderwal et al., 2015). Functional images covering the whole brain including the cerebellum were acquired during two resting states and the task performance. A resting state scan of 8 minutes during which participants were instructed to lie as still as possible with their eyes open and not to think about anything in particular, while a fixation cross was shown on the screen, was followed by the learning task during approximately 30 minutes. Before the second resting state acquisition, where the same instructions applied, we performed a T1-weighted structural scan during which the same video as before was played. See Figure 3.3 for the illustration of the protocol. Functional images covering the whole brain with the cerebellum were acquired using a multiband BOLD-sensitive gradient echo-planar imaging (EPI) sequence (Moeller et al., 2010) (TR 1224 ms, TE 30 ms, flip angle 66°, 54 slices with thickness 2.5 mm, FOV 210 x 210 mm², matrix 84 x 84, multiband factor 3). Structural images were acquired with a T2-weighted SPACE sequence (voxel size 1 x 1 x 1 mm³, data matrix 256 x 240, TR/TE 3200/408 ms, bandwidth = 725 Hz/pix) and T1-weighted MPRAGE sequence (voxel size 1 x 1 x 1 mm³, data matrix 256 x 240, TR/TI/TE 2300/900/2.98 ms, flip angle 9°, bandwidth = 240 Hz/Pix). Prior to fMRI image, a pair of spin-echo echo-planar images with opposite phase encoding directions and with matched geometric dimensions (resolution, FOV, echo spacing) with fMRI images was used for EPI related distortion correction. Participants' movements during data acquisition were controlled using Framewise Integrated Real Time MRI Monitoring (Dosenbach et al., 2017).

3.2.5. Behavioral data analyses

Commonly used indices of motor sequence learning, including in speech production, are error rates, reaction times and performance durations (Segawa et al., 2015). Evaluating the accuracy of the participants' production was based on the transcription of their utterances. Following the criteria used in Segawa's study (Segawa et al., 2015), syllable productions were labeled as incorrect if they contained insertions, omissions, hesitations, self-reparations, if they were impartial or missing, and if a syllable was pronounced as two or more syllables; they were labeled as correct in all other cases. The reaction times were considered less informative for our experimental paradigm, as the participants were presented with the syllable, in an auditory and written form, before they had to pronounce it, and were instructed not to speak until the

3. Motor and Cognitive Aspects of Performance Monitoring During Novel Speech Sequence Production: an fMRI Study – 3.2. Methods

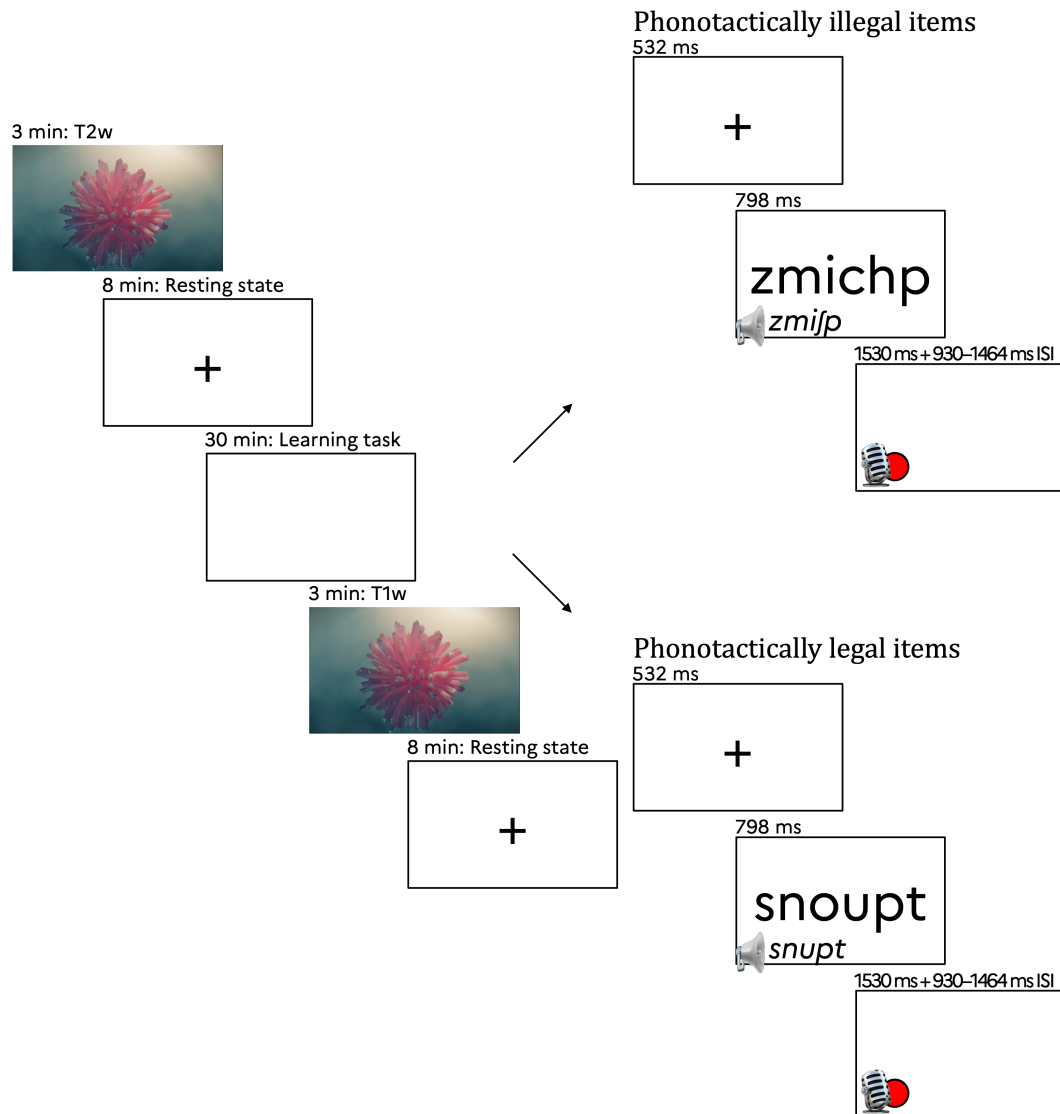


Figure 3.3. – *Experimental procedure*. During the first part of the anatomical acquisition (T2w), participants watched a movie followed by a first resting state acquisition with their eyes open. They then performed the learning task, articulation of novel syllables according to the model presented in written and auditory form; the syllables either follow a structure that is familiar to the participant (phonotactically legal, low monitoring load condition), or are formed in a way that is novel to the participant (phonotactically illegal, high monitoring load learning condition). Finally, the same movie was presented during the second part of the anatomical acquisition (T1w) and followed by a second resting state acquisition.

3. *Motor and Cognitive Aspects of Performance Monitoring During Novel Speech Sequence Production: an fMRI Study* – 3.2. *Methods*

syllable disappears from the screen. Reaction times were also shown of no interest in a study using a very similar paradigm to ours (Segawa et al., 2015). As for the utterance duration, a technical limitation arising from the noise of the fMRI makes it hard to determine when precisely each utterance ends. A single rater, myself, transcribed and evaluated all data. This was judged as reasonable given that I have received university level training in phonetics including phonetic transcription and annotation, and that all the used materials are readily pronounceable for a native speaker of Serbian.

Categorization of each utterance as correct or incorrect allowed us to create three regressors that were used to differentiate between internal and external monitoring, namely, phonotactically illegal, high monitoring load correctly pronounced utterances, phonotactically legal, low monitoring load correctly pronounced utterances, and incorrectly pronounced utterances regardless of their phonotactic legality. Indeed, given that more monitoring is required for actions that have a higher probability of committing an error, one way to study monitoring in an experiment is to manipulate the difficulty of performing the action. For instance, pronouncing syllables that violate the phonotactic constraints of the languages that one is familiar with would require more monitoring than pronouncing novel syllables in which phonemes are phonotactically legal in one's native language. In order to index internal monitoring, which happens prior to articulation, we can rely on performance outcome. As successful internal monitoring results in preventing errors, it can be revealed by looking into correctly pronounced syllables, contrasting the difficult, phonotactically illegal ones, that require more monitoring, than the easier, phonotactically legal ones. Successfully avoiding probable errors arguably requires more monitoring than successfully avoiding committing an error that was unlikely to be committed in the first place. In other words, contrasting correctly produced sequences of greater and lower difficulty, such as phonotactically illegal and legal syllables, gives an indication of internal monitoring. External monitoring, based on the sensorimotor feedback provided by one's own speech, could be revealed by contrasting productions containing an error with correct ones, as hearing something different from what we expected arguably triggers more monitoring than hearing what we expected to hear. This reasoning should not be interpreted as a claim that avoiding an error excludes external monitoring, nor that committing one means there was no internal monitoring. Rather, we considered that when an error was avoided, internal monitoring was recruited more than external monitoring. Similarly, hearing oneself commit an error triggers external monitoring to a greater extent than internal monitoring, which occurs prior to articulation

3. Motor and Cognitive Aspects of Performance Monitoring During Novel Speech Sequence Production: an fMRI Study – 3.2. Methods

(Runnqvist et al., 2021).

3.2.6. Neuroimaging data preprocessing

Functional and anatomical preprocessing was performed using fMRIPrep 20.2.0 (Esteban et al., 2018; 2019), which is based on Nipype 1.5.1 (Gorgolewski et al., 2011; 2018). A Singularity image for fMRIPrep was used and it was run on the high performance computing cluster of the Centre de Calcul Intensif of Aix Marseille University. MRIQC was used for image quality control (Esteban et al., 2017) and was run on the same cluster.

3.2.6.1. Anatomical data

The T1-weighted (T1w) image was corrected for intensity non-uniformity (INU) with N4BiasFieldCorrection (Tustison et al., 2010), distributed with ANTs 2.3.3 (Avants et al., 2008), and used as T1w-reference throughout the workflow. The T1w-reference was then skull-stripped with a Nipype implementation of the antsBrainExtraction.sh workflow (from ANTs), using OASIS30ANTs as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using fast (FSL 5.0.9, Zhang, Brady & Smith, 2001). Volume-based spatial normalization to one standard space was performed through nonlinear registration with antsRegistration (ANTs 2.3.3), using brain-extracted versions of both T1w reference and the T1w template. The following template was selected for spatial normalization: ICBM 152 Nonlinear Asymmetrical template version 2009c (Fonov et al., 2009; TemplateFlow ID: MNI152Nlin2009cAsym).

3.2.6.2. Functional data

Same preprocessing steps were applied for the five experimental runs and the two resting state runs of each participant (Figure 3.4). First, a reference volume and its skull-stripped version were generated by aligning and averaging 1 single-band reference (SBref). A B0-non-uniformity map (or fieldmap) was estimated based on two (or more) echo-planar imaging (EPI) references with opposing phase-encoding directions, with 3dQwarp Cox and Hyde (1997) (AFNI 20160207). Based on the estimated susceptibility distortion, a corrected EPI reference was calculated for a more accurate co-registration with the anatomical reference. The BOLD reference was then co-registered to the T1w reference using flirt (FSL 5.0.9, Jenkinson & Smith 2001) with

3. *Motor and Cognitive Aspects of Performance Monitoring During Novel Speech Sequence Production: an fMRI Study* – 3.2. *Methods*

the boundary-based registration (Greve & Fischl, 2009) cost-function. Co-registration was configured with nine degrees of freedom to account for distortions remaining in the BOLD reference. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) were estimated before any spatiotemporal filtering using *mcflirt* (FSL 5.0.9, Jenkinson et al. 2002). The BOLD time-series were resampled onto their original, native space by applying a single, composite transform to correct for head-motion and susceptibility distortions. These resampled BOLD time-series are referred to as preprocessed BOLD in original space, or just preprocessed BOLD. The BOLD time-series were resampled into standard space, generating a preprocessed BOLD run in MNI152Nlin2009cAsym space. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. Several confounding time-series were calculated based on the preprocessed BOLD: framewise displacement (FD), DVARS and three region-wise global signals. FD was computed using two formulations following Power (absolute sum of relative motions, Power et al., 2014) and Jenkinson (relative root mean square displacement between affines, Jenkinson et al., 2002). FD and DVARS are calculated for each functional run, both using their implementations in *Nipype* (following the definitions by Power et al., 2014). The three global signals were extracted within the CSF, the WM, and the whole-brain masks. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction (*CompCor*, Behzadi et al. 2007). Principal components were estimated after high-pass filtering the preprocessed BOLD time-series (using a discrete cosine filter with 128s cut-off) for anatomical *CompCor* (*aCompCor*). For *aCompCor*, two probabilistic masks (CSF, WM) were generated in anatomical space. The implementation differs from that of Behzadi et al (2007). Instead of eroding the masks by 2 pixels on BOLD space, the *aCompCor* masks were subtracted from a mask of pixels that likely contain a volume fraction of GM. This mask was obtained by thresholding the corresponding partial volume map at .05, and it ensured components were not extracted from voxels containing a minimal fraction of GM. Finally, these masks were resampled into BOLD space and binarized by thresholding at .99 (as in the original implementation). For each *CompCor* decomposition, the 12 components with the largest singular values were retained. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confounded time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each (Satterthwaite et al. 2013). All resamplings

3. *Motor and Cognitive Aspects of Performance Monitoring During Novel Speech Sequence Production: an fMRI Study* – 3.2. *Methods*

were performed with a single interpolation step by composing all the pertinent transformations (i.e., head-motion transform matrices, susceptibility distortion correction, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using `antsApplyTransforms` (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels (Lanczos, 1964).

The preprocessed data were imported to and analyzed using the Statistical Parametric Mapping software (SPM12, <http://www.fil.ion.ucl.ac.uk/spm/software/spm12/>) in MATLAB R2018b (Mathworks Inc., Natick, MA). Smoothing was performed with an isotropic Gaussian kernel (full-width at half-maximum = 5 mm). Regressors of no interest included 24 head movement regressors, the mean signal of the CSF and WM mask, and 24 `aCompCor` regressors related to cerebrospinal fluid and white matter.

3.2.7. Whole brain analysis

For the univariate whole brain analysis, we created a general linear model for each participant. The GLM included, for each of the 5 runs, three regressors of interest corresponding to the occurrences of inaccurate productions (Inc), and accurate productions further divided by the monitoring load condition (HighCor and LowCor). Regressors of interest were convolved with the canonical hemodynamic response function. Functional data were filtered with a 128 s high-pass filter. For each subject, we estimated two contrasts:

External monitoring: Inc - (HighCor + LowCor) [2 -1 -1]

Internal monitoring: HighCor - LowCor [0 1 -1]

Finally, we performed the group level analysis. All statistical comparisons were performed with a voxelwise threshold of $p < .001$ and a cluster extent threshold of 25 voxels.

3.2.8. Region of interest analysis

For the univariate analysis of ROI, 7 anatomical ROI were created based on the previous literature (Table 3.1). Spheres with a MNI coordinates center and a 10 mm radius were created using the MarsBar SPM toolbox (Brett et al., 2002). Each participant's anatomical images were normalized to the MNI space. A mask was created by keeping only voxels within each sphere whose gray matter probability content, as determined by fMRIPrep segmentation, was higher than .2. For a given

3. Motor and Cognitive Aspects of Performance Monitoring During Novel Speech Sequence Production: an fMRI Study – 3.2. Methods

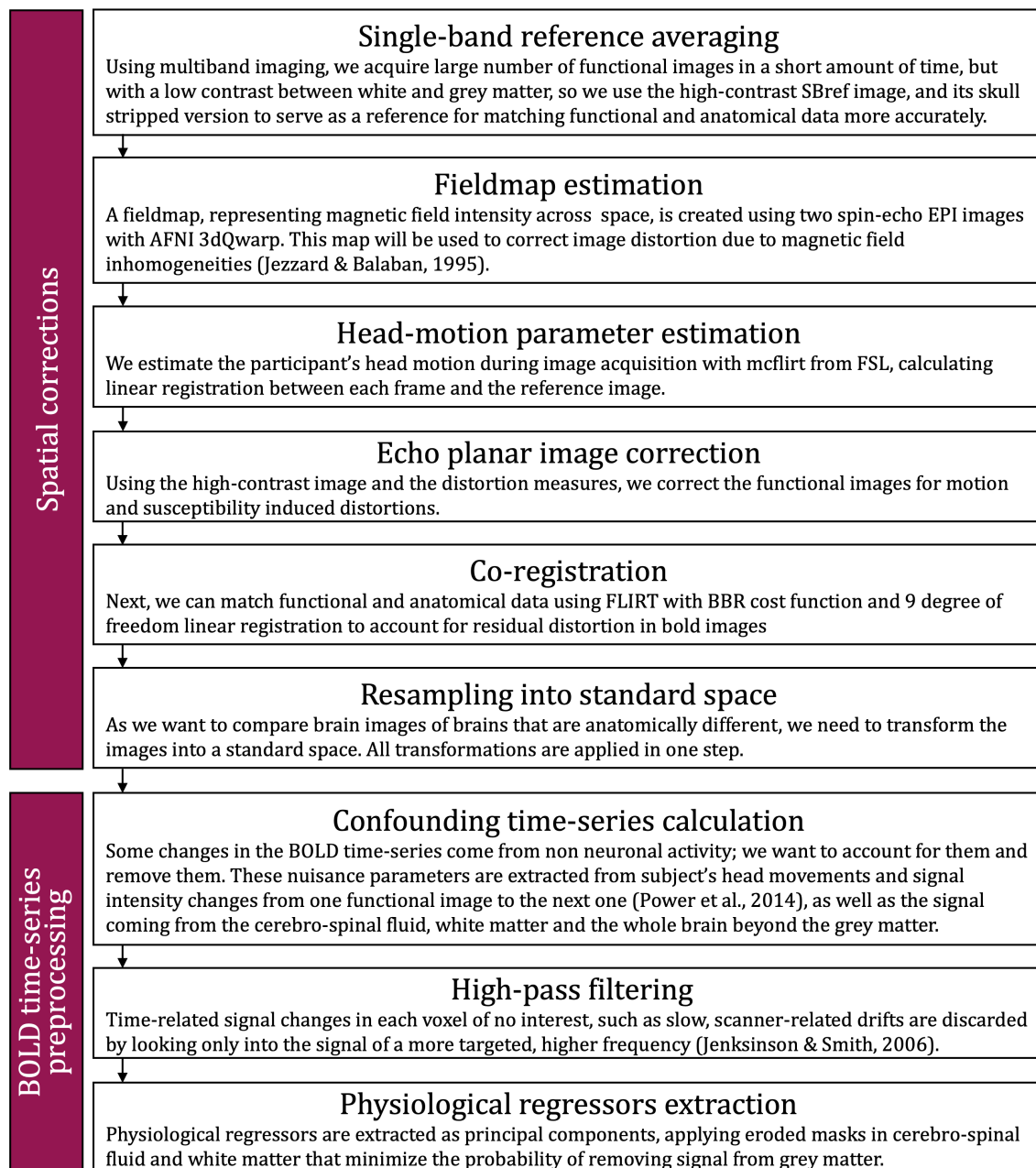


Figure 3.4. – *Functional data preprocessing flowchart*. SBref = single-band reference, EPI = echo planar imaging, AFNI = Analysis of Functional NeuroImages software, 3dQwarp = AFNI's non-linear transformation function, FSL = FMRIB [Functional Magnetic Resonance Imaging of the Brain] Software Library, FLIRT = FMRIB's Linear Image Registration Tool, BBR = boundary-based registration, BOLD = blood-oxygen-level dependent.

3. *Motor and Cognitive Aspects of Performance Monitoring During Novel Speech Sequence Production: an fMRI Study* – 3.3. Results

Region	ROI center	Reference
Cerebellum		
Right crus I	38, -64, -30	Stoodley & Schmahmann, 2009*
Right SMC	16, -59, -23	Golfinopoulos et al., 2010*
Right lobule VIII	28, -56, -54	Rauschecker et al., 2008
Medial frontal cortex		
Left ACC	-6, 20, 34	Gauvin et al., 2016
Left preSMA	-6, 8, 49	Gauvin et al., 2016
Temporo-parietal cortex		
Left pSTG	-65, -33, 14	Golfinopoulos et al., 2010*
Left Spt	-54, -30, 14	Okada & Hickok, 2006

Table 3.1. – *MNI coordinates and references of the ROI classified by anatomical regions. Asterisks indicate meta-analysis or model-based coordinates. MNI = Montreal Neurological Institute, ROI = region of interest, SMC = superior medial cerebellum, ACC = anterior cingulate cortex, preSMA = presupplementary motor area, pSTG = posterior superior temporal gyrus, Spt = Sylvian fissure at the parietal-temporal boundary.*

ROI mask and on the basis of unsmoothed functional images, we extracted the mean of the beta value differences between the contrasted experimental conditions as specified above. For each ROI, we performed one tailed one sample t-test comparing the distribution of the beta value differences to the null hypothesis (no difference). Bonferonni correction was used to correct for multiple comparisons.

3.3. Results

3.3.1. Behavioral results

In the phonotactically illegal, high monitoring load condition, in which the syllables had to be learned, participants clearly produced more errors (49.17%), mean standard error (MSE) .04, standard deviation (SD) .2) than in the legal, low monitoring load condition (16.54%, MSE .03, SD .2). This validates the assumption that the syllables containing new consonant clusters were more error prone and required more monitoring. Production accuracy improved in both conditions. We fitted a logistic mixed model to predict production accuracy as a function of the number of repetitions and the phonotactic legality (formula: Accuracy ~Repetition + Legality + Repetition * Legality). The model included participants, items and stimuli randomization lists as random effects (formula: list (~1 | Participant, ~1 |Item, ~1 | List). Within this model,

3. Motor and Cognitive Aspects of Performance Monitoring During Novel Speech Sequence Production: an fMRI Study – 3.3. Results

both the effect of the repetition number and the phonotactic legality were statistically significant and positive (repetition: $\beta = .09$, 95% confidence interval (CI) [.08, .1], $p < .001$; Std. $\beta = .64$, 95% CI [.57, .71]); legality: ($\beta = 3.11$, 95% CI [2.30, 3.93], $p < .001$; Std. $\beta = 2.48$, 95% CI [1.68, 3.27]). The interaction effect of the repetition number and the phonotactic legality was significant at $p < .001$. Standardized parameters were obtained by fitting the model on a standardized version of the dataset. 95% Confidence intervals and p-values were computed using a Wald z-distribution approximation. Slopes of the learning curves were higher for the illegal compared to the legal items, as shown in Figure 3.5. Finally, the production accuracy in the two conditions differed even in the last experimental run ($p = 2.84 \times 10^{-40}$), indicating that the monitoring load remained high for the new consonant combinations throughout the experiment.

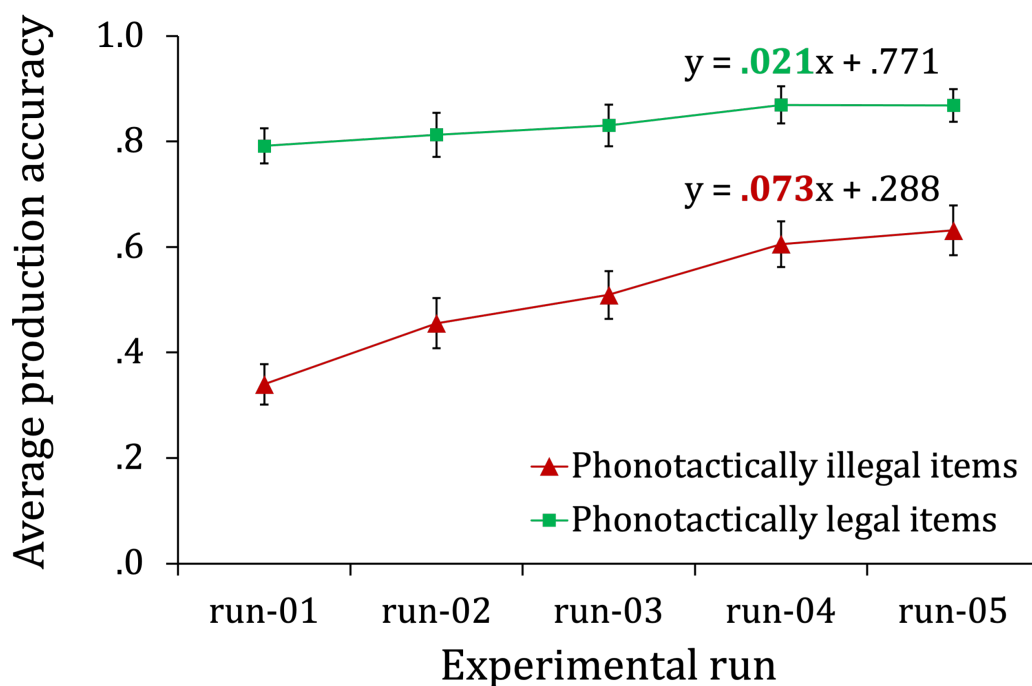


Figure 3.5. – Average production accuracy over five experimental runs for phonotactically illegal, high monitoring load items and phonotactically legal, low monitoring load items. Error bars represent standard errors of the mean. Slopes of the learning curves are shown in bold.

3.3.2. Neuroimaging results

3.3.2.1. Regions of interest

Using a region of interest approach, we examined the beta values difference in 7 predefined regions located in the cerebellum, medial frontal, temporal and temporo-parietal cortices (Table 3.1) in two contrasts of interest, namely external monitoring, contrasting incorrect trials with the correct ones, and internal monitoring, contrasting high and low monitoring load on correctly produced items (Figure 3.6). Two cerebellar regions, right crus I and right superior medial cerebellum, were involved both in the contrast targeting external monitoring and in the contrast targeting internal monitoring of sequence production (crus I, internal monitoring: $p = .04$, external monitoring: $p = .02$; superior medial cerebellum, $p = .01$ for each contrast). This was also the case for the preSMA in the left medial frontal cortex ($p = .02$ in internal and $.001$ in external monitoring). Internal monitoring was further linked to the left pSTG ($p = .01$), while external monitoring was linked to the left ACC ($p = .007$).

3.3.2.2. Whole brain

To examine the specificity of the findings from the ROI analyses, we also conducted a whole-brain analysis (Table 3.2) and Figure 3.7). After correcting for multiple comparisons, both monitoring contrasts engaged cerebellar lobules VI and VIII, as well as the SMA in the medial frontal cortex and the Brodmann area 22 in the temporal cortex. The internal monitoring contrast additionally confirmed the Brodmann area 41 activation in the temporal cortex. Both the internal and the external monitoring contrasts also revealed significant clusters in the parietal cortex as well as in the insula. Figure 3.8 summarizes all the analyses that were carried out.

3.4. Discussion

The aim of the current study was to investigate the cortico-cerebellar correlates of error monitoring during the production of novel syllable sequences. A post-lexical variable, phonotactic legality, was manipulated in two conditions that differed in their degree of required monitoring. The critical condition consisted of novel syllable sequences that had to be learned and heavily monitored. The baseline condition consisted of known syllable sequences, hence required less monitoring. Contrasting these two conditions for correct trials provided an index of internal monitoring at the

3. Motor and Cognitive Aspects of Performance Monitoring During Novel Speech Sequence Production: an fMRI Study – 3.4. Discussion

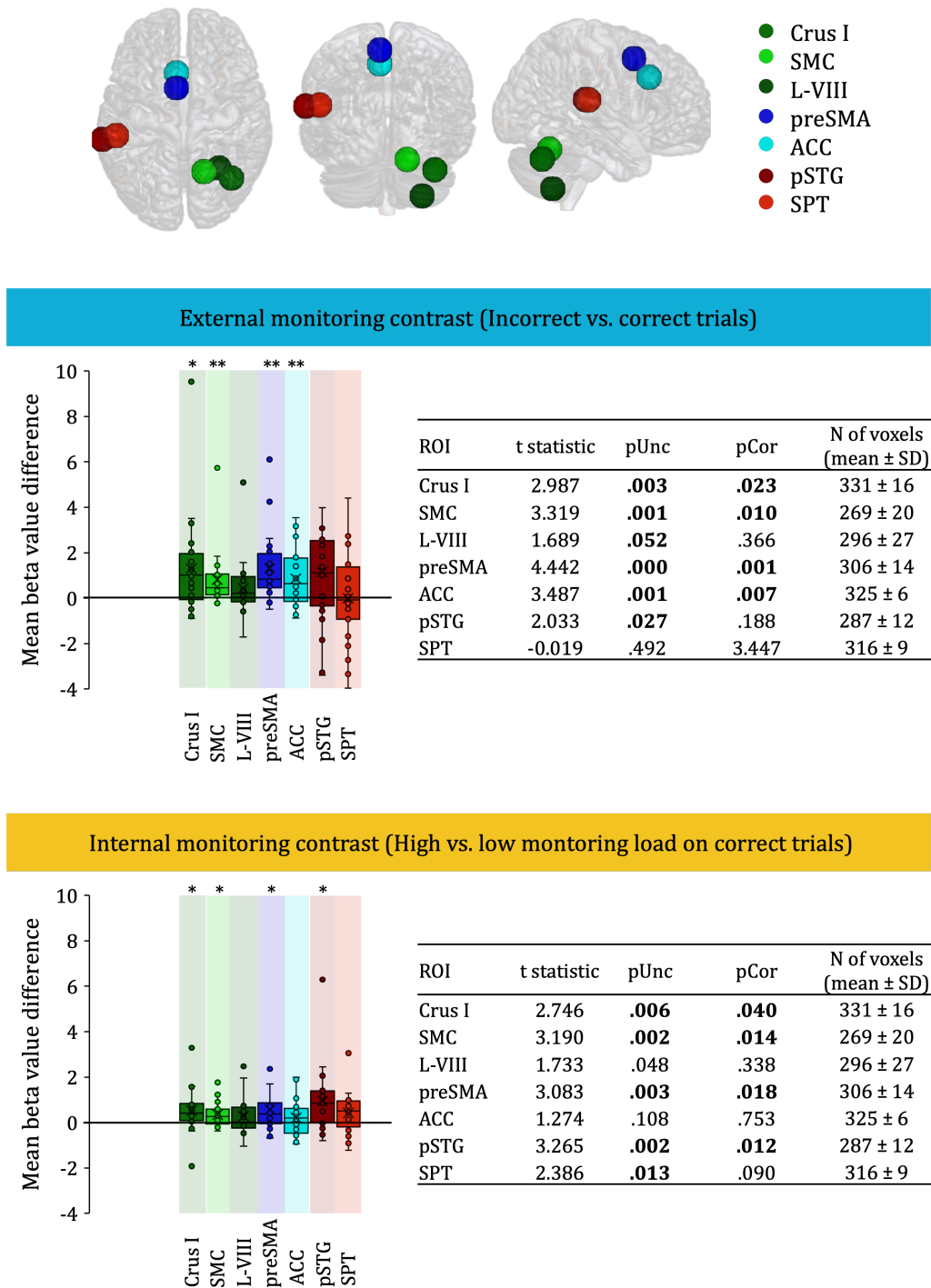


Figure 3.6. – Results of the ROI analysis. Top panel represents the locations of the 7 predefined ROI. Middle panel and bottom panel represent the external monitoring contrast and the internal monitoring contrast respectively. ROI in the medial frontal cortex are represented with blue tones, ROI in the cerebellum in green tones and ROI in the temporal cortex in red tones. The asterisks indicate significant effects $< .05$ (*) or $< .005$ (**) after applying Bonferonni correction for multiple comparisons.

3. Motor and Cognitive Aspects of Performance Monitoring During Novel Speech Sequence Production: an fMRI Study – 3.4. Discussion

Anatomical location	External monitoring (Incorrect > correct)			Internal monitoring (High > low load correct)		
	k	t	MNI	k	t	MNI
Frontal cortex						
L PCG	760	6.27	-56, 8, 36			
L PoCG	73	5.64	-69, -15, 24			
L IFG	32	4.99	-36, 25, 2	34	5.27	-36, 18, 9
Medial frontal cortex						
L SMA	430	6.18	-2, 5, 59	200	6.13	-6, 0, 64
Temporal cortex						
L STG				1120	7.36	-64, -18, 9
R STG				247	7.91	61, 2, -8
L MidTG	99	5.55	-62, -10, -1			
Parietal cortex						
L SPG				51	5.12	-16, -78, 59
L IPG	124	5.62	-36, -45, 39	163	5.79	-32, -45, 44
Occipital cortex						
L SOG	96	4.92	-22, -68, 39			
L MidOG				63	5.26	-26, -72, 32
L IOG	170	6.05	-46, -68, -16	104	5.69	-44, -68, -6
R IOG	27	4.58	28, -98, -11			
L Calc	92	6.24	-6, -98, -8	662	6.23	-16, -68, 9
R Calc	240	5.59	21, -68, 6			
L LG				39	4.38	-14, -55, -1
Cerebellum						
R L-VI	414	6.46	26, -65, -24	90	5.50	28, -65, -21
	27	5.51	8, -72, -14			
R L-VIII	145	5.40	24, -72, -54	154	5.67	21, -72, -54

Table 3.2. – Results of the whole-brain analyses of the external and internal monitoring contrasts. Local maxima of BOLD response separated by > 8 mm. Regions were automatically labeled using the automated anatomical labeling atlas 3 (Rolls et al., 2020) in SPM. MNI coordinates are presented as x, y, z, i.e. in the left-right, anterior-posterior and inferior-superior dimensions, respectively. All peaks are significant at a voxel-wise threshold of $p < .001$ (extent threshold = 25 voxels) and at a cluster threshold of $p < .05$ with an FDR correction for multiple comparisons. L = left, R = right, k = cluster extent, t = t-statistic, MNI = Montreal Neurological Institute, PCG = precentral gyrus, PoCG = postcentral gyrus, IFG = inferior frontal gyrus, SMA = supplementary motor area, STG = superior temporal gyrus, MidTG = middle temporal gyrus, SPG = superior parietal gyrus, IPG = inferior parietal gyrus, SOG = superior occipital gyrus, MidOG = middle occipital gyrus, IOG = inferior occipital gyrus, Calc = calcarine fissure, LG = lingual gyrus, L-VI = cerebellar lobule VI, L-VIII = cerebellar lobule VIII.

3. Motor and Cognitive Aspects of Performance Monitoring During Novel Speech Sequence Production: an fMRI Study – 3.4. Discussion

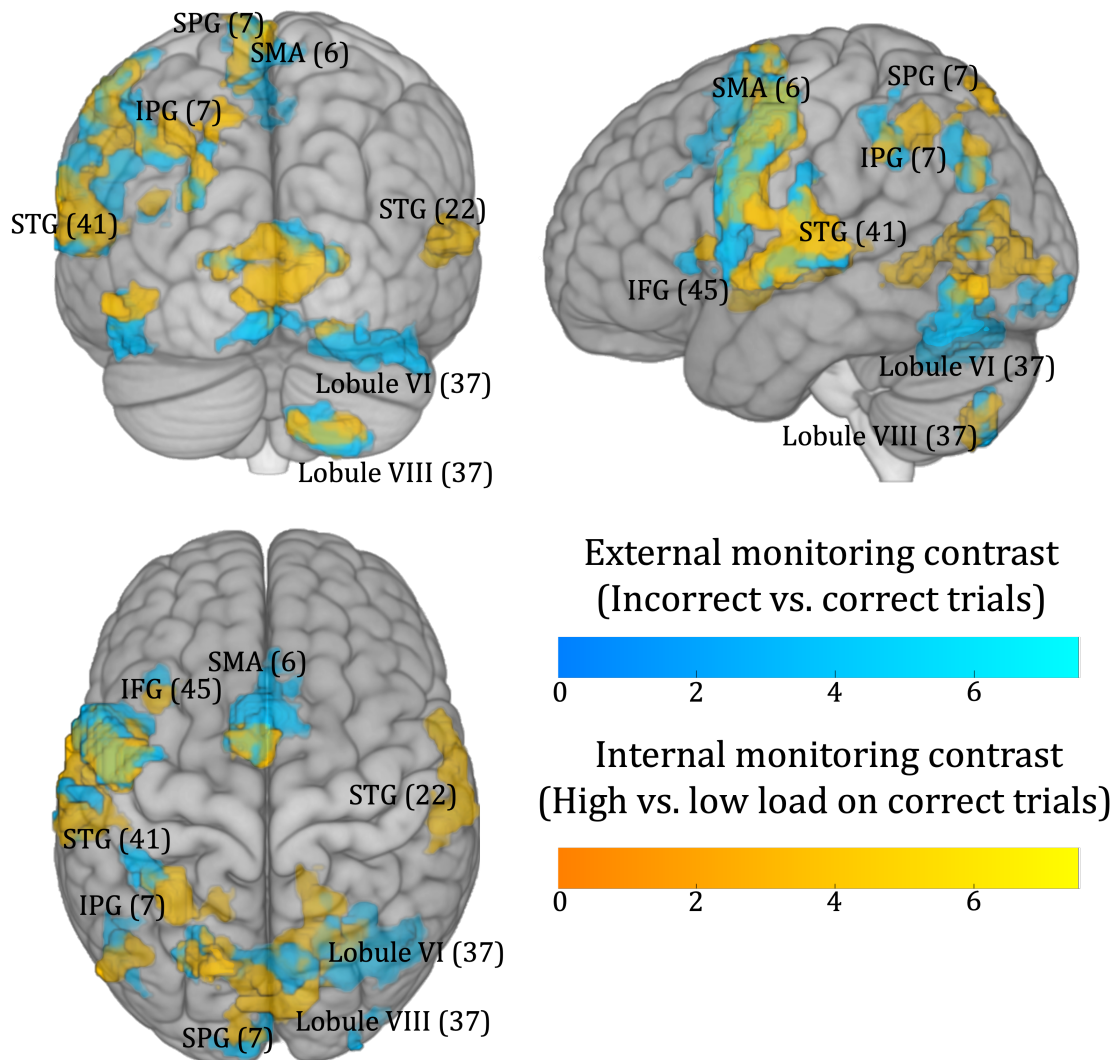


Figure 3.7. – Results of the whole brain analysis. BOLD response of external monitoring (errors vs. correct trials, blue) and internal monitoring (high vs. low monitoring load for correct trials, yellow) contrasts. Overlap across contrasts is visible as green. Statistical t-maps are overlaid on an MNI template using a voxelwise threshold of $p < .001$ and an extent threshold of 25 voxels. Numbers in brackets represent Brodmann areas. IFG = inferior frontal gyrus, STG = superior temporal gyrus, SPG = superior parietal gyrus, IPG = inferior parietal gyrus, SMA = supplementary motor area.

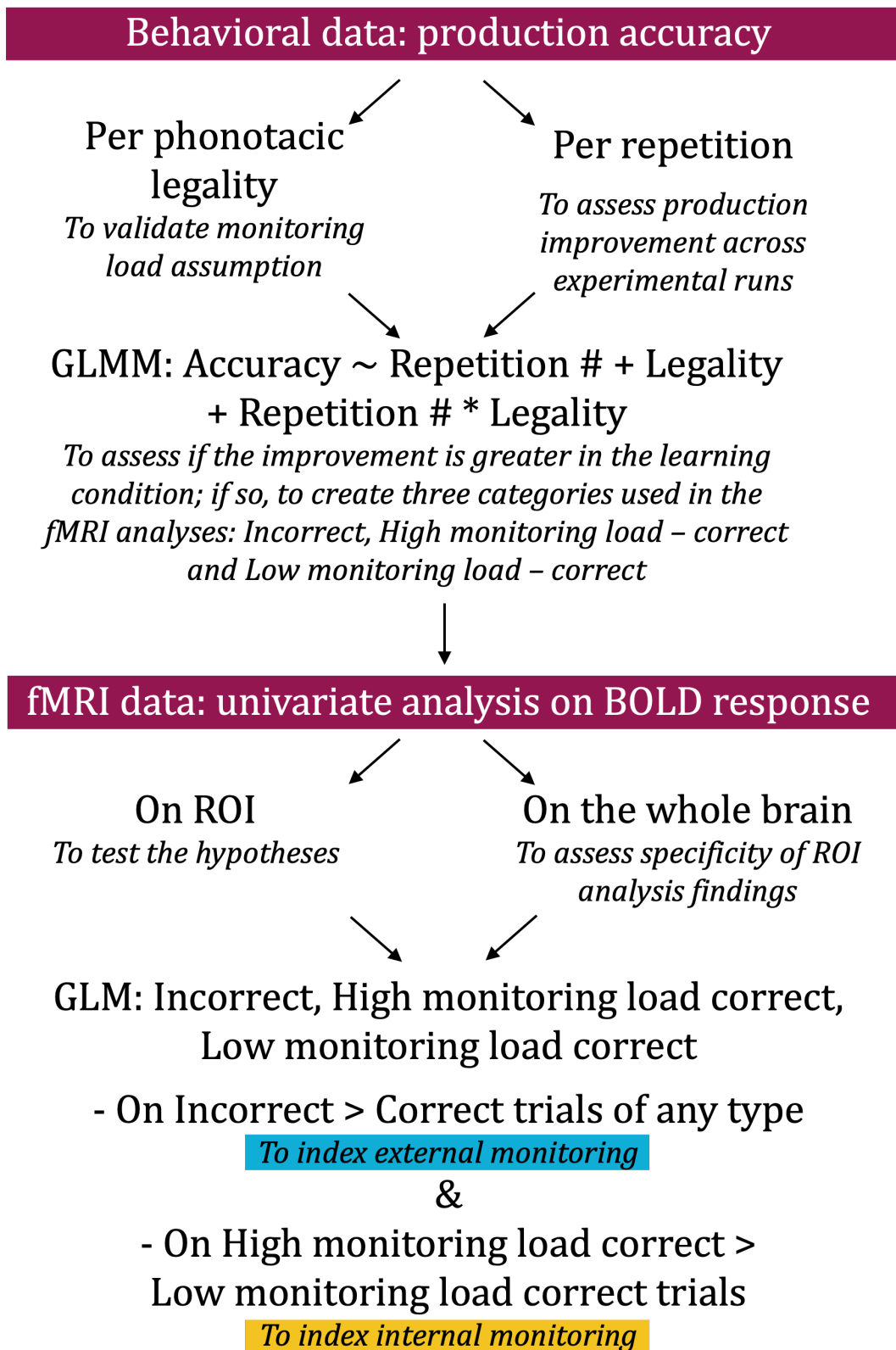


Figure 3.8. – Summary of the analyses conducted in the task-related fMRI analysis. GLMM = generalized linear mixed model, ROI = region of interest, GLM = general linear model.

post-lexical level. Contrasting errors with correct production gave an index of external monitoring. Both types of monitoring were examined in each anatomical region of interest to address our specific hypotheses. A whole-brain analysis was used to inform speech production monitoring debates more generally.

3.4.1. Transmodal and sensorimotor-driven monitoring in distinct cerebellar regions

The first and primary goal of the fMRI analysis presented in this chapter was to refine our understanding of how distinct cerebellar regions contribute to error monitoring in speech. We were interested in contrasting the hypothesis that cerebellar monitoring engages motor regions for monitoring of motor behavior and cognitive regions only for cognition, with the hypothesis that it is transmodal, engaging across the board regions typically classified as cognitive. In line with the latter hypothesis, a region known to be implicated in a broad number of tasks and cognitive skills, crus I, was differentially activated in both contrasts. However, a ROI in the SMC, motor region pertaining to the spinocerebellum, was also differentially activated in both contrasts. Furthermore, in the whole brain analysis, differential activation of cerebro-cerebellar motor regions lobules VI and VIII, not far from the ROI in the lobule VIII, was observed for both internal and external monitoring. The presence of these motor regions for both internal and external monitoring might indicate sensorimotor-driven, hierarchically subordinate speech motor control in addition to the hierarchically superior transmodal performance monitoring. Activation of the superior medial cerebellum was previously detected for external, but not internal monitoring (Runnqvist et al., 2021) in a lexical task. However, as the number of errors in that study did not allow distinguishing between fluent and disfluent errors, activation of the superior medial cerebellum was likely due to articulatory errors, fitting well with the idea that this region engages in controlling speech motor implementation. The superior medial cerebellum, as well as the motor lobules VIII and VI could have a slightly different, not only hierarchically subordinate function. For instance, the SMC is considered to be involved in the generation of proprioceptive input during internal modeling (as opposed to predictive input through crus I). Yet an alternative account can be found in the DIVA model, in which the superior lateral cerebellum, encompassing the crus I, seems to contribute to sensory predictions that form auditory and somatosensory targets. In turn, projections from the superior medial cerebellum are thought to

3. *Motor and Cognitive Aspects of Performance Monitoring During Novel Speech Sequence Production: an fMRI Study* – 3.4. *Discussion*

contribute to precisely-timed feedforward commands as evidenced by speech motor deficits such as ataxic dysarthria (e.g., Golfopoulos et al., 2010). Similarly, the lobule VIII has been linked to the implementation of motor sequences and to efficiency in the representation of articulatory patterns as opposed to learning a novel sequence (e.g., Orban et al., 2010; Rauschecker et al., 2008). Interestingly, learning a new motor sequence, thus taxing monitoring heavily, has been found to also engage the lobule VI (e.g., Orban et al., 2010). The lobule VI was also differentially activated (though on the left side) for internal monitoring of a lexical variable during speech production in addition to crus I (Runnqvist et al., 2021). In the present study, a cluster revealed by the whole brain analysis peaked in the lobule VI, though it likely also comprised the ROI in crus I. Collectively, these studies suggest that the lobule VI might be more transmodal than the lobule VIII or the superior medial cerebellum. Regardless, globally the current results fit the idea of a sensorimotor to transmodal functional gradient in the cerebellum, as proposed by Guell and colleagues (Guell et al., 2018).

3.4.2. Sensorimotor-driven feedback monitoring through the medial frontal cortex

We looked into two regions of the medial frontal cortex, the preSMA and the ACC, to contrast transmodal monitoring through conflict with sensorimotor-driven monitoring of articulatory representations or sensory feedback. Differential preSMA activation was observed for both internal and external monitoring, but the ROI in the ACC was only activated for external monitoring. Regarding the preSMA, the whole brain analysis largely confirmed this pattern by revealing significant clusters peaking in the SMA, presumably comprising the ROI in the preSMA (see Figure 3.7), for both internal and external monitoring. This pattern is consistent with the idea of transmodal, conflict-based monitoring and with the notion that the preSMA is involved when motor-related difficulties arise. This suggests that the preSMA plays a role in monitoring articulatory aspects of speech production, and is thus sensorimotor-driven. Concerning the ACC, our findings are more consistent with a role of the targeted cingulate region in processing overt feedback as indicated by overt errors, as opposed to performing conflict-based monitoring. This is in line with the observation of the ACC exclusively for external monitoring in Runnqvist et al. (2021), who proposed an integrative account of the ACC activations for error and feedback monitoring following the proposed vocal control network put forward by Loh and colleagues (2020).

3. *Motor and Cognitive Aspects of Performance Monitoring During Novel Speech Sequence Production: an fMRI Study* – 3.4. Discussion

These authors argue that, across primates, area 44 is in charge of cognitive control of orofacial and non-speech vocal responses, and the midcingulate cortex is in charge of analyzing vocal non-speech feedback driving response adaptation. Furthermore, the cognitive control of human-specific speech vocal information would require the additional recruitment of area 45 and preSMA. In this framework, it would not be the conflict that generates activation in the ACC and preSMA observed here, but rather the feedback provided through the articulated error. While the present ACC pattern fits well with this explanation, the preSMA pattern would require assuming that the internal monitoring contrast reflects some form of sensory feedback, such as pre-articulatory proprioception. Further supporting the hypothesis of a vocal feedback control network regarding the current results is activation in BA 45 (Figure 3.7), just as in Runnqvist et al. (2021). Thus, globally speaking, the current results support the idea that the medial frontal cortex is involved in sensorimotor-driven feedback monitoring. Nevertheless, a caveat to this conclusion is the considerable size and functional heterogeneity of the medial frontal cortex. Due to this, a better understanding of its contributions to monitoring will likely require both the consideration of a larger set of anatomical subregions and the use of functional localizers.

3.4.3. Transmodal predictive monitoring through temporo-parietal cortices

Finally, we looked into two temporo-parietal regions, the pSTG and the Spt, to contrast transmodal monitoring through internal modeling or speech comprehension with sensorimotor-driven monitoring through auditory feedback. The ROI approach unexpectedly only revealed a significant differential activation of the pSTG in the contrast of internal monitoring. This result differs from previous results that reported pSTG activation in conditions taxing external but not internal monitoring. However, when completing this picture with the results of the whole brain analysis, one can see that while only the internal monitoring contrast results in differential activation of Brodmann area 41, both monitoring contrasts engage Brodmann area 22, though with a different laterality. A recent meta-analysis (e.g., Meekings & Scott, 2021) indicates that active regions in speech error monitoring often differ across studies, and consequently renders the crucial functional role of the pSTG controversial. Clearly, the present results contribute to this impression of disparity, but at the same time they also indicate that the temporal cortex engages in error monitoring. A systematic

3. *Motor and Cognitive Aspects of Performance Monitoring During Novel Speech Sequence Production: an fMRI Study* – 3.5. *Conclusion*

research approach seems necessary to increase our understanding of the relevant variables governing the location of such monitoring related activations. The whole brain analysis further revealed a significant cluster in the inferior parietal gyrus for both contrasts, possibly comprising part of the ROI in Spt. Moreover, a cluster in the superior parietal gyrus was revealed for the contrast of internal monitoring. Besides possibly comprising a part of Spt, these parietal clusters are of interest because the parietal cortex has been linked to motor imagery and the retrieval of a stored internal model (e.g., Blakemore & Sirigu, 2003). A functional loop between the parietal cortex and the cerebellum, estimating the current status of the motor system throughout movement execution and allowing for predictive monitoring, has been proposed. Thus, the combined but differential activation of temporal, parietal and cerebellar regions for internal and external monitoring suggests that their involvement goes beyond sensory feedback and speech comprehension, and is more in line with proposals linking these regions to auditory and somatosensory targets and predictive internal modeling-like mechanisms at play during speech production (e.g., Hickok, 2012; Hickok, 2014; Tourville & Guenther, 2011; Guenther et al., 2006).

3.5. Conclusion

The study reported in this chapter examined the domain-general, transmodal, or representation specific nature of internal and external monitoring of speech production. Current models imply that monitoring mechanisms are the same during different processing stages (e.g., Levelt et al., 1999; Nozari et al., 2011; Pickering & Garrod, 2013), supporting the transmodal account. Our results showed that this assumption is correct to some extent as we obtained evidence for transmodal monitoring presumably carried out through predictive internal modeling, via the crus I in the cerebellum, and possibly linked to regions in temporal and parietal cortices (e.g., Stockert et al., 2021; Clower et al., 2001; Glickstein, 2000). However, we also obtained evidence for a sensorimotor-driven type of monitoring or control in motor regions of the cerebellum and in the medial frontal cortex. As most models of speech production agree that not all levels of language are directly linked to sensorimotor aspects (e.g., Levelt et al., 1999; Dell, 1988; but see Strijkers, 2016 and Fairs et al., 2021 for an account of parallel and distributed processing in language production), it is perhaps not surprising that levels directly connected to sensorimotor aspects of speech capitalize on monitoring mechanisms based on auditory and proprioceptive

3. *Motor and Cognitive Aspects of Performance Monitoring During Novel Speech Sequence Production: an fMRI Study* – 3.5. *Conclusion*

feedback. An open question to be addressed in future research is whether the coexistence of transmodal and sensorimotor-driven monitoring indicates complementary functions of the involved brain regions. In this case, different regions could monitor different representations or be triggered to deal with specific needs, such as repairing or compensating for overt errors.

4. Motor Tuning and Internal Model Refinement During Novel Skill Acquisition: a Resting-state Functional Connectivity Analysis

As we are acquiring new motor skills, our internal models of the outside world are refined, and our motor performance becomes smoother and increasingly automatized. This achievement involves various cortical areas and both the cerebellum and the basal ganglia. Although the two regions are functionally interconnected, forming loops with the cerebral cortex, there is little evidence on how these loops are created at the beginning of learning, or how they evolve as learning progresses. In this chapter, we focused on functional brain networks reconfiguration that takes place over the course of learning. Using a novel syllable production paradigm, we looked into the changes in the connectivity between both these regions with the rest of the brain and the factors that affect these changes, such as motor improvement and learning progress. The connectivity of both the cerebellum and the basal ganglia with frontal and medial frontal cortical areas changed after the participants' repetition-induced motor improvement. Learning-related changes, arguably reflecting internal model updating, resulted in an increase in the connectivity between the right crus I in the cerebellum and the left precuneus in the parietal cortex. This connectivity change could be explained by the cerebellum's increasing capacity to predict the future state of the parietal cortical activity and synchronize with it. The parietal activity could then sustain motor imagery that precedes movement articulation, further improving cerebellar prediction.

Keywords: crus I, precuneus, speech motor sequence learning, resting state functional connectivity, motor tuning, internal models

4.1. Introduction

Language learning is a form of adapting one's behavior to a novel environment. If a speaker wants to make themselves understood in a foreign language, whose rules of combining sounds into syllables are different than what they are familiar with, they will need to adapt their articulatory movements to this new environment. Such adaptations occur through the creation and continuous refinement of a neural representation of an action, in this case, the action of pronouncing a novel sequence of sounds, and result in progressively automatized, smooth movement synchronization.

4.1.1. Cerebellum and internal modeling

As the speaker's pronunciation approaches the desired outcome, through repeated practice, the internal model is refined. This capacity of creation, maintenance and updating of internal models is mainly sustained by the cerebellum and its predictive capacity. Concretely, predictions that are sustained by the cerebellar activity are compared with actual input in subcortical or cortical areas. The difference between the two then serves as the generator of the error signal (Sokolov, Miall, & Ivry, 2017). The error signals can be used to update the model and maximize motor learning (Blakemore & Sirigu, 2003), as the movements become smoother and more refined (Sokolov, Miall, & Ivry, 2017). Additionally, the cerebellum contributes to novel speech sequence learning, or language processing more generally, through working memory, temporal representation and attentional system, whose activity is thought to increase during error detection, activating the forward model (Ito, 2008). Motor imagery, covert repetition, rehearsal, conflict detection and outcome prediction are all sustained by the cerebellum, and, to some extent, the parietal cortex (Sokolov, Miall, & Ivry, 2017). In non-learning language tasks, the posterior cerebellum is coactivated, in a contralateral fashion, with left prefrontal and left parietal cortices (Sokolov, Miall, & Ivry, 2017; Ito, 2008), and reciprocally connected with the prefrontal cortex in monkeys (Ito, 2008).

4.1.2. Basal ganglia and motor learning

Automatization and smoothness in the production of novel motor sequences, in language and beyond, relies on motor sequence planning. Planning of upcoming movements in a sequence, both in humans and in rhesus monkeys, is sustained by

4. *Motor Tuning and Internal Model Refinement During Novel Skill Acquisition: a Resting-state Functional Connectivity Analysis – 4.2. Methods*

the basal ganglia (de Long & Alexander, 1987; Golfinopoulos et al., 2010). Getting to this automatization phase, where less and less feedback is necessary, is reflected at the neural level in a transition from frontal cortical activity towards the recruitment of the basal ganglia. In the DIVA model, timing of the motor command release is dependent on reciprocal connections between the SMA and different parts of the basal ganglia, such as caudate nucleus, putamen and globus pallidus (Golfinopoulos et al., 2010). In motor learning in general, basal ganglia have been identified as a part of cortico-subcortical networks contributing to performance improvement (Fernandez-Seara et al., 2009), automaticity (Wu, Chan & Hallett, 2010) and consolidation (Debas et al., 2014).

4.1.3. Objectives

As the cerebellum and the basal ganglia both have crucial roles in a set of processes that enable novel speech sequence learning and motor learning in general, we wanted to examine how their connectivity with cortical areas evolves with progress made in learning of a skill that combines both of them, novel syllable production. To this end, we recorded resting state fMRI data prior to and following a task in which participants learned to articulate novel types of syllables. We focused on the changes in the functional connectivity of regions in the basal ganglia and the cerebellum with the rest of the brain. We explored how are these changes modulated as a function of the participant's motor articulation improvement and learning progress.

4.2. Methods

4.2.1. Participants

The participant pool was identical to that reported in Chapter 3.

4.2.2. Experimental design and behavioral data analysis

The experimental design was identical to that reported in Chapter 3. Additionally, production accuracy in each of the five experimental runs was used to attribute a single learning coefficient to each participant. This coefficient was based on their individual progress as measured by the slope of their learning curves. The slope of the learning curve for phonotactically illegal items was used to index the learning progress

4. *Motor Tuning and Internal Model Refinement During Novel Skill Acquisition: a Resting-state Functional Connectivity Analysis* – 4.2. Methods

of each participant. The slope of phonotactically legal items was used as a control, as we expected motor performance to improve in both conditions due to training, but more substantially in the illegal, learning condition.

4.2.3. Imaging data acquisition, anatomical and functional data preprocessing

As these steps were identical for the resting state and for the task related data, please refer to the respective sections of Chapter 3 for more information. After the common preprocessing steps, the anatomical and functional volumes were imported to the functional connectivity toolbox CONN (Whitfield-Gabrieli & Nieto-Castañón, 2012) together with tissue masks, potential confound regressor variables and realignment, scrubbing and QC-timeseries covariates. Using CONN's denoising procedure, a noise component for each outlier identified by fMRIPrep during the outlier identification preprocessing step was used to remove any influence on the BOLD signal scan through censoring. Data were then bandpass-filtered (.008-.09 Hz), spatially smoothed with a Gaussian kernel (FWHM 5 mm) and denoised using the following regressors: 12 WM dimensions, 12 CSF dimensions, 24 realignment parameters (6 dimensions, quadratic effects and first-order derivatives, non-filtered), 58 scrubbing (censoring) dimensions, effect of rest-pre with 1 dimension and one first-order derivative, non-filtered, effect of rest-post with 1 dimension and one first-order derivative, non-filtered. Regression was followed by filtering (RegBP), detrending was set to 1 and despiking to zero.

4.2.4. Functional connectivity analyses

In order to perform seed-to-voxel analysis, we selected six seed ROI (Figure 4.1). Cerebellar seed regions were identical to those reported in Chapter 3 (Table 3.1). The anatomical form of the regions of interest in the basal ganglia did not allow us to use 10 mm radius spheres. We therefore used anatomically defined regions of interest taken from FSL Harvard-Oxford maximum likelihood subcortical atlas (Kennedy et al., 2016), merging left and right putamen, caudate nucleus and pallidum in order to obtain three bilateral ROI.

Resting state functional connectivity analyses were conducted in CONN. We created individual resting state functional connectivity maps of correlations between the time series of the individual seed regions (right cerebellar crus I, right cerebellar

4. *Motor Tuning and Internal Model Refinement During Novel Skill Acquisition: a Resting-state Functional Connectivity Analysis – 4.2. Methods*

lobule VIII, right superior medial cerebellum, bilateral putamen, bilateral globus pallidus, bilateral caudate nucleus) and those of every voxel in the whole brain for each participant and for both resting states, before and after the learning task.

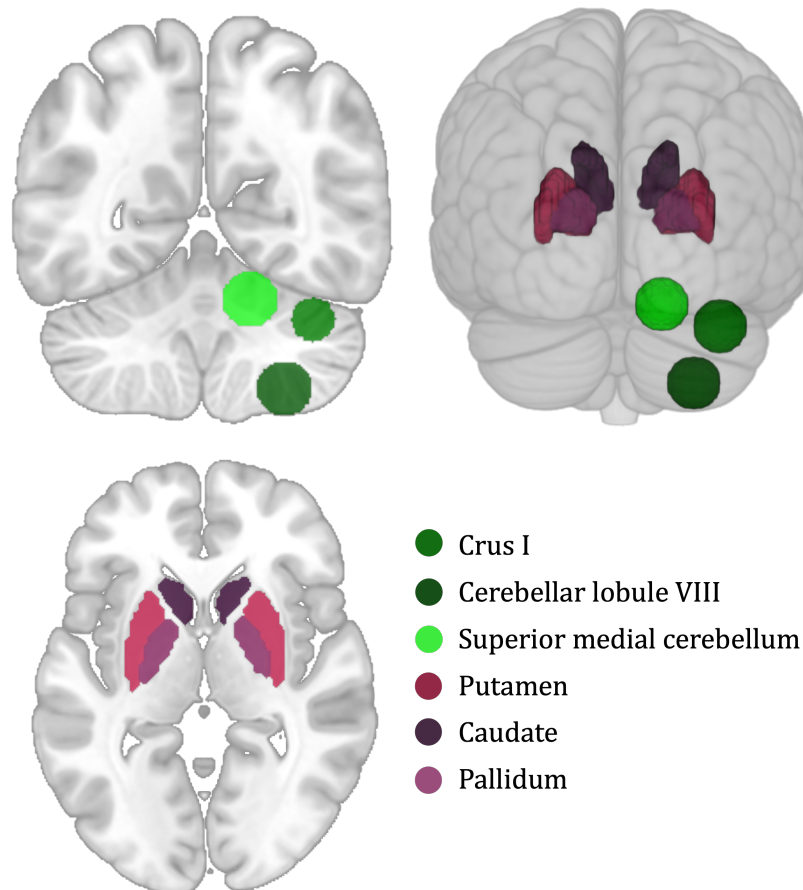


Figure 4.1. – *Seed regions of interest used in seed-to-voxel connectivity analysis.*

4.2.4.1. Statistical analyses

A general linear model was used for second-level analyses of functional connectivity data. We first compared the connectivity values, i.e. Fisher-transformed correlation coefficient values of each seed between the two sessions, prior and following the learning task.

We used the resulting maps to further disentangle the changes that are due to the actual learning and not to a habituation-induced improvement in motor performance, by examining the correlation of these changes with the learning index that was

4. *Motor Tuning and Internal Model Refinement During Novel Skill Acquisition: a Resting-state Functional Connectivity Analysis* – 4.3. Results

attributed to each participant based on the slope of their learning curve calculated for the items in the illegal, learning condition. If the connection changes were due to motor improvement, we expected to observe them when comparing the resting state before the task with the one after it. If the changes were due to learning, and not to an increase in the performance accuracy, we expected to see a significant correlation of the mean connection change measure (the difference in the Fisher-transformed correlation coefficient values before and after the learning task) with the learning index (subject-effect covariate). There should be no correlation, however, between the connectivity change and the slope of the learning curve for the legal items that served as the control condition. We used Bonferonni-corrected FDR values to account for the fact that we had six seed regions of interest.

4.3. Results

4.3.1. Behavioral results

Please refer to the corresponding section in Chapter 3 and to Figure 3.5 in particular.

4.3.2. Neuroimaging results

4.3.2.1. Task-related changes

In order to look into the average connectivity changes induced by the behavioral task, we compared the resting state after the task to that before the task performance, between each of our six cerebellar and subcortical seeds and the rest of the brain. As shown in Table 4.1. and Figure 4.2, both the cerebellum and the basal ganglia differentially connected with the regions in frontal and medial frontal cortices.

4.3.2.2. Learning-related changes

In order to disentangle the connectivity changes induced by learning from those related to task habituation, we looked at the correlation between (a) the mean connectivity change, obtained via a paired t-test between the Fisher-transformed correlation coefficient values pre and post learning and (b) the slope of the learning curve in the phonotactically illegal condition. This correlation was only significant for the seed in the right crus I, which was differentially connected with a cluster

4. Motor Tuning and Internal Model Refinement During Novel Skill Acquisition: a Resting-state Functional Connectivity Analysis – 4.3. Results

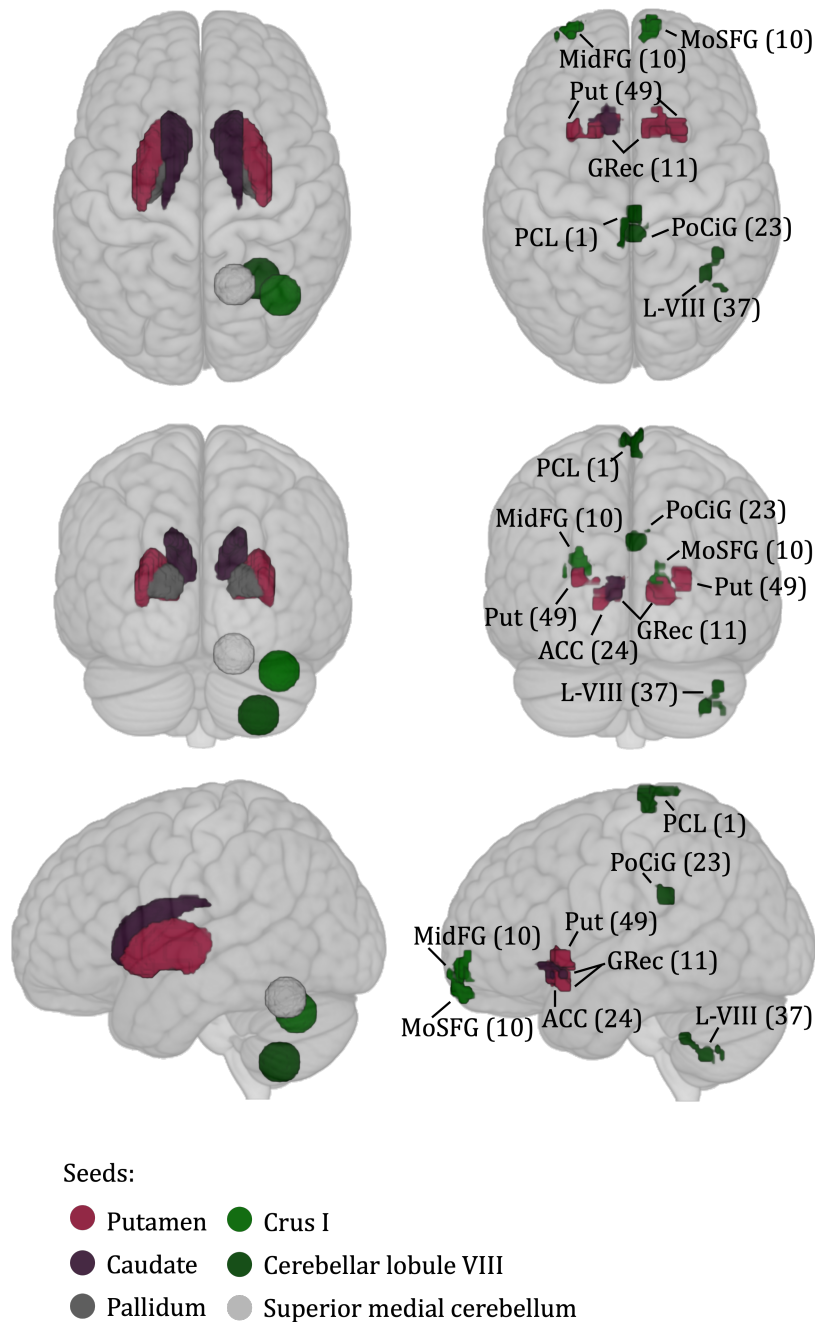


Figure 4.2. – *Task-related functional connectivity changes*. Left: seeds. Right: targets. ACC = anterior cingulate cortex, L-VIII = cerebellar lobule VIII, GRec = gyrus rectus, MidFG = middle frontal gyrus, MoSFG = medial orbital superior frontal gyrus, PoCiG = posterior cingulate gyrus, PCL = paracentral lobule, Put = putamen. Numbers in brackets correspond to Brodmann areas. Seeds whose connection to the rest of the brain did not change in function of task are presented in grey tones. p-values are FDR Bonferonni-corrected.

4. Motor Tuning and Internal Model Refinement During Novel Skill Acquisition: a Resting-state Functional Connectivity Analysis – 4.4. Discussion

Seed ROI	Connected to	k	p	MNI
Bi Caudate	L GRec (BA11)	91	.011	-12, 20, -12
Bi Putamen	L ACC (BA24)	112	.002	-6, 22, -4
	R GRec (BA11)	105	.002	12, 20, -12
	R Putamen (BA49)	81	.009	22, 18, -2
	L Putamen (BA49)	63	.031	-28, 14, -4
R L-VIII	L PCL (BA1)	89	.022	-4, -32, 78
	R L-VIII (BA37)	77	.028	36, -52, -48
	R PoCG (BA23)	68	.039	2, -36, 26
R Crus I	R MoSFG (BA10)	86	.020	8, 68, -14
	L MidFG (BA10)	75	.024	-28, 64, -8

Table 4.1. – *Task-induced connectivity changes in the resting state connectivity of the basal ganglia and the cerebellum with the rest of the brain.* ROI = region of interest, MNI = Montreal Neurological Institute, k = cluster size in voxels, Bi = bilateral, L = left, R = right, BA = Brodmann area, GRec = gyrus rectus, ACC = anterior cingulate cortex, L-VIII = cerebellar lobule VIII, PCL = paracentral lobule, PoCG = posterior cingulate gyrus, MoSFG = medial orbital superior frontal gyrus, MidFG = middle frontal gyrus.

peaking in the precuneus in the left posterior parietal cortex (-16, -70, +36; size 95 voxels; p-FDR = .001445; Figure 4.3a). A post hoc analysis of connectivity between the right crus I and the cluster in the left precuneus showed no correlation between the connectivity change in these two regions from the first to the second resting state and the slope of legal items, indicating that the change was learning-related and did not reflect practice-induced increase in the motor performance. In order to visualize the direction of this learning-related connectivity change, we plotted its Fisher-transformed connectivity values against the values of the slopes of the two learning curves (Figure 4.3b and c, Appendix C). Correlation between the connectivity values and those of the illegal and legal slope confirmed our observation ($R^2 = .7515$ and .0064 respectively).

4.4. Discussion

In this study, we were interested in functional connectivity changes over the course of learning to perform a complex motor skill, a process that reflects a continuous update of an internal model and results in fine tuning of motor articulatory movements. Given that internal model updating relies heavily on the cerebellum, and motor tuning on the basal ganglia, we chose to focus on these regions and their

4. Motor Tuning and Internal Model Refinement During Novel Skill Acquisition: a Resting-state Functional Connectivity Analysis – 4.4. Discussion

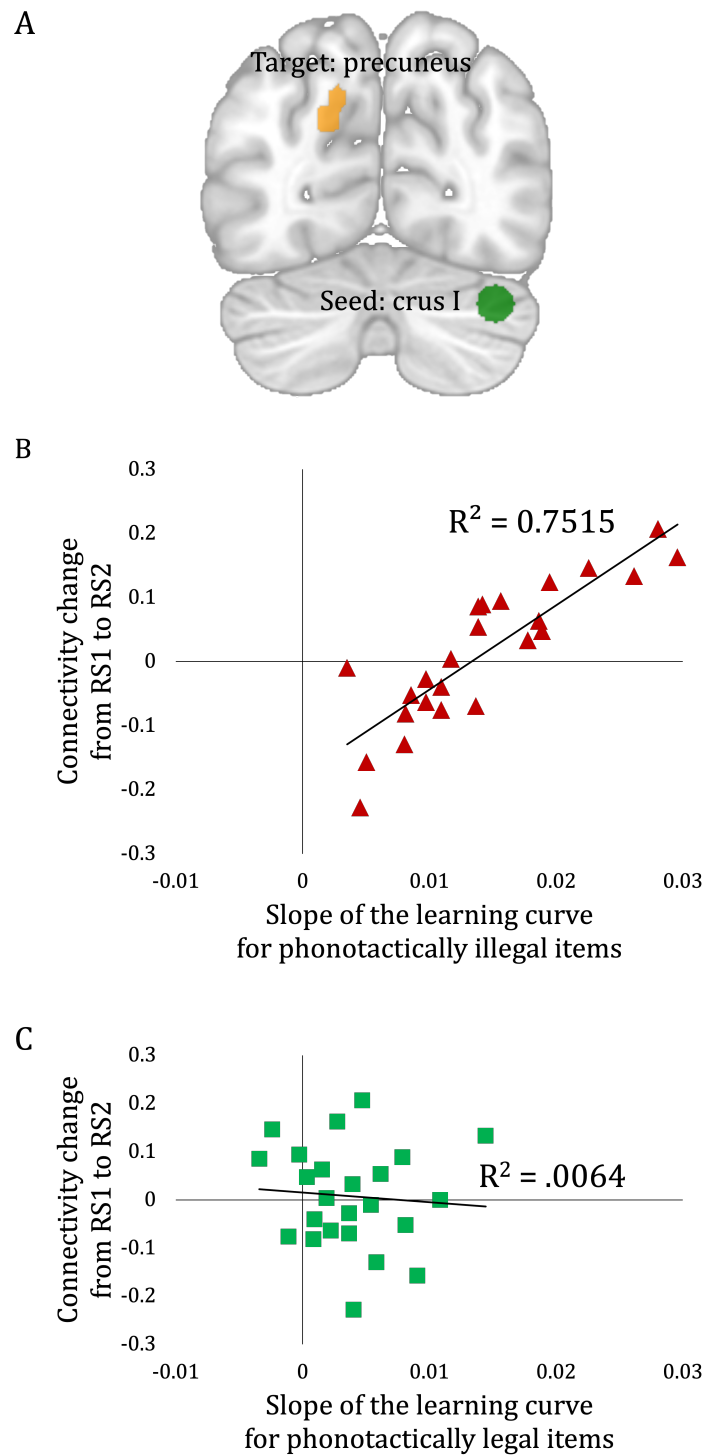


Figure 4.3. – *Learning-related functional connectivity changes*. A: Connectivity change between the right cerebellar crus I and left precuneus from first to the second resting state. B & C: Correlation between the mean connectivity change of functional connectivity between the right crus I and the left precuneus and the slope of the learning curve for phonotactically illegal and legal items respectively.

4. Motor Tuning and Internal Model Refinement During Novel Skill Acquisition: a Resting-state Functional Connectivity Analysis – 4.4. Discussion

connectivity with the rest of the brain. To this end, we recorded resting state fMRI data of participants before and after they learned to pronounce new syllables. Participants' learning progress was correlated with the changes in the resting state connectivity.

4.4.1. Syllable production accuracy improvement as motor tuning

When a speaker's pronunciation of a novel syllable improves, this improvement can be due to the refinement of the way the speaker's articulators move. The increasing accuracy in the production of the syllables observed in all the participants indicates presence of motor tuning. Comparing the maps of functional connectivity of the seeds in the basal ganglia and in the cerebellum with each voxel of the brain at resting state before and after performing the syllable production task, we observed changes in the connectivity of both regions.

Regarding the basal ganglia, we found a change in the connectivity of putamen with the frontal and the medial frontal cortical areas, and specifically the ACC. Changes in the activity of both putamen and the ACC as a function of fast, initial learning of a motor skill have already been reported in non-linguistic tasks (Floyer-Lea & Matthews, 2005; Fernández-Seara et al., 2009). Using perfusion fMRI, previous motor learning studies detected an increase in the bloodflow related to activity in putamen and in the SMA in the medial frontal cortex (Fernández-Seara et al., 2009). Our finding brings closer together motor learning and speech production studies, as it indicates the existence of a shared domain-independent motor improvement network. Furthermore, it expands the nature of this network from functional *co-activation* to functional *connectivity*.

In the cerebellum, we detected connectivity changes in brain regions that have been associated with skilled motor practice, such as the middle frontal gyrus (Zwicker et al., 2011), but also with the posterior cingulate gyrus, whose activation on a longer temporal scale of motor learning increases together with that of the basal ganglia (Lohse et al., 2014). This finding suggests a possible larger network whose connectivity continuously changes until learning consolidation.

4.4.2. Syllable learning as an internally modeled mental activity

Certain connectivity changes are the result of the learning progress itself, beyond motor improvements. In order to identify those changes, we compared learning curves of each participant that describe the change in their syllable production accuracy over the course of the experiment. In the learning condition, performance improvement is the result of novel skill acquisition, while in the control condition it may be explained by practice. We used the measure of the curve of the learning condition as a within-subject covariate to distinguish between the changes due to motor improvement from those due to learning of a novel skill, and therefore acquisition of an internal model. Adding this covariate while comparing maps of functional connectivity between each subcortical seed and each voxel of the brain at resting state before and after performing the learning task allowed us to detect a learning-related connectivity change between the right crus I in the cerebellum and the left precuneus in the posterior parietal cortex. The correlation between the mean connectivity change values and the index of the learning progress revealed that augmented learning was positively correlated with the change in the connectivity in these two regions before and after the learning task.

In a novel situation, such as learning a novel language in which syllable types differ from the existing patterns, the internal model is refined through repeated practice, leading to smooth, accurate articulation achieved through better movement control (Ito, 2008). In the DIVA model, this refinement leads to increased correspondence between the sensory and auditory targets and feedback, and a decrease in the number of errors. According to Ito's internal model hypothesis, modeling mental activities would require a large-scale network including both the cerebellum and the parietal cortex (Ito, 2008). These regions are known to be functionally coupled in resting state data, in particular through connectivity between supramodal zones of the cerebellum, including crus I and II, and the inferior parietal lobule (O'Reilly et al., 2010). Our study shows that the coupling between the crus I and the precuneus is reinforced during learning. Interesting hypotheses can be made to guide further studies, that could verify the directionality of this coupling through effective connectivity analyses. Through its internal model, the cerebellum could recognize current neural states in the precuneus. Refining the internal model of novel syllables during learning could be mediated through increasingly accurate predictions by the cerebellum of *upcoming* parietal cortical activity, enabling the two regions to synchronize (Sokolov, Miall & Ivry, 2017). Next, the parietal cortical activity could reflect motor imagery that precedes

4. Motor Tuning and Internal Model Refinement During Novel Skill Acquisition: a Resting-state Functional Connectivity Analysis – 4.5. Conclusion

movement articulation (Blakemore & Sirigu, 2003). Stored movement representations that are necessary for generating more accurate predictions (Blakemore & Sirigu, 2003) would then be communicated to the cerebellum and used both to predict sensory outcome of the articulation and to provide information on how to correct the articulation in the future if necessary (Sokolov, Miall & Ivry, 2017).

4.5. Conclusion

When learning a new motor skill, we refine both the motor performance itself and the internal model of that performance (Kawato, 1999). While it is known that both the basal ganglia and the cerebellum participate in broader networks that enable motor learning, little is known about the changes in functional connectivity during the skill acquisition. There is even less evidence on what these reconfigurations could mean in the domain of speech, given that most motor learning studies focus on manual, non-speech articulation learning. Our results imply that motor articulation improvements induce changes in the way the cerebellum and the basal ganglia are connected with frontal and medial frontal cortical regions. Internal model improvements, on the other hand, could be mediated through a cerebello-parietal loop, in which the cerebellum predicts the upcoming activations in the parietal cortex, and uses that information to further refine the model.

5. General discussion

5.1. Overview

The acquisition of new skills that enable us to function in everyday life relies on domain-general cognitive mechanisms sustained by large portions of cortical and sub-cortical brain areas. Neurocognitive models of motor sequence learning acknowledge that vast portions of cerebral cortex, including motor cortical, frontal, parietal and temporal areas, as well as the cerebellum and the basal ganglia support the learning process, particularly in its early phases (Doyon et al., 2002; Doyon & Benali, 2005; Doyon et al., 2009; Doyon et al., 2018). Major linguistic neurocognitive models of speech production, comprehension and motor control involve the same broad areas (Hickok, 2012; Guenther, 1994; 2006; Guenther, Ghosh & Tourville, 2006; Golfopoulos, Tourville & Guenther, 2010; Tourville & Guenther, 2011; Kotz & Schwartz, 2010). Drawing from both traditions, we aimed to shed light on the cognitive processes and the neural underpinnings that support learning of a skill that is both motor and linguistic. We asked to what extent the cognitive processes linked to brain areas during motor sequence learning and speech production are domain-general or task-specific. A meta-analytic approach was used to specify the cognitive processes in the brain areas related to motor sequence learning that could differ as a function of experimental design. In an fMRI study, we focused on motor sequence learning in speech production, and aimed to further examine two cognitive processes in particular: performance monitoring and internal model refinement.

In the study presented in Chapter 2, we performed a coordinate-based meta-analysis of neuroimaging studies of motor sequence learning. A literature search allowed us to discern the most commonly used task, the serial reaction time task, and the most commonly used baseline conditions, randomly ordered sequences and previously practiced or known sequences. A first, general meta-analysis confirmed the validity of existing neurobiological models of motor sequence learning (Doyon et al., 2002; Doyon & Benali, 2005; Doyon et al., 2009; Doyon et al., 2018). Subsequent meta-

analyses focusing on the most commonly used task and baseline conditions revealed possible roles of the medial frontal cortex and the cerebellum in error monitoring, cerebellar processing of sequence complexity and visuo-proprioceptive integration in the superior parietal lobule. These results indicate a functional specialization of different areas in learning-related cognitive processes, as well as the importance of taking into account experimental design when interpreting the study results.

In our fMRI study, native French speakers repeatedly produced novel phoneme combinations of different complexity. In Chapter 3, performance monitoring was examined by comparing situations requiring different monitoring loads both prior to and upon articulation. Our main focus was on the regions previously detected in self-monitoring of non-motor aspects of speech production, situated in the medial frontal cortex, the temporo-parietal cortex and the cerebellum (Runnqvist et al., 2021). While different cerebellar areas, including posterior lobules and in particular crus I could be linked to both cognitive and sensorimotor-driven monitoring, the temporo-parietal and the medial frontal cortex were distinctively involved in higher cognitive and sensorimotor-driven monitoring respectively. These activations suggest multiple domain-general and task-specific performance monitoring mechanisms.

In Chapter 4, we assessed internal model refinement through an analysis of the changes in spontaneous brain activity that could emerge as a function of learning progress and motor performance improvement. We were particularly interested in functional networks containing the cerebellum and the basal ganglia, known to be involved in internal models, performance improvement, automatization and learning consolidation (Lisberger, 2010; Fernandez-Seara et al., 2009; Wu, Chan & Hallett, 2010; Debas et al., 2014). Connections between both the cerebellum and the basal ganglia and the rest of the brain changed with repetition-induced motor improvement. Learning progress was positively correlated with connectivity changes between the cerebellar crus I and the precuneus in the parietal cortex, suggesting a role of these regions in internal model refinement.

5.2. Main results

The scope of our first study, reported in Chapter 2, was motor sequence learning beyond speech. Arguably, at least some of the involved cognitive processes were expected to be shared between the two behaviors, as speech articulation is a subset of motor behaviors. Extensive studies of motor sequence learning using neuroimaging

techniques yielded proposals of a variety of neurocognitive models accounting for the acquisition of this skill. In this dissertation, we chose to focus on the model proposed by Doyon and his colleagues (Doyon et al., 2002; Doyon & Benali, 2005; Doyon et al., 2009; Doyon et al., 2018), which acknowledges the importance of the task being performed for the recruitment of a specific brain region during initial learning, as different cognitive processes are thought to be involved contingent on what is being learned. We performed a meta-analysis of neuroimaging studies of motor sequence learning, both in a global fashion, encompassing all the studies meeting our inclusion criteria, and with a focus on subsets of studies that arguably, through their choice of the task or the control condition, may have isolated specific cognitive processes necessary for achieving learning of the novel motor sequences. Concretely, we transformed the foci of the activation yielded by the relevant contrasts of previous studies to probability distributions and used the union of all the probabilities to estimate the likelihood of the activation in any given voxel. As expected, our global meta-analysis revealed a consistent involvement of frontal, medial frontal, parietal and cerebellar regions during initial motor sequence learning, in line with previous research (Doyon et al., 2002; Doyon & Benali, 2005; Doyon et al., 2009; Doyon et al., 2018). Concentrating on each of these areas and their activation or lack thereof when focusing on studies using specific tasks or baseline conditions, we were able to discern which cognitive processes they are likely to sustain during learning. We found evidence of the involvement of the medial frontal cortex in error monitoring, in line with studies on both human and non-human primates and speech production (Fu et al., 2023; Runnqvist et al., 2021). In the parietal cortex, task-general but not baseline-specific implication of the precuneus and the inferior parietal lobule pointed to their role in visual-sensorimotor integration (Doyon et al., 2002; Hardwick et al., 2013), while the superior parietal lobule may be specialized in integrating proprioceptive information (Hardwick et al., 2013). In the basal ganglia, putamen was detected using the random sequence baselines specific to motor sequence learning studies (Janacsek et al., 2020). In the cerebellum, contrasting the novel sequences to those that were known or previously practiced indicated a possible role in error monitoring, online tuning of movements and sequence complexity processing (Grafton et al., 2008; Bernard & Seidler, 2013; Orban et al., 2010; Boecker et al., 2002).

Chapters 3 and 4 report the results of our fMRI study, which focused on motor sequence learning in speech. In Chapter 3, we investigated performance monitoring during the production of novel speech sequences. Prior research has shown that

performance monitoring in speech production engages the right cerebellum, with a differential activation of medial frontal and temporo-parietal areas, as well as the superior medial cerebellum for external but not internal monitoring (Runnqvist et al., 2021). In this study, monitoring of speech production was measured for a pre-articulatory, lexical level of language processing. Previous studies have examined the neural correlates of speech production, focusing on specific articulatory variables and tentatively assigning them particular roles (Segawa et al., 2015, Basilakos et al., 2018). Specifically, these studies found activations related to novelty or articulatory complexity not only in medial frontal and superior temporal cortices and the cerebellum, but also in premotor and parietal cortices and in the basal ganglia. While both parameters can be used to tax monitoring load and therefore indicate a potential role of the detected regions in performance monitoring, these studies either did not take into account production accuracy (Basilakos et al., 2018) or focused only on correctly pronounced syllables (Segawa et al., 2015). Our study was, to the best of our knowledge, the first to directly examine neural correlates of performance monitoring during novel syllable production. We were interested in both internal and external monitoring. Internal monitoring was indexed by the difference between the two conditions of different error probabilities: production of phonotactically illegal and legal sequences, on the correct trials, where the errors were successfully intercepted. External monitoring was indexed by the contrast between errors and correct trials. Neural correlates of the two types of monitoring were examined in predefined regions of interest where we had reasonable expectation to observe activation based on previous studies (Rauschecker et al., 2008; Golfinopoulos et al. 2010; Gauvin et al. 2016; Stoodley & Schmahmann, 2009; Okada & Hickok, 2006). These were located in the cerebellum and in the medial frontal and temporo-parietal cortices. We found differential activation for both types of monitoring in the cerebellar crus I, the superior medial cerebellum and the preSMA in the medial frontal cortex. We also detected the ACC in the medial frontal cortex in external and the pSTG in internal monitoring. These results were complemented by a whole-brain analysis that showed a larger pattern encompassing larger portions of the right cerebellum and the left temporo-parietal, frontal and medial frontal cortex, as well as bilateral activation of the STG for the internal monitoring contrast. Our results indicate that in some regions, such as the cerebellum, the temporal and the medial frontal cortex, both external and internal monitoring rely on the same mechanisms both before and after articulation (e.g., Levelt et al., 1999; Nozari et al., 2011; Pickering & Garrod, 2013), suggesting a more general, and less task or representation-specific

monitoring mechanism.

In Chapter 4, we report the results of our resting state fMRI data analysis. In this chapter, we were interested in the functional brain networks reconfiguration taking place as a function of learning progress that could reflect internal model refinement. Previous research has shown that refining such models involves the cerebellum, but also various cortical regions, depending on what types of representations are being learned (Imamizu et al., 2000). Achieving automatized performance is followed by changes in both the cerebellum and the basal ganglia (Lisberger, 2010). Internal model refinement of how a syllable should be pronounced when the auditory feedback does not match the targeted production has already been studied in a paradigm where the feedback was artificially altered and an online adaptation was necessary (Keough & Jones, 2009). Focusing on singers whose own auditory feedback was altered while they were producing syllables, this study found that both singers and non-singers rely on the external feedback to refine their model of what the syllable should sound like when they move their articulators in a particular manner. While indicating that internal models of syllable production can be refined with practice, this study focused on adaptation rather than on learning and did not aim to reveal the underlying neural correlates of this process. Ours was the first study to capture the neural substrate of this process, in which internal models are being set as learning becomes refined through repeated practice and increased production accuracy. A general increase in production accuracy could be conceptualized as motor improvement; we also aimed to quantify individual learning progress of each participant. To this end, we measured the slope of their learning curve on the production of phonotactically illegal items, whose articulation was completely novel to the participant and had to be learned. Task-induced changes between the spontaneous brain activity of subcortical brain regions and the rest of the brain were found between both the cerebellum and the basal ganglia with medial frontal cortex. Furthermore, we found that the coupling between the right crus I and the left precuneus was positively correlated with the learning progress, stronger coupling predicting greater learning progress. We interpret this coupling as a possible reflection of the cerebellum's increasing capacity to predict the future state of the parietal cortical activity and synchronize with it (Sokolov, Miall & Ivry, 2017). The observed parietal cortical activity could then sustain motor imagery that precedes movement articulation (Blakemore & Sirigu, 2003), leading to improved pronunciation.

5.3. Speech motor sequence learning as skill acquisition at the interface of language and movement

Speech motor sequence learning is a blooming field. Behavioral protocols are being developed to establish the best research practices in the field (Buchwald & Cheng, 2023), and, for the first time, neuroimaging experiments are determining the neural correlates of the acquisition of this skill. The results of our fMRI experiment can be situated within the context of the previous related studies. Both Rauschecker and colleagues' (2008) and Segawa and colleagues' (2015) studies found an effect of novelty in temporal (pSTS/pSTG) and medial frontal (preSMA/SMA) cortices. Rauschecker observed a decrease in cortical and cerebellar activity with the number of repetitions, and Segawa compared novel and learned illegal syllables. We found similar activations in our study, despite the fact that all syllables were novel and the number of repetitions of each syllable was not taken into account in the task-related fMRI analysis. This similarity suggests that the detected regions do not mediate novel syllable processing per se, but rather monitor novel speech. In particular, our internal monitoring contrast, in which we contrasted correctly pronounced illegal and legal syllables, is comparable to Segawa and colleagues' analysis of the main effects of phontactic legality, in which they compared illegal and legal syllables that were already learned *and* correctly pronounced. Both our regions of interest analysis and Segawa's whole brain analysis showed activations in the pSTG and the preSMA. Sequence novelty is an important factor as previous studies found no significant effect in either of the two regions during internal monitoring of known speech production (Runnqvist et al., 2021).

Existing neurocognitive models of speech production and motor control account for the neural correlates of speech motor sequence learning, and our findings can be discussed in the light of these models. In our fMRI experiment, we chose to focus on the areas in medial frontal and temporo-parietal cortices, as well as on the cerebellum and potential functional connections between the cerebellum and the basal ganglia with the rest of the brain. Of particular interest, in the GODIVA model, sequential structure of the upcoming syllable is stored in the preSMA (Bohland & Guenther, 2006). The internal monitoring contrast related to the planning of novel syllables in our study, as reported in Chapter 3, revealed an activation in this area. In combination with the reported lack of activation when syllabic structure is of lesser interest (Runnqvist et al., 2021), this finding supports the role of the preSMA in the storage of sequential

5. *General discussion – 5.3. Speech motor sequence learning as skill acquisition at the interface of language and movement*

information. At the babbling, learning stage, according to the DIVA model, motor and sensory information tune the projections from temporal and parietal sensory error maps to the feedforward control map (Tourville & Guenther, 2011), and cerebellar projections contribute to precisely-timed feedforward commands (Golfinopoulos, Tourville & Guenther, 2010). Our finding that functional connectivity changes between the cerebellum and the parietal cortex correlate with learning progress, reported in Chapter 4, fits well with this view and adds functional connections to the synaptic ones already accounted for in the model.

Finally, motor sequence learning is a well-established field of research, and our body of results shows that it could serve as the basis for speech production studies. For instance, many of the regions involved in learning of speech motor sequences, as shown in our study presented in Chapters 3 and 4, overlap with those underpinning motor sequence learning, as modelled by Doyon and his colleagues (Doyon et al., 2018). In particular, frontal, medial frontal and parietal cortices and the cerebellum seem to support skill acquisition in both domains. It would follow that general, modality-independent processes involved in sequence learning, such as chunking, are sustained by same areas, in the basal ganglia, in both motor and speech learning. In this light, it could seem surprising that the whole brain analysis of our fMRI experiment, reported in Chapter 3, failed to detect activation in the basal ganglia. Looking into our internal monitoring contrast that compared the production of phonotactically illegal sequences to legal ones from the perspective of the meta-analytic approach we took in Chapter 3, we can find a similarity with the contrast in which the sequences that had to be learned are compared to those that were already known or practiced. This similarity is due to the assumption that the pronunciation of the legal sequences, which served as the baseline condition, did not have to be learned, but only improved by practice, as we argued in Chapter 4. Our meta-analysis focusing on the same contrast in manual sequence learning also failed to detect activity in the basal ganglia. A future speech motor sequence learning experiment could benefit from this finding by using a design in which the syllables in the control condition would be constantly changing, akin to our random sequence meta-analysis that did detect activity in the basal ganglia. A further comparison of our whole brain analysis and that particular meta-analysis shows striking similarities, with activations in frontal, medial frontal and parietal cortices and the cerebellum in both cases. Activations in temporal and occipital cortices detected in our fMRI experiment, but not in the comparable meta-analysis, likely reflect the visual and auditory aspects of our paradigm.

Another parallel can be drawn between our experimental task and the SRTT in motor sequence learning studies, both of which showed activations in the frontal, medial frontal and parietal cortices. Indeed, both employ learning paradigms that involve the formation of sequences combining familiar elements in a novel manner. The control condition, on the other hand, consists of sequences whose probability transitions are predictable. In our fMRI experiment, the predictability comes from the fact that the consonant clusters in the sequence are familiar to the participant, as they exist in their native language. In the vast majority of the SRTT studies, baseline sequences have transitional probabilities that are up to 100% predictable (Lum et al., 2018).

5.4. Limitations

The scope of our studies was limited to healthy adults who, during short sessions, learned to pronounce isolated monosyllabic sequences of equal complexity, but different phonotactic probabilities, and containing only phonemes that were familiar to them. This is clearly an oversimplification of real-life learning situations. Our setup involved a single speaker who recorded the stimuli played to the participants, and evaluated their productions for accuracy. Ideally, scores should be provided by different raters, and then compared to improve reliability. Our findings should be considered together with previous behavioral and neuroimaging studies which aimed to gain understanding of the various factors that could impact learning in order to provide a more complete account of this phenomenon. For instance, linguistic parameters such as consonant voicing (Cheng & Buchwald, 2021), word stress and phoneme novelty (Meigh, Cobun & Yunusova, 2020), word frequency or phonological complexity (Quinn, 2021) are important factors in real-life learning situations. Learners' characteristics, such as age (Quinn, 2021) or specific conditions like Parkinson's disease or stuttering, could also impact speech motor sequence learning (Whitfield & Goberman, 2017; Masapollo et al., 2021) and it is therefore important to perform studies on these populations. Finally, our experiment took place in a single day and the learning portion lasted around 30 minutes. This paradigm allowed us to test our hypotheses on performance monitoring and internal model refinement, and we could observe learning. However, both language learning and motor sequence learning literatures suggest the importance of subsequent sleep for learning consolidation and long-term retention (Albouy et al., 2015; 2012; 2008; Boutin et al., 2018, Fernandez & Lüthi, 2020; Thompson et al., 2021; Schimke et al., 2021; March, Ricketts & Tamminen,

2023). Hence, future studies should apply protocols like ours in more longitudinal studies.

5.5. Conclusion

A significant portion of both the motor learning and psycholinguistic literature has addressed how neural and cognitive mechanisms underlying learning and skilled sequential behaviors support our daily actions and communication. As an example of such behavior, pronouncing novel consonant clusters is both a linguistic and a motor skill. Some models, such as GODIVA (Bohland, Bullock & Guenther, 2010), and studies on stuttering (Watkins et al., 2008) or perturbation anticipation (Bonnard, de Graaf & Pailhous, 2004) aimed to bridge the gap between motor and cognitive aspects of common speech production and adaptive behaviors. Despite this, motor programming has often been underspecified in psycholinguistic models, if addressed at all (Miller & Guenther, 2021). When defining a research question, we consider the appropriate theoretical background, and adopt a methodological approach that we consider to be the only possible one in the given situation, or the most common, or the “safest bet”. Our choices, even when we believe they are a result of best practices, may lead us to interpretations that only show one side of the story. The time is ripe for an evolution in research that would bring distinct theoretical and methodological frameworks closer together. As our studies illustrate, when examining speech production and learning, linguists would greatly benefit from working with motion scientists and vice versa, while closely collaborating with technicians and data engineers, at every state of the research process.

5.6. Perspectives

Based on the work carried out in this dissertation, several recommendations could be made for future research. Drawing from both motor learning and speech production literatures, the studies carried out in this dissertation allowed us not only to bridge the two together, but also to better understand the findings in both domains and situate them in a broader context. In this light, results generalizability should be systematically addressed, as different cognitive skills can rely on the same underlying mechanisms. In this light, future research could benefit from integrating research traditions that have been developed independently. Our meta-analyses highlighted

the importance of experimental design choices and the impact they may have on results interpretation. Indeed, an attempt to disentangle a process of interest might shift the focus from the process in question as a whole to its particular subcomponent, while neglecting the others. This bigger picture may also be obtained in future fMRI studies by systematically including resting state acquisitions before and after the experimental task. Their relatively short, task-free nature is easy to implement and provides a large field of possibilities for exploring functional network reconfigurations as a complementary insight to the phenomenon of interest, for instance, through more naturalistic protocols. This addition, combined with open science practices, could then not only enable replications of existing studies, but also make it possible to address future research questions and test hypotheses on large study cohorts and through meta-analyses.

References

- Alario, F. X., Chainay, H., Lehericy, S., & Cohen, L. (2006). The role of the supplementary motor area (SMA) in word production. *Brain research*, 1076(1), 129-143.
- Albarqi, G., & Tavakoli, P. (2022). The effects of proficiency level and dual-task condition on L2 self-monitoring behavior. *Studies in Second Language Acquisition*, 1-22.
- Albouy, G., Fogel, S., King, B. R., Laventure, S., Benali, H., Karni, A., ... & Doyon, J. (2015). Maintaining vs. enhancing motor sequence memories: respective roles of striatal and hippocampal systems. *Neuroimage*, 108, 423-434.
- Albouy, G., King, B. R., Maquet, P., & Doyon, J. (2013). Hippocampus and striatum: Dynamics and interaction during acquisition and sleep-related motor sequence memory consolidation. *Hippocampus*, 23(11), 985-1004.
- Albouy, G., Sterpenich, V., Balteau, E., Vandewalle, G., Desseilles, M., Dang-Vu, T., ... & Maquet, P. (2008). Both the hippocampus and striatum are involved in consolidation of motor sequence memory. *Neuron*, 58(2), 261-272.
- Albouy, G., Sterpenich, V., Vandewalle, G., Darsaud, A., Gais, S., Rauchs, G., ... & Maquet, P. (2012). Neural correlates of performance variability during motor sequence acquisition. *Neuroimage*, 60(1), 324-331.*
- Amaro Jr, E., & Barker, G. J. (2006). Study design in fMRI: basic principles. *Brain and cognition*, 60(3), 220-232.
- Argyropoulos, G. P. (2016). The cerebellum, internal models and prediction in 'non-motor' aspects of language: a critical review. *Brain and language*, 161, 4-17.
- Avants, B.B., C.L. Epstein, M. Grossman, and J.C. Gee. 2008. "Symmetric Diffeomorphic Image Registration with Cross-Correlation: Evaluating Automated Labeling of Elderly and Neurodegenerative Brain." *Medical Image Analysis* 12(1): 26-41.
- Bahrnick, H. P., Bahrnick, L. E., Bahrnick, A. S., & Bahrnick, P. E. (1993). Maintenance of foreign language vocabulary and the spacing effect. *Psychological Science*,

4(5), 316-321.

- Basilakos, A., Smith, K. G., Fillmore, P., Fridriksson, J., & Fedorenko, E. (2018). Functional characterization of the human speech articulation network. *Cerebral Cortex*, 28(5), 1816-1830.
- Behzadi, Y., Restom, K., Liau, J., & Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage*, 37(1), 90-101.
- Berlot, E., Popp, N. J., & Diedrichsen, J. (2020). A critical re-evaluation of fMRI signatures of motor sequence learning. *Elife*, 9, e55241.
- Bernard, J. A., & Seidler, R. D. (2013). Cerebellar contributions to visuomotor adaptation and motor sequence learning: an ALE meta-analysis. *Frontiers in human neuroscience*, 7, 27.**
- Berry, D. C., Berry, D., Dienes, Z., & Dienes, Z. P. (1993). *Implicit learning: Theoretical and empirical issues*. Psychology Press.
- Bischoff-Grethe, A., Goedert, K. M., Willingham, D. T., & Grafton, S. T. (2004). Neural substrates of response-based sequence learning using fMRI. *Journal of cognitive neuroscience*, 16(1), 127-138.*
- Biswal, B., Zerrin Yetkin, F., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic resonance in medicine*, 34(4), 537-541.
- Blakemore, S. J., & Sirigu, A. (2003). Action prediction in the cerebellum and in the parietal lobe. *Experimental Brain Research*, 153(2), 239-245.
- Blakemore, S. J., Frith, C. D., & Wolpert, D. M. (2001). The cerebellum is involved in predicting the sensory consequences of action. *Neuroreport*, 12(9), 1879-1884.
- Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature neuroscience*, 1(7), 635-640.
- Boecker, H., Ceballos-Baumann, A. O., Bartenstein, P., Dagher, A., Forster, K., Haslinger, B., ... & Conrad, B. (2002). A H215O positron emission tomography study on mental imagery of movement sequences—the effect of modulating sequence length and direction. *NeuroImage*, 17(2), 999-1009.
- Bohland, J. W., & Guenther, F. H. (2006). An fMRI investigation of syllable sequence production. *Neuroimage*, 32(2), 821-841.
- Bohland, J. W., Bullock, D., & Guenther, F. H. (2010). Neural representations and mechanisms for the performance of simple speech sequences. *Journal of*

cognitive neuroscience, 22(7), 1504-1529.

- Bonnard, M., de Graaf, J., & Pailhous, J. (2004). Interactions between cognitive and sensorimotor functions in the motor cortex: evidence from the preparatory motor sets anticipating a perturbation. *Reviews in the Neurosciences*, 15(5), 371-382.
- Bosco, G., & Poppele, R. E. (2001). Proprioception from a spinocerebellar perspective. *Physiological reviews*, 81(2), 539-568.
- Boutin, A., Pinsard, B., Boré, A., Carrier, J., Fogel, S. M., & Doyon, J. (2018). Transient synchronization of hippocampo-striato-thalamo-cortical networks during sleep spindle oscillations induces motor memory consolidation. *Neuroimage*, 169, 419-430.
- Boyd, L. A., Edwards, J. D., Siengsukon, C. S., Vidoni, E. D., Wessel, B. D., & Lindsell, M. A. (2009). Motor sequence chunking is impaired by basal ganglia stroke. *Neurobiology of learning and memory*, 92(1), 35-44.
- Brandt, T. H. (2006). How to see what you are looking for in fMRI and PET - or the crucial baseline condition. *Journal of neurology*, 253(5), 551-555.
- Brett, M., Anton, J. L., Valabregue, R., & Poline, J. B. (2002, June). Region of interest analysis using an SPM toolbox. In 8th international conference on functional mapping of the human brain (Vol. 16, No. 2, p. 497).
- Buccino, G., Vogt, S., Ritzl, A., Fink, G. R., Zilles, K., Freund, H. J., & Rizzolatti, G. (2004). Neural circuits underlying imitation learning of hand actions: an event-related fMRI study. *Neuron*, 42(2), 323-334.*
- Buchwald, A., & Cheng, H. S. (2023). Factors Affecting Nonnative Consonant Cluster Learning. *Journal of Speech, Language, and Hearing Research*, 1-14.
- Buckner, R. L. (2013). The cerebellum and cognitive function: 25 years of insight from anatomy and neuroimaging. *Neuron*, 80(3), 807-815.
- Callan, D. E., Jones, J. A., Callan, A. M., & Akahane-Yamada, R. (2004). Phonetic perceptual identification by native-and second-language speakers differentially activates brain regions involved with acoustic phonetic processing and those involved with articulatory-auditory/orosensory internal models. *NeuroImage*, 22(3), 1182-1194.
- Canese, K., & Weis, S. (2013). PubMed: the bibliographic database. *The NCBI handbook*, 2(1).
- Carp, J. (2012). The secret lives of experiments: methods reporting in the fMRI literature. *Neuroimage*, 63(1), 289-300.

- Chanoine, V. & Badier, J-M. (2016). Techniques de neuro-imagerie. In S. Pinto & M. Sato (Eds.), *Traité de neurolinguistique: Du cerveau au langage*(pp. 45–64). De Boeck Supérieur.
- Cheng, H. S., & Buchwald, A. (2021). Does voicing affect patterns of transfer in nonnative cluster learning?. *Journal of Speech, Language, and Hearing Research*, 64(6S), 2103-2120.
- Cholin, J. (2008). The mental syllabary in speech production: An integration of different approaches and domains. *Aphasiology*, 22(11), 1127-1141.
- Civier, O., Tasko, S. M., & Guenther, F. H. (2010). Overreliance on auditory feedback may lead to sound/syllable repetitions: simulations of stuttering and fluency-inducing conditions with a neural model of speech production. *Journal of fluency disorders*, 35(3), 246-279.
- Cleeremans, A., & Dienes, Z. (2008). Computational models of implicit learning. *Cambridge handbook of computational psychology*, 396-421.
- Clower, D. M., West, R. A., Lynch, J. C., & Strick, P. L. (2001). The inferior parietal lobule is the target of output from the superior colliculus, hippocampus, and cerebellum. *Journal of Neuroscience*, 21(16), 6283-6291.
- Cox, R. W., & Hyde, J. S. (1997). Software tools for analysis and visualization of fMRI data. *NMR in Biomedicine: An International Journal Devoted to the Development and Application of Magnetic Resonance In Vivo*, 10(4-5), 171-178.
- Cross, E. S., Schmitt, P. J., & Grafton, S. T. (2007). Neural substrates of contextual interference during motor learning support a model of active preparation. *Journal of cognitive neuroscience*, 19(11), 1854-1871.*
- Dandolo, L. C., & Schwabe, L. (2019). Time-dependent motor memory representations in prefrontal cortex. *Neuroimage*, 197, 143-155.
- Darainy, M., Vahdat, S., & Ostry, D. J. (2019). Neural basis of sensorimotor plasticity in speech motor adaptation. *Cerebral Cortex*, 29(7), 2876-2889.
- Daselaar, S. M., Rombouts, S. A., Veltman, D. J., Raaijmakers, J. G., & Jonker, C. (2003). Similar network activated by young and old adults during the acquisition of a motor sequence. *Neurobiology of aging*, 24(7), 1013-1019.*
- de Long, M. R., & Alexander, G. E. (1987). The basal ganglia and sensorimotor integration. *Clinical Aspects of Sensory Motor Integration*, 203-211.
- de Zeeuw, C. I., & Ten Brinke, M. M. (2015). Motor learning and the cerebellum. *Cold Spring Harbor perspectives in biology*, 7(9), a021683.

- de Zubicaray, G. I., Williams, S. C., Wilson, S. J., Rose, S. E., Brammer, M. J., Bullmore, E. T., ... & Doddrell, D. M. (1998). Prefrontal cortex involvement in selective letter generation: a functional magnetic resonance imaging study. *Cortex*, 34(3), 389-401.
- Debaere, F., Wenderoth, N., Sunaert, S., Van Hecke, P., & Swinnen, S. P. (2004). Changes in brain activation during the acquisition of a new bimanual coordination task. *Neuropsychologia*, 42(7), 855-867.*
- Debas, K., Carrier, J., Barakat, M., Marrelec, G., Bellec, P., Tahar, A. H., ... & Doyon, J. (2014). Off-line consolidation of motor sequence learning results in greater integration within a cortico-striatal functional network. *Neuroimage*, 99, 50-58.
- Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of memory and language*, 27(2), 124-142.
- Dennis, N. A., & Cabeza, R. (2011). Age-related dedifferentiation of learning systems: an fMRI study of implicit and explicit learning. *Neurobiology of aging*, 32(12), 2318-e17.*
- Desmurget, M., & Grafton, S. (2000). Forward modeling allows feedback control for fast reaching movements. *Trends in cognitive sciences*, 4(11), 423-431.
- Diedrichsen, J., & Kornysheva, K. (2015). Motor skill learning between selection and execution. *Trends in cognitive sciences*, 19(4), 227-233.
- dos Santos Palma, G. C., Correa, U. C., & Torriani-Pasin, C. (2020). The differentiated effect of the task complexity on retention and transfer of stroke survivors. *Human movement science*, 69, 102545.
- Dosenbach, N. U., Koller, J. M., Earl, E. A., Miranda-Dominguez, O., Klein, R. L., Van, A. N., ... & Fair, D. A. (2017). Real-time motion analytics during brain MRI improve data quality and reduce costs. *Neuroimage*, 161, 80-93.
- Doyon, J., & Benali, H. (2005). Reorganization and plasticity in the adult brain during learning of motor skills. *Current opinion in neurobiology*, 15(2), 161-167.
- Doyon, J., Bellec, P., Amsel, R., Penhune, V., Monchi, O., Carrier, J., ... & Benali, H. (2009). Contributions of the basal ganglia and functionally related brain structures to motor learning. *Behavioural brain research*, 199(1), 61-75.
- Doyon, J., Gabbitov, E., Vahdat, S., Lungu, O., & Boutin, A. (2018). Current issues related to motor sequence learning in humans. *Current Opinion in Behavioral Sciences*, 20, 89-97.

- Doyon, J., Orban, P., Barakat, M., Debas, K., Lungu, O., Albouy, G., ... & Benali, H. (2011). Functional brain plasticity associated with motor learning. *Medecine Sciences: M/S*, 27(4), 413-420.
- Doyon, J., Owen, A. M., Petrides, M., Sziklas, V., & Evans, A. C. (1996). Functional anatomy of visuomotor skill learning in human subjects examined with positron emission tomography. *European Journal of Neuroscience*, 8(4), 637-648.*
- Doyon, J., Penhune, V., & Ungerleider, L. G. (2003). Distinct contribution of the cortico-striatal and cortico-cerebellar systems to motor skill learning. *Neuropsychologia*, 41(3), 252-262.
- Doyon, J., Song, A. W., Karni, A., Lalonde, F., Adams, M. M., & Ungerleider, L. G. (2002). Experience-dependent changes in cerebellar contributions to motor sequence learning. *Proceedings of the National Academy of Sciences*, 99(2), 1017-1022.
- Doyon, J., Ungerleider, L. G., Squire, L., & Schacter, D. (2002). Functional anatomy of motor skill learning. *Neuropsychology of memory*, 3, 225-238.
- Du, Y., & Clark, J. E. (2018). The “Motor” in Implicit Motor Sequence Learning: A Footstepping Serial Reaction Time Task. *Journal of Visualized Experiments: JoVE*, 135.
- Duyck, W., Desmet, T., Verbeke, L. P., & Brysbaert, M. (2004). WordGen: A tool for word selection and nonword generation in Dutch, English, German, and French. *Behavior Research Methods, Instruments, & Computers*, 36(3), 488-499.
- Eickhoff, S. B., Bzdok, D., Laird, A. R., Kurth, F., & Fox, P. T. (2012). Activation likelihood estimation meta-analysis revisited. *Neuroimage*, 59(3), 2349-2361.
- Eickhoff, S. B., Nichols, T. E., Laird, A. R., Hoffstaedter, F., Amunts, K., Fox, P. T., ... & Eickhoff, C. R. (2016). Behavior, sensitivity, and power of activation likelihood estimation characterized by massive empirical simulation. *Neuroimage*, 137, 70-85.
- Esteban, O., Blair, R. W., Markiewicz, C. J., Berleant, S. L., Moodie, C., Ma, F., Isik, A. I. et al. 2018. “*fMRIPrep*.” *Software*. Zenodo.
- Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., ... & Gorgolewski, K. J. (2019). fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature methods*, 16(1), 111-116.
- Ettinger, U., Corr, P. J., Mofidi, A., Williams, S. C., & Kumari, V. (2013). Dopaminergic basis of the psychosis-prone personality investigated with functional

- magnetic resonance imaging of procedural learning. *Frontiers in human neuroscience*, 7, 130.*
- Fairs, A., Michelas, A., Dufour, S., & Strijkers, K. (2021). The same ultra-rapid parallel brain dynamics underpin the production and perception of speech. *Cerebral Cortex Communications*, 2(3), tgab040.
- Faivre, N., Vuillaume, L., Bernasconi, F., Salomon, R., Blanke, O., & Cleeremans, A. (2020). Sensorimotor conflicts alter metacognitive and action monitoring. *Cortex*, 124, 224-234.
- Favaro, L., Gamba, M., Cresta, E., Fumagalli, E., Bandoli, F., Pilenga, C., ... & Reby, D. (2020). Do penguins' vocal sequences conform to linguistic laws?. *Biology letters*, 16(2), 20190589.
- Fernández-Seara, M. A., Aznárez-Sanado, M., Mengual, E., Loayza, F. R., & Pastor, M. A. (2009). Continuous performance of a novel motor sequence leads to highly correlated striatal and hippocampal perfusion increases. *NeuroImage*, 47(4), 1797-1808.
- Fernandez, L. M., & Lüthi, A. (2020). Sleep spindles: mechanisms and functions. *Physiological reviews*, 100(2), 805-868.
- Fiez, J. A. (1996). Cerebellar contributions to cognition. *Neuron*, 16(1), 13-15.
- Fischer, S., Nitschke, M. F., Melchert, U. H., Erdmann, C., & Born, J. (2005). Motor memory consolidation in sleep shapes more effective neuronal representations. *Journal of Neuroscience*, 25(49), 11248-11255.*
- Flament, D., Ellermann, J. M., Kim, S. G., Uğurbil, K., & Ebner, T. J. (1996). Functional magnetic resonance imaging of cerebellar activation during the learning of a visuomotor dissociation task. *Human brain mapping*, 4(3), 210-226.
- Floyer-Lea, A., & Matthews, P. M. (2005). Distinguishable brain activation networks for short-and long-term motor skill learning. *Journal of neurophysiology*, 94(1), 512-518.*
- Fonov, V. S., Evans, A. C., McKinstry, R. C., Almlí, C. R., & Collins, D. L. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*, 47, S102.
- Ford, J. M., & Mathalon, D. H. (2012). Anticipating the future: automatic prediction failures in schizophrenia. *International Journal of Psychophysiology*, 83(2), 232-239.
- Fox, P. T., Laird, A. R., Eickhoff, S. B., Lancaster, J. L., Fox, M., Uecker, A. M., & Ray, K. L. (2013). User manual for GingerALE 2.3. *San Antonio, TX: UT Health Science*

Center San Antonio.

- Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D., & Frackowiak, R. S. (1994). Statistical parametric maps in functional imaging: a general linear approach. *Human brain mapping, 2*(4), 189-210.
- Fu, Z., Sajad, A., Errington, S. P., Schall, J. D., & Rutishauser, U. (2023). Neurophysiological mechanisms of error monitoring in human and non-human primates. *Nature Reviews Neuroscience, 1-20.*
- Fuentes-García, J. P., Pulido, S., Morales, N., & Menayo, R. (2022). Massed and distributed practice on learning the forehand shot in tennis. *International Journal of Sports Science & Coaching, 17*(2), 318-324.
- Gabitov, E., Manor, D., & Karni, A. (2016). Learning from the other limb's experience: sharing the 'trained' M1 representation of the motor sequence knowledge. *The Journal of physiology, 594*(1), 169-188.*
- Gauvin, H. S., & Hartsuiker, R. J. (2020). Towards a new model of verbal monitoring. *Journal of Cognition, 3*(1).
- Gauvin, H. S., De Baene, W., Brass, M., & Hartsuiker, R. J. (2016). Conflict monitoring in speech processing: An fMRI study of error detection in speech production and perception. *NeuroImage, 126*, 96-105.
- Gharbawie, O. A., Karl, J. M., & Whishaw, I. Q. (2007). Recovery of skilled reaching following motor cortex stroke: do residual corticofugal fibers mediate compensatory recovery?. *European Journal of Neuroscience, 26*(11), 3309-3327.
- Gheysen, F., Van Opstal, F., Roggeman, C., Van Waelvelde, H., & Fias, W. (2010). Hippocampal contribution to early and later stages of implicit motor sequence learning. *Experimental brain research, 202*(4), 795-807.
- Glickstein, M. (1992). The cerebellum and motor learning. *Current opinion in neurobiology, 2*(6), 802-806.
- Glickstein, M. (2000). How are visual areas of the brain connected to motor areas for the sensory guidance of movement?. *Trends in neurosciences, 23*(12), 613-617.
- Gobel, E. W., Parrish, T. B., & Reber, P. J. (2011). Neural correlates of skill acquisition: decreased cortical activity during a serial interception sequence learning task. *Neuroimage, 58*(4), 1150-1157.*
- Gobet, F., Lane, P. C., Croker, S., Cheng, P. C., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in cognitive sciences,*

5(6), 236-243.

- Goh, H. T., Lee, Y. Y., & Fisher, B. E. (2013). Neural correlates of dual-task practice benefit on motor learning: a repetitive transcranial magnetic stimulation study. *European Journal of Neuroscience*, 37(11), 1823-1829.
- Goh, H. T., Sullivan, K. J., Gordon, J., Wulf, G., & Winstein, C. J. (2012). Dual-task practice enhances motor learning: a preliminary investigation. *Experimental Brain Research*, 222, 201-210.
- Golfinopoulos, E., Tourville, J. A., & Guenther, F. H. (2010). The integration of large-scale neural network modeling and functional brain imaging in speech motor control. *Neuroimage*, 52(3), 862-874.
- Gonzalez, C. C., & Burke, M. R. (2018). Motor sequence learning in the brain: The long and short of it. *Neuroscience*, 389, 85-98.*
- Gorgolewski, K., C. D. Burns, C. Madison, D. Clark, Y. O. Halchenko, M. L. Waskom, and S. Ghosh. 2011. "Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python." *Frontiers in Neuroinformatics* 5: 13.
- Gorgolewski, K., Esteban, O., Markiewicz, C. J., Ziegler, E., Ellis, D. G., Notter, M. P., Jarecka, D. et al. 2018. "Nipype." Software. Zenodo.
- Grafton, S. T., Hazeltine, E., & Ivry, R. (1995). Functional mapping of sequence learning in normal humans. *Journal of cognitive neuroscience*, 7(4), 497-510.*
- Grafton, S. T., Hazeltine, E., & Ivry, R. B. (1998). Abstract and effector-specific representations of motor sequences identified with PET. *Journal of Neuroscience*, 18(22), 9420-9428.*
- Grafton, S. T., Hazeltine, E., & Ivry, R. B. (2002). Motor sequence learning with the nondominant left hand. *Experimental Brain Research*, 146(3), 369-378.*
- Grafton, S. T., Salidis, J., & Willingham, D. B. (2001). Motor learning of compatible and incompatible visuomotor maps. *Journal of Cognitive Neuroscience*, 13(2), 217-231.*
- Grafton, S. T., Schmitt, P., Van Horn, J., & Diedrichsen, J. (2008). Neural substrates of visuomotor learning based on improved feedback control and prediction. *Neuroimage*, 39(3), 1383-1395.
- Graybiel, A. M. (1998). The basal ganglia and chunking of action repertoires. *Neurobiology of learning and memory*, 70(1-2), 119-136.
- Graybiel, A. M., Aosaki, T., Flaherty, A. W., & Kimura, M. (1994). The basal ganglia and adaptive motor control. *Science*, 265(5180), 1826-1831.

- Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *Neuroimage*, 48(1), 63-72.
- Guell, X., Schmahmann, J. D., Gabrieli, J. D., & Ghosh, S. S. (2018). Functional gradients of the cerebellum. *Elife*, 7, e36652.
- Guenther, F. H. (1994). A neural network model of speech acquisition and motor equivalent speech production. *Biological cybernetics*, 72(1), 43-53.
- Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of communication disorders*, 39(5), 350-365.
- Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and language*, 96(3), 280-301.
- Haaland, K. Y., Elsinger, C. L., Mayer, A. R., Durgerian, S., & Rao, S. M. (2004). Motor sequence complexity and performing hand produce differential patterns of hemispheric lateralization. *Journal of cognitive neuroscience*, 16(4), 621-636.*
- Hansen, S. J., McMahon, K. L., & de Zubicaray, G. I. (2019). Neural mechanisms for monitoring and halting of spoken word production. *Journal of Cognitive Neuroscience*, 31(12), 1946-1957.
- Hardwick, R. M., Rottschy, C., Miall, R. C., & Eickhoff, S. B. (2013). A quantitative meta-analysis and review of motor learning in the human brain. *Neuroimage*, 67, 283-297.**
- Hartsuiker, R. J., & Kolk, H. H. (2001). Error monitoring in speech production: A computational test of the perceptual loop theory. *Cognitive psychology*, 42(2), 113-157.
- Haswell, C. C., Izawa, J., Dowell, L. R., Mostofsky, S. H., & Shadmehr, R. (2009). Representation of internal models of action in the autistic brain. *Nature neuroscience*, 12(8), 970-972.
- Hayes, H. A., Hunsaker, N., & Dibble, L. E. (2015). Implicit motor sequence learning in individuals with Parkinson disease: a meta-analysis. *Journal of Parkinson's disease*, 5(3), 549-560.
- Hechavarría, J. C., Beetz, M. J., Macias, S., & Kössl, M. (2016). Distress vocalization sequences broadcasted by bats carry redundant information. *Journal of Comparative Physiology A*, 202, 503-515.
- Hemond, C., Brown, R. M., & Robertson, E. M. (2010). A distraction can impair or enhance motor performance. *Journal of Neuroscience*, 30(2), 650-654.

- Hermer-Vazquez, L., & Moshtagh, N. (2009). Rats' learning of a new motor skill: insight into the evolution of motor sequence learning. *Behavioural processes*, 81(1), 50-59.
- Heun, R., Freymann, N., Granath, D. O., Stracke, C. P., Jessen, F., Barkow, K., & Reul, J. (2004). Differences of cerebral activation between superior and inferior learners during motor sequence encoding and retrieval. *Psychiatry Research: Neuroimaging*, 132(1), 19-32.*
- Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature reviews neuroscience*, 13(2), 135-145.
- Hickok, G. (2014). The architecture of speech production and the role of the phoneme in speech processing. *Language, Cognition and Neuroscience*, 29(1), 2-20.
- Hikosaka, O. (1991). Basal ganglia—possible role in motor coordination and learning. *Current opinion in neurobiology*, 1(4), 638-643.
- Hikosaka, O., Nakamura, K., Sakai, K., & Nakahara, H. (2002). Central mechanisms of motor skill learning. *Current opinion in neurobiology*, 12(2), 217-222.
- Hirano, S., Kojima, H., Naito, Y., Honjo, I., Kamoto, Y., Okazawa, H., ... & Konishi, J. (1997). Cortical processing mechanism for vocalization with auditory verbal feedback. *Neuroreport*, 8(9), 2379-2382.
- Houde, J. F., & Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science*, 279(5354), 1213-1216.
- Houk, J. C., Buckingham, J. T., & Barto, A. G. (1996). Models of the cerebellum and motor learning. *Behavioral and brain sciences*, 19(3), 368-383.
- Imamizu, H., Miyauchi, S., Tamada, T., Sasaki, Y., Takino, R., Pütz, B., ... & Kawato, M. (2000). Human cerebellar activity reflecting an acquired internal model of a new tool. *Nature*, 403(6766), 192-195.
- Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition*, 92(1-2), 101-144.
- Ito, M. (2000). Mechanisms of motor learning in the cerebellum. *Brain research*, 886(1-2), 237-245.
- Ito, M. (2008). Control of mental activities by internal models in the cerebellum. *Nature Reviews Neuroscience*, 9(4), 304-313.
- Ito, M. (2013). Error detection and representation in the olivo-cerebellar system. *Frontiers in neural circuits*, 7, 1.
- Izawa, J., Pekny, S. E., Marko, M. K., Haswell, C. C., Shadmehr, R., & Mostofsky, S. H. (2012). Motor learning relies on integrated sensory inputs in ADHD, but

- over-selectively on proprioception in autism spectrum conditions. *Autism research*, 5(2), 124-136.
- Janacsek, K., Shattuck, K. F., Tagarelli, K. M., Lum, J. A., Turkeltaub, P. E., & Ullman, M. T. (2020). Sequence learning in the human brain: A functional neuroanatomical meta-analysis of serial reaction time studies. *NeuroImage*, 207, 116387.**
- Jenkinson, M., & Smith, S. (2001). A global optimization method for robust affine registration of brain images. *Medical image analysis*, 5(2), 143-156.
- Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*, 17(2), 825-841.
- Jezzard, P., & Balaban, R. S. (1995). Correction for geometric distortion in echo planar images from B0 field variations. *Magnetic resonance in medicine*, 34(1), 65-73.
- Jiménez, L., Vaquero, J. M., & Lupiáñez, J. (2006). Qualitative differences between implicit and explicit sequence learning. *Journal of experimental psychology: Learning, Memory, and Cognition*, 32(3), 475.
- Jolles, D. D., van Buchem, M. A., Crone, E. A., & Rombouts, S. A. (2013). Functional brain connectivity at rest changes after working memory training. *Human brain mapping*, 34(2), 396-406.
- Jordan, M. I., & Rumelhart, D. E. (1992). Forward models: Supervised learning with a distal teacher. *Cognitive science*, 16(3), 307-354.
- Jouen, A. L., Verwey, W. B., Van Der Helden, J., Scheiber, C., Neveu, R., Dominey, P. F., & Ventre-Dominey, J. (2013). Discrete sequence production with and without a pause: the role of cortex, basal ganglia, and cerebellum. *Frontiers in human neuroscience*, 7, 492.*
- Kaipa, R. (2016). Is there an interaction between task complexity and practice variability in speech-motor learning?. *Annals of Neurosciences*, 23(3), 134-138.
- Kaipa, R., Howard, B., Kaipa, R., Turcat, E., & Prema, L. (2020). Role of massed versus distributed practice in learning novel foreign language utterances. *Motor Control*, 24(1), 17-38.
- Kal, E., Prosée, R., Winters, M., & Van Der Kamp, J. (2018). Does implicit motor learning lead to greater automatization of motor skills compared to explicit motor learning? A systematic review. *PloS one*, 13(9), e0203591.

- Karim, H. T., Huppert, T. J., Erickson, K. I., Wollam, M. E., Sparto, P. J., Sejdić, E., & VanSwearingen, J. M. (2017). Motor sequence learning-induced neural efficiency in functional brain connectivity. *Behavioural brain research*, 319, 87-95.*
- Katz, D. I., & Lackner, J. R. (1977). Adaptation to delayed auditory feedback. *Perception & Psychophysics*, 22(5), 476-486.
- Kawashima, R., Matsumura, M., Sadato, N., Naito, E., Waki, A., Nakamura, S., ... & Yonekura, Y. (1998a). Regional cerebral blood flow changes in human brain related to ipsilateral and contralateral complex hand movements—a PET study. *European Journal of Neuroscience*, 10(7), 2254-2260.*
- Kawashima, R., Tanji, J., Okada, K., Sugiura, M., Sato, K., Kinomura, S., ... & Fukuda, H. (1998b). Oculomotor sequence learning: a positron emission tomography study. *Experimental Brain Research*, 122(1), 1-8.*
- Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current opinion in neurobiology*, 9(6), 718-727.
- Kawato, M., Furukawa, K., & Suzuki, R. (1987). A hierarchical neural-network model for control and learning of voluntary movement. *Biological cybernetics*, 57(3), 169-185.
- Kearney, E., Nieto-Castañón, A., Weerathunge, H. R., Falsini, R., Daliri, A., Abur, D., ... & Guenther, F. H. (2020). A simple 3-parameter model for examining adaptation in speech and voice production. *Frontiers in Psychology*, 10, 2995.
- Keele, S. W., Ivry, R., Mayr, U., Hazeltine, E., & Heuer, H. (2003). The cognitive and neural architecture of sequence representation. *Psychological review*, 110(2), 316.
- Kennedy, D., Haselgrove, C., Fischl, B., Breeze, J., Frazier, J., Seidman, L., ... & Kosofsky, B. (2016). Harvard-Oxford cortical and subcortical structural atlases. *Harvard Center for Morphometric Analysis*.
- Kennerley, S. W., Sakai, K., & Rushworth, M. F. S. (2004). Organization of action sequences and the role of the pre-SMA. *Journal of neurophysiology*, 91(2), 978-993.
- Keough, D., & Jones, J. A. (2009). The sensitivity of auditory-motor representations to subtle changes in auditory feedback while singing. *The Journal of the Acoustical Society of America*, 126(2), 837-846.

- Kincses, Z. T., Johansen-Berg, H., Tomassini, V., Bosnell, R., Matthews, P. M., & Beckmann, C. F. (2008). Model-free characterization of brain functional networks for motor sequence learning using fMRI. *Neuroimage*, 39(4), 1950-1958.*
- King, B. R., Fogel, S. M., Albouy, G., & Doyon, J. (2013). Neural correlates of the age-related changes in motor sequence learning and motor adaptation in older adults. *Frontiers in human neuroscience*, 7, 142.
- King, B. R., Hoedlmoser, K., Hirschauer, F., Dolfen, N., & Albouy, G. (2017). Sleeping on the motor engram: the multifaceted nature of sleep-related motor memory consolidation. *Neuroscience & Biobehavioral Reviews*, 80, 1-22.
- Knolle, F., Schröger, E., & Kotz, S. A. (2013). Cerebellar contribution to the prediction of self-initiated sounds. *Cortex*, 49(9), 2449-2461.
- Konovalov, A., & Krajbich, I. (2018). Neurocomputational dynamics of sequence learning. *Neuron*, 98(6), 1282-1293.*
- Kotz, S. A., & Schwartz, M. (2010). Cortical speech processing unplugged: a timely subcortico-cortical framework. *Trends in cognitive sciences*, 14(9), 392-399.
- Kotz, S. A., Schwartz, M., & Schmidt-Kassow, M. (2009). Non-motor basal ganglia functions: A review and proposal for a model of sensory predictability in auditory language perception. *Cortex*, 45(8), 982-990.
- Kumari, V., Gray, J. A., Honey, G. D., Soni, W., Bullmore, E. T., Williams, S. C., ... & Sharma, T. (2002). Procedural learning in schizophrenia: a functional magnetic resonance imaging investigation. *Schizophrenia research*, 57(1), 97-107.*
- Kwon, Y. H., Kwon, J. W., & Lee, M. H. (2015). Effectiveness of motor sequential learning according to practice schedules in healthy adults; distributed practice versus massed practice. *Journal of physical therapy science*, 27(3), 769-772.
- Laird, A. R., Fox, P. T., Eickhoff, S. B., Lancaster, J. L., Fox, M., & Uecker, A. M. (2011). User Manual for GingerALE 2.1. Verfügbar unter <http://www.brainmap.org/ale/manual.pdf>.
- Laird, A. R., Lancaster, J. L., Fox, M., & Uecker, A. M. (2003). Users' Manual for Brain-Map GingerALE 2.0.
- Lancaster, J. L., Tordesillas-Gutiérrez, D., Martínez, M., Salinas, F., Evans, A., Zilles, K., ... & Fox, P. T. (2007). Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. *Human brain mapping*,

28(11), 1194-1205.

- Lanczos, C. 1964. "Evaluation of Noisy Data." *Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis* 1(1): 76–85.
- Lee, T. D., & Genovese, E. D. (1988). Distribution of practice in motor skill acquisition: Learning and performance effects reconsidered. *Research Quarterly for exercise and Sport*, 59(4), 277-287.**
- Lehéricy, S., Benali, H., Van de Moortele, P. F., Péligrini-Issac, M., Waechter, T., Ugurbil, K., & Doyon, J. (2005). Distinct basal ganglia territories are engaged in early and advanced motor sequence learning. *Proceedings of the National Academy of Sciences*, 102(35), 12566-12571.
- Lesage, E., Hansen, P. C., & Miall, R. C. (2017). Right lateral cerebellum represents linguistic predictability. *Journal of Neuroscience*, 37(26), 6231-6241.
- Levelt, W. J. (1983). Monitoring and self-repair in speech. *Cognition*, 14(1), 41-104.
- Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and brain sciences*, 22(1), 1-38.
- Lin, C. H. J., Knowlton, B. J., Chiang, M. C., Iacoboni, M., Udompholkul, P., & Wu, A. D. (2011). Brain–behavior correlates of optimizing learning through interleaved practice. *Neuroimage*, 56(3), 1758-1772.*
- Lisberger, S. G. (2010). Visual guidance of smooth-pursuit eye movements: sensation, action, and what happens in between. *Neuron*, 66(4), 477-491.
- Llinás, R., & Welsh, J. P. (1993). On the cerebellum and motor learning. *Current opinion in neurobiology*, 3(6), 958-965.
- Loh, K. K., Procyk, E., Neveu, R., Lamberton, F., Hopkins, W. D., Petrides, M., & Amiez, C. (2020). Cognitive control of orofacial motor and vocal responses in the ventrolateral and dorsomedial human frontal cortex. *Proceedings of the National Academy of Sciences*, 117(9), 4994-5005.
- Lohse, K. R., Wadden, K., Boyd, L. A., & Hodges, N. J. (2014). Motor skill acquisition across short and long time scales: a meta-analysis of neuroimaging data. *Neuropsychologia*, 59, 130-141.
- Loonis, R. F., Brincat, S. L., Antzoulatos, E. G., & Miller, E. K. (2017). A meta-analysis suggests different neural correlates for implicit and explicit learning. *Neuron*, 96(2), 521-534.**
- Luft, A. R., & Buitrago, M. M. (2005). Stages of motor skill learning. *Molecular neurobiology*, 32, 205-216.

- Lum, J. A., Mills, A., Plumridge, J. M., Sloan, N. P., Clark, G. M., Hedenius, M., & Enticott, P. G. (2018). Transcranial direct current stimulation enhances retention of a second (but not first) order conditional visuo-motor sequence. *Brain and cognition, 127*, 34-41.
- Lungu, O., Monchi, O., Albouy, G., Jubault, T., Ballarin, E., Burnod, Y., & Doyon, J. (2014). Striatal and hippocampal involvement in motor sequence chunking depends on the learning strategy. *PLoS One, 9*(8), e103885.
- MacIntosh, B. J., Mraz, R., McIlroy, W. E., & Graham, S. J. (2007). Brain activity during a motor learning task: An fMRI and skin conductance study. *Human brain mapping, 28*(12), 1359-1367.
- March, J. A., Ricketts, J., & Tamminen, J. (2023). Is word learning capacity restored after a daytime nap?. *Cortex, 159*, 142-166.
- Martin, C. D., Niziolek, C. A., Duñabeitia, J. A., Perez, A., Hernandez, D., Carreiras, M., & Houde, J. F. (2018). Online adaptation to altered auditory feedback is predicted by auditory acuity and not by domain-general executive control resources. *Frontiers in Human Neuroscience, 12*, 91.
- Masapollo, M., Segawa, J. A., Beal, D. S., Tourville, J. A., Nieto-Castañón, A., Heyne, M., ... & Guenther, F. H. (2021). Behavioral and neural correlates of speech motor sequence learning in stuttering and neurotypical speakers: an fMRI investigation. *Neurobiology of Language, 2*(1), 106-137.
- McCloskey, D. I. (1981). *Corollary discharge and motor commands and perception. Handbook of physiology.*
- Medina, J. F., & Mauk, M. D. (2000). Computer simulation of cerebellar information processing. *Nature neuroscience, 3*(11), 1205-1211.
- Meekings, S., & Scott, S. K. (2021). Error in the superior temporal gyrus? A systematic review and activation likelihood estimation meta-analysis of speech production studies. *Journal of Cognitive Neuroscience, 33*(3), 422-444.
- Meigh, K. M., Cobun, E., & Yunusova, Y. (2020). Phoneme and stress programming interact during nonword repetition learning. *Journal of Speech, Language, and Hearing Research, 63*(7), 2219-2228.
- Mennes, M., Kelly, C., Zuo, X. N., Di Martino, A., Biswal, B. B., Castellanos, F. X., & Milham, M. P. (2010). Inter-individual differences in resting-state functional connectivity predict task-induced BOLD activity. *Neuroimage, 50*(4), 1690-1701.

- Metalis, S. A. (1985). Effects of massed versus distributed practice on acquisition of video game skill. *Perceptual and Motor Skills*, 61(2), 457-458.
- Middleton, F. A., & Strick, P. L. (1994). Anatomical evidence for cerebellar and basal ganglia involvement in higher cognitive function. *Science*, 266(5184), 458-461.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, 63(2), 81.
- Miller, H. E., & Guenther, F. H. (2021). Modelling speech motor programming and apraxia of speech in the DIVA/GODIVA neurocomputational framework. *Aphasiology*, 35(4), 424-441.
- Moberget, T., & Ivry, R. B. (2016). Cerebellar contributions to motor control and language comprehension: searching for common computational principles. *Annals of the New York Academy of Sciences*, 1369(1), 154-171.
- Moberget, T., Gullesen, E. H., Andersson, S., Ivry, R. B., & Endestad, T. (2014). Generalized role for the cerebellum in encoding internal models: evidence from semantic processing. *Journal of Neuroscience*, 34(8), 2871-2878.
- Moeller, S., Yacoub, E., Olman, C. A., Auerbach, E., Strupp, J., Harel, N., & Ugurbil, K. (2010). Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magnetic resonance in medicine*, 63(5), 1144-1153.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine*, 151(4), 264-269.
- Monti, M. M. (2011). Statistical analysis of fMRI time-series: a critical review of the GLM approach. *Frontiers in human neuroscience*, 5, 28.
- Motley, M. T., Camden, C. T., & Baars, B. J. (1982). Covert formulation and editing of anomalies in speech production: Evidence from experimentally elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior*, 21(5), 578-594.
- Müller, P., Rehfeld, K., Schmicker, M., Hökelmann, A., Dordevic, M., Lessmann, V., ... & Müller, N. G. (2017). Evolution of neuroplasticity in response to physical activity in old age: the case for dancing. *Frontiers in aging neuroscience*, 9, 56.
- Müller, R. A., Cauich, C., Rubio, M. A., Mizuno, A., & Courchesne, E. (2004). Abnormal activity patterns in premotor cortex during sequence learning in autistic

- patients. *Biological psychiatry*, 56(5), 323-332.*
- Müller, R. A., Kleinhans, N., Kemmotsu, N., Pierce, K., & Courchesne, E. (2003). Abnormal variability and distribution of functional maps in autism: an fMRI study of visuomotor learning. *American Journal of Psychiatry*, 160(10), 1847-1862.*
- Müller, R. A., Kleinhans, N., Pierce, K., Kemmotsu, N., & Courchesne, E. (2002). Functional MRI of motor sequence acquisition: effects of learning stage and performance. *Cognitive Brain Research*, 14(2), 277-293.*
- Müller, V. I., Cieslik, E. C., Laird, A. R., Fox, P. T., Radua, J., Mataix-Cols, D., ... & Eickhoff, S. B. (2018). Ten simple rules for neuroimaging meta-analysis. *Neuroscience & Biobehavioral Reviews*, 84, 151-161.
- Nair, D. G., Purcott, K. L., Fuchs, A., Steinberg, F., & Kelso, J. S. (2003). Cortical and cerebellar activity of the human brain during imagined and executed unimanual and bimanual action sequences: a functional MRI study. *Cognitive brain research*, 15(3), 250-260.
- Naismith, S. L., Lagopoulos, J., Ward, P. B., Davey, C. G., Little, C., & Hickie, I. B. (2010). Fronto-striatal correlates of impaired implicit sequence learning in major depression: an fMRI study. *Journal of affective disorders*, 125(1-3), 256-261.*
- Nejati, V., Garusi Farshi, M. T., Ashayeri, H., & Aghdasi, M. T. (2008). Dual task interference in implicit sequence learning by young and old adults. *International journal of geriatric psychiatry*, 23(8), 801-804.
- New, B., & Spinelli, E. (2013). Diphones-fr: A French database of diphone positional frequency. *Behavior research methods*, 45(3), 758-764.
- New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet : LEXIQUETM // A lexical database for contemporary french : LEXIQUETM. *L'année psychologique*, 101(3), 447-462.
- Nichols, T., & Hayasaka, S. (2003). Controlling the familywise error rate in functional neuroimaging: a comparative review. *Statistical methods in medical research*, 12(5), 419-446.
- Ninokura, Y., Mushiake, H., & Tanji, J. (2004). Integration of temporal order and object information in the monkey lateral prefrontal cortex. *Journal of neurophysiology*, 91(1), 555-560.
- Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive psychology*, 19(1), 1-32.

- Nooteboom, S. G., & Quené, H. (2017). Self-monitoring for speech errors: Two-stage detection and repair with and without auditory feedback. *Journal of Memory and Language*, *95*, 19-35.
- Nozari, N., Dell, G. S., & Schwartz, M. F. (2011). Is comprehension necessary for error detection? A conflict-based account of monitoring in speech production. *Cognitive psychology*, *63*(1), 1-33.
- Nyberg, L., Eriksson, J., Larsson, A., & Marklund, P. (2006). Learning by doing versus learning by thinking: an fMRI study of motor and mental training. *Neuropsychologia*, *44*(5), 711-717.*
- Nyberg, L., Larsson, A., Eriksson, J., Birgander, R., Sundström, T., & Riklund Åhlström, K. (2007). Comparing 1, 5T and 3T BOLD fMRI imaging of finger tapping with familiar and novel sequences. *Neuroscience Imaging*, *2*(1), 53-64.*
- O'Reilly, J. X., Beckmann, C. F., Tomassini, V., Ramnani, N., & Johansen-Berg, H. (2010). Distinct and overlapping functional zones in the cerebellum defined by resting state functional connectivity. *Cerebral cortex*, *20*(4), 953-965.
- Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences*, *87*(24), 9868-9872.
- Okada, K., & Hickok, G. (2006). Left posterior auditory-related cortices participate both in speech perception and speech production: Neural overlap revealed by fMRI. *Brain and language*, *98*(1), 112-117.
- Okada, K., Matchin, W., & Hickok, G. (2018). Neural evidence for predictive coding in auditory cortex during speech production. *Psychonomic bulletin & review*, *25*(1), 423-430.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(1), 97-113.
- Olson, I. R., Rao, H., Moore, K. S., Wang, J., Detre, J. A., & Aguirre, G. K. (2006). Using perfusion fMRI to measure continuous changes in neural activity with learning. *Brain and cognition*, *60*(3), 262-271.*
- Oomen, C. C., & Postma, A. (2002). Limitations in processing resources and speech monitoring. *Language and Cognitive Processes*, *17*(2), 163-184.
- Oppenheim, G. M., & Dell, G. S. (2008). Inner speech slips exhibit lexical bias, but not the phonemic similarity effect. *Cognition*, *106*(1), 528-537.
- Orban, P., Peigneux, P., Lungu, O., Albouy, G., Breton, E., Laberenne, F, ... & Doyon, J. (2010). The multifaceted nature of the relationship between performance

- and brain activity in motor sequence learning. *Neuroimage*, 49(1), 694-702.
- Orban, P., Peigneux, P., Lungu, O., Debas, K., Barakat, M., Bellec, P., ... & Doyon, J. (2011). Functional neuroanatomy associated with the expression of distinct movement kinematics in motor sequence learning. *Neuroscience*, 179, 94-103.
- Parsons, M. W., Harrington, D. L., & Rao, S. M. (2005). Distinct neural systems underlie learning visuomotor and spatial representations of motor skills. *Human brain mapping*, 24(3), 229-247.*
- Peterburs, J., & Desmond, J. E. (2016). The role of the human cerebellum in performance monitoring. *Current opinion in neurobiology*, 40, 38-44.
- Pickering, M. J., & Garrod, S. (2013). How tightly are production and comprehension interwoven?. *Frontiers in psychology*, 4, 238.
- Polskaia, N., St-Amant, G., Fraser, S., & Lajoie, Y. (2023). Neural correlates of dual-task processing following motor sequence learning: a functional near-infrared spectroscopy (fNIRS) study. *Journal of Motor Behavior*, 55(1), 92-101.
- Poon, C. S., & Merfeld, D. M. (2005). Internal models: the state of the art. *Journal of Neural Engineering*, 2(3), S147-S151.
- Postma, A. (2000). Detection of errors during speech production: A review of speech monitoring models. *Cognition*, 77(2), 97-132.
- Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *Neuroimage*, 84, 320-341.
- Purves, D., Augustine, G. J., Fitzpatrick, D., Katz, L. C., LaMantia, A. S., McNamara, J. O., & Williams, S. M. (2001). Projections from the Cerebellum. In *Neuroscience. 2nd edition*. Sinauer Associates.
- Quinn, R. (2021). *Exploring the Role of Phonological and Motor Learning in Word and Phrase Frequency Effects in Speech Production: the Effect of Word Frequency Phrase Frequency and Phonological Complexity on Speech Production in Children and Adults*. The University of Liverpool (United Kingdom).
- Ramnani, N., Toni, I., Josephs, O., Ashburner, J., & Passingham, R. E. (2000). Learning- and expectation-related changes in the human brain during motor learning. *Journal of Neurophysiology*, 84(6), 3026-3035.
- Rauch, S. L., Whalen, P. J., Savage, C. R., Curran, T., Kendrick, A., Brown, H. D., ... & Rosen, B. R. (1997). Striatal recruitment during an implicit sequence learning task as measured by functional magnetic resonance imaging. *Human*

brain mapping, 5(2), 124-132.*

- Rauschecker, A. M., Pringle, A., & Watkins, K. E. (2008). Changes in neural activity associated with learning to articulate novel auditory pseudowords by covert repetition. *Human brain mapping*, 29(11), 1231-1242.
- Reason, J. T. (1978). Motion sickness adaptation: a neural mismatch model. *Journal of the Royal Society of Medicine*, 71(11), 819-829.
- Reber, P. J. (2008). Cognitive neuroscience of declarative and nondeclarative memory. *Advances in psychology*, 139, 113-123.
- Reichmuth, C., & Casey, C. (2014). Vocal learning in seals, sea lions, and walruses. *Current opinion in neurobiology*, 28, 66-71.
- Rieckmann, A., Fischer, H., & Bäckman, L. (2010). Activation in striatum and medial temporal lobe during sequence learning in younger and older adults: relations to performance. *Neuroimage*, 50(3), 1303-1312.*
- Riès, S., Janssen, N., Dufau, S., Alario, F. X., & Burle, B. (2011). General-purpose monitoring during speech production. *Journal of cognitive neuroscience*, 23(6), 1419-1436.
- Roche, R. A., Commins, S., Agnew, F., Cassidy, S., Corapi, K., Leibbrand, S., ... & O'Mara, S. M. (2007). Concurrent task performance enhances low-level visuomotor learning. *Perception & psychophysics*, 69, 513-522.
- Roelofs, A. (2020). On (correctly representing) comprehension-based monitoring in speaking: Rejoinder to Nozari (2020). *Journal of Cognition*, 3(1).
- Roelofs, A. (2020). Self-monitoring in speaking: In defense of a comprehension-based account. *Journal of Cognition*, 3(1).
- Rolls, E. T., Huang, C. C., Lin, C. P., Feng, J., & Joliot, M. (2020). Automated anatomical labeling atlas 3. *Neuroimage*, 206, 116189.
- Rose, M., Haider, H., Salari, N., & Büchel, C. (2011). Functional dissociation of hippocampal mechanism during implicit learning based on the domain of associations. *Journal of Neuroscience*, 31(39), 13739-13745.
- Runnqvist, E., Bonnard, M., Gauvin, H. S., Attarian, S., Trébuchon, A., Hartsuiker, R. J., & Alario, F. X. (2016). Internal modeling of upcoming speech: A causal role of the right posterior cerebellum in non-motor aspects of language production. *Cortex*, 81, 203-214.
- Runnqvist, E., Chanoine, V., Strijkers, K., Pattamadilok, C., Bonnard, M., Nazarian, B., ... & Alario, F. X. (2021). Cerebellar and cortical correlates of internal and external speech error monitoring. *Cerebral Cortex Communications*, 2(2),

tgab038.

- Sakai, K., Ramnani, N., & Passingham, R. E. (2002). Learning of sequences of finger movements and timing: frontal lobe and action-oriented representation. *Journal of neurophysiology*, *88*(4), 2035-2046.*
- Sakreida, K., Higuchi, S., Di Dio, C., Ziessler, M., Turgeon, M., Roberts, N., & Vogt, S. (2018). Cognitive control structures in the imitation learning of spatial sequences and rhythms—An fMRI study. *Cerebral Cortex*, *28*(3), 907-923.*
- Sanchez, D. J., & Reber, P. J. (2013). Explicit pre-training instruction does not improve implicit perceptual-motor sequence learning. *Cognition*, *126*(3), 341-351.
- Satterthwaite, T. D., Elliott, M. A., Gerraty, R. T., Ruparel, K., Loughead, J., Calkins, M. E., ... & Wolf, D. H. (2013). An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *Neuroimage*, *64*, 240-256.
- Schendan, H. E., Searl, M. M., Melrose, R. J., & Stern, C. E. (2003). An FMRI study of the role of the medial temporal lobe in implicit and explicit sequence learning. *Neuron*, *37*(6), 1013-1025.
- Schiller, N. O., Meyer, A. S., Baayen, R. H., & Levelt, W. J. (1996). A comparison of lexeme and speech syllables in Dutch. *Journal of Quantitative Linguistics*, *3*(1), 8-28.
- Schimke, E. A., Angwin, A. J., Cheng, B. B., & Copland, D. A. (2021). The effect of sleep on novel word learning in healthy adults: A systematic review and meta-analysis. *Psychonomic Bulletin & Review*, *28*(6), 1811-1838.
- Schubotz, R. I., & von Cramon, D. Y. (2001). Interval and ordinal properties of sequences are associated with distinct premotor areas. *Cerebral Cortex*, *11*(3), 210-222.*
- Schumacher, E. H., & Schwarb, H. (2009). Parallel response selection disrupts sequence learning under dual-task conditions. *Journal of Experimental Psychology: General*, *138*(2), 270.
- Segawa, J. A., Tourville, J. A., Beal, D. S., & Guenther, F. H. (2015). The neural correlates of speech motor sequence learning. *Journal of cognitive neuroscience*, *27*(4), 819-831.
- Seidler, R. D., Purushotham, A., Kim, S. G., Ugurbil, K., Willingham, D., & Ashe, J. (2002). Cerebellum activation associated with performance change but not motor learning. *Science*, *296*(5575), 2043-2046.*

- Shadmehr, R., Mussa-Ivaldi, F. A., & Bizzi, E. (1993). Postural force fields of the human arm and their role in generating multijoint movements. *Journal of Neuroscience*, *13*(1), 45-62.
- Shergill, S. S., Brammer, M. J., Williams, S. C., Murray, R. M., & McGuire, P. K. (2000). Mapping auditory hallucinations in schizophrenia using functional magnetic resonance imaging. *Archives of general psychiatry*, *57*(11), 1033-1038.
- Shima, K., Mushiake, H., Saito, N., & Tanji, J. (1996). Role for cells in the presupplementary motor area in updating motor plans. *Proceedings of the National Academy of Sciences*, *93*(16), 8694-8698.
- Siegert, R. J., Taylor, K. D., Weatherall, M., & Abernethy, D. A. (2006). Is implicit sequence learning impaired in Parkinson's disease? A meta-analysis. *Neuropsychology*, *20*(4), 490.**
- Siegert, R. J., Weatherall, M., & Bell, E. M. (2008). Is implicit sequence learning impaired in schizophrenia? A meta-analysis. *Brain and Cognition*, *67*(3), 351-359.**
- Singer, R. N. (1965). Massed and distributed practice effects on the acquisition and retention of a novel basketball skill. *Research Quarterly. American Association for Health, Physical Education and Recreation*, *36*(1), 68-77.
- Singer, R. N. (1980). *Motor learning and human performance: An application to motor skills and movement behaviors*. New York: Macmillan.
- Soares, J. M., Magalhães, R., Moreira, P. S., Sousa, A., Ganz, E., Sampaio, A., ... & Sousa, N. (2016). A Hitchhiker's guide to functional magnetic resonance imaging. *Frontiers in neuroscience*, *10*, 515.
- Sokolov, A. A., Miall, R. C., & Ivry, R. B. (2017). The cerebellum: adaptive prediction for movement and cognition. *Trends in cognitive sciences*, *21*(5), 313-332.
- Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of comparative and physiological psychology*, *43*(6), 482.
- Stansbury, A. L., & Janik, V. M. (2021). The role of vocal learning in call acquisition of wild grey seal pups. *Philosophical Transactions of the Royal Society B*, *376*(1836), 20200251.
- Steele, C. J. (2008). *Tides of change: Identifying the neural correlates of motor sequence learning* (Doctoral dissertation, Concordia University).*
- Steele, C. J., & Penhune, V. B. (2010). Specific increases within global decreases: a functional magnetic resonance imaging investigation of five days of motor sequence learning. *Journal of Neuroscience*, *30*(24), 8332-8341.*

- Stelmach, G. E. (1969). Efficiency of motor learning as a function of intertrial rest. *Research Quarterly. American Association for Health, Physical Education and Recreation*, 40(1), 198-202.
- Stockert, A., Schwartz, M., Poeppel, D., Anwander, A., & Kotz, S. A. (2021). Temporocerebellar connectivity underlies timing constraints in audition. *Elife*, 10, e67303.
- Stoodley, C. J., & Schmahmann, J. D. (2009). Functional topography in the human cerebellum: a meta-analysis of neuroimaging studies. *Neuroimage*, 44(2), 489-501.
- Stoodley, C. J., & Schmahmann, J. D. (2010). Evidence for topographic organization in the cerebellum of motor control versus cognitive and affective processing. *Cortex*, 46(7), 831-844.
- Strangman, G., Heindel, W. C., Anderson, J. A., & Sutton, J. P. (2005). Learning motor sequences with and without knowledge of governing rules. *Neurorehabilitation and neural repair*, 19(2), 93-114.*
- Strick, P. L., Dum, R. P., & Fiez, J. A. (2009). Cerebellum and nonmotor function. *Annual review of neuroscience*, 32, 413-434.
- Strijkers, K. (2016). Can hierarchical models display parallel cortical dynamics? A non-hierarchical alternative of brain language theory. *Language, Cognition and Neuroscience*, 31(4), 465-469.
- Tanaka, H., Ishikawa, T., Lee, J., & Kakei, S. (2020). The cerebro-cerebellum as a locus of forward model: a review. *Frontiers in systems neuroscience*, 14, 19.
- Terband, H., Maassen, B., Guenther, F. H., & Brumberg, J. (2014). Auditory-motor interactions in pediatric motor speech disorders: Neurocomputational modeling of disordered development. *Journal of Communication Disorders*, 47, 17-33.
- Terband, H., Rodd, J., & Maas, E. (2020). Testing hypotheses about the underlying deficit of apraxia of speech through computational neural modelling with the DIVA model. *International Journal of Speech-Language Pathology*, 22(4), 475-486.
- Terzic, L., Voegtle, A., Farahat, A., Hartong, N., Galazky, I., Nasuto, S. J., ... & Sweeney-Reed, C. M. (2022). Deep brain stimulation of the ventrointermediate nucleus of the thalamus to treat essential tremor improves motor sequence learning. *Human Brain Mapping*, 43(15), 4791-4799.

- Thompson, K., Gibbings, A., Shaw, J., Ray, L., Hébert, G., De Koninck, J., & Fogel, S. (2021). Sleep and Second-Language Acquisition Revisited: The Role of Sleep Spindles and Rapid Eye Movements. *Nature and Science of Sleep*, 1887-1902.
- Toni, I., Krams, M., Turner, R., & Passingham, R. E. (1998). The time course of changes during motor sequence learning: a whole-brain fMRI study. *Neuroimage*, 8(1), 50-61.*
- Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and cognitive processes*, 26(7), 952-981.
- Travers, B. G., Kana, R. K., Klinger, L. G., Klein, C. L., & Klinger, M. R. (2015). Motor learning in individuals with autism spectrum disorder: activation in superior parietal lobule related to learning and repetitive behaviors. *Autism Research*, 8(1), 38-51.
- Turkeltaub, P. E., Eden, G. F., Jones, K. M., & Zeffiro, T. A. (2002). Meta-analysis of the functional neuroanatomy of single-word reading: method and validation. *Neuroimage*, 16(3), 765-780.
- Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: improved N3 bias correction. *IEEE transactions on medical imaging*, 29(6), 1310-1320.
- Tzvi, E., Stoldt, A., Witt, K., & Krämer, U. M. (2015). Striatal-cerebellar networks mediate consolidation in a motor sequence learning task: An fMRI study using dynamic causal modelling. *Neuroimage*, 122, 52-64.
- van den Wildenberg, W. P., van Boxtel, G. J., van der Molen, M. W., Bosch, D. A., Speelman, J. D., & Brunia, C. H. (2006). Stimulation of the subthalamic region facilitates the selection and inhibition of motor responses in Parkinson's disease. *Journal of cognitive neuroscience*, 18(4), 626-636.
- van der Graaf, F. H. C. E., De Jong, B. M., Maguire, R. P., Meiners, L. C., & Leenders, K. L. (2004). Cerebral activation related to skills practice in a double serial reaction time task: striatal involvement in random-order sequence learning. *Cognitive Brain Research*, 20(2), 120-131.*
- van der Graaf, F. H. C. E., Maguire, R. P., Leenders, K. L., & de Jong, B. M. (2006). Cerebral activation related to implicit sequence learning in a Double Serial Reaction Time task. *Brain research*, 1081(1), 179-190.*

- Vanderwal, T., Kelly, C., Eilbott, J., Mayes, L. C., & Castellanos, F. X. (2015). Inscapes: A movie paradigm to improve compliance in functional magnetic resonance imaging. *Neuroimage*, *122*, 222-232.
- Verwey, W. B., Shea, C. H., & Wright, D. L. (2015). A cognitive framework for explaining serial processing and sequence execution strategies. *Psychonomic bulletin & review*, *22*, 54-77.
- Verwey, W. B., & Wright, D. L. (2004). Effector-independent and effector-dependent learning in the discrete sequence production task. *Psychological research*, *68*(1), 64-70.
- Wadden, K., Brown, K., Maletsky, R., & Boyd, L. A. (2013). Correlations between brain activity and components of motor learning in middle-aged adults: an fMRI study. *Frontiers in human neuroscience*, *7*, 169.
- Watkins, K. E., Smith, S. M., Davis, S., & Howell, P. (2008). Structural and functional abnormalities of the motor system in developmental stuttering. *Brain*, *131*(1), 50-59.
- Weber, R., Mangus, J. M., & Huskey, R. (2015). Brain imaging in communication research: A practical guide to understanding and evaluating fMRI studies. *Communication Methods and Measures*, *9*(1-2), 5-29.
- Weerathunge, H. R., Alzamendi, G. A., Cler, G. J., Guenther, F. H., Stepp, C. E., & Zañartu, M. (2022). LaDIVA: A neurocomputational model providing laryngeal motor control for speech acquisition and production. *PLOS Computational Biology*, *18*(6), e1010159.
- Whitfield-Gabrieli, S., & Nieto-Castañón, A. (2012). Conn: A functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain connectivity*, *2*(3), 125-141.
- Whitfield, J. A., & Goberman, A. M. (2017). Speech motor sequence learning: Acquisition and retention in Parkinson disease and normal aging. *Journal of Speech, Language, and Hearing Research*, *60*(6), 1477-1492.
- Whitfield, J. A., & Goberman, A. M. (2017). Speech motor sequence learning: Effect of Parkinson disease and normal aging on dual-task performance. *Journal of Speech, Language, and Hearing Research*, *60*(6S), 1752-1765.
- Willingham, D. B. (1998). A neuropsychological theory of motor skill learning. *Psychological review*, *105*(3), 558.
- Witt, S. T., Laird, A. R., & Meyerand, M. E. (2008). Functional neuroimaging correlates of finger-tapping task variations: an ALE meta-analysis. *Neuroimage*, *42*(1),

343-356.**

- Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science*, 269(5232), 1880-1882.
- Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in cognitive sciences*, 2(9), 338-347.
- Wu, T., Chan, P., & Hallett, M. (2010). Effective connectivity of neural networks in automatic movements in Parkinson's disease. *Neuroimage*, 49(3), 2581-2587.
- Zhang, Y., Brady, M., & Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE transactions on medical imaging*, 20(1), 45-57.
- Zuk, J., Loui, P., & Guenther, F. (2022). Neural Control of Speaking and Singing: The DIVA Model for Singing.
- Zwicker, J. G., Missiuna, C., Harris, S. R., & Boyd, L. A. (2011). Brain activation associated with motor skill practice in children with developmental coordination disorder: an fMRI study. *International Journal of Developmental Neuroscience*, 29(2), 145-152.

*Study included in the meta-analyses

**A meta-analysis

Notes

1. An analogy for external monitoring would be seeing a pedestrian in the rear-view mirror reappear after passing them. This provides us with external feedback, informing us that we safely passed the pedestrian. On the other hand, internal monitoring allow one to anticipate a way of moving the car that will be safe for the pedestrian instead of simply waiting to see the pedestrian reappear in the rear-view mirror.
2. Twenty-five subject groups accounted for the motor execution, 12 through a random sequence, and 11 through a sequence that was previously known or already practiced, leaving two subject groups that used different motor-accounting control conditions. In Konovalov et al., 2018, activations were compared with those during the previous trial, i.e. the sequence has been practiced less in the baseline than in the learning condition. In Heun et al., 2004, the participants were asked to tap their fingers in the control condition as they please. These two studies were included in the meta-analysis focusing on the studies that accounted for motor execution, and excluded for the subsequent meta-analyses focusing on the studies that used random or previously known or practiced sequences.
3. Including several contrasts from the same subject group as a single group is only relevant for the meta-analyses including all the studies, the studies that used a motor baseline, regardless of the task, or the SRTT regardless of the baseline.

APPENDICES

A. Motor sequence learning and motor adaptation in table-top games

Some games have rules that may be easy to explain, but require sequences that are complex and have to be learned. For instance, in Caroline, a table-top game played at medieval festivities in the city of Carcassonne in France (Figure 1, left), you manipulate a ball with sticks that are attached on the far side of the table, while you hold the other end. Manipulating the distance, the angle and the elevation of the sticks, you should drop the ball in one of the holes on the table, with the goal of the ball landing into the farthest possible hole, which scores the most points. In the beginning, this is an impossible task, but with the time, you learn the right sequence of movements and improve your performance. Other games are easier, as they have simple rules and require sequence movements that are easily learned. This is the case with air-hockey (Figure 1, right). What is more difficult in this game is the motor adaptation you need if you want to play it with a three year old, giving them impression it is them playing. In order for this to work, you need to adapt the force you use, and you would likely have to lean to get to their height, partially changing your reference frame and the established internal model.



Figure 1. – *Motor sequence learning and motor adaptation in table-top games.* Left: Playing Caroline requires acquisition of a novel sequential movement that includes variations in the distance, angle and elevation of the sticks to drop the ball in the right place. Right: Motor adaptation. Playing air-hockey with a three year old requires changes in the force, aiming and reactivity of your movements.

B. Foci and converging clusters for individual meta-analyses

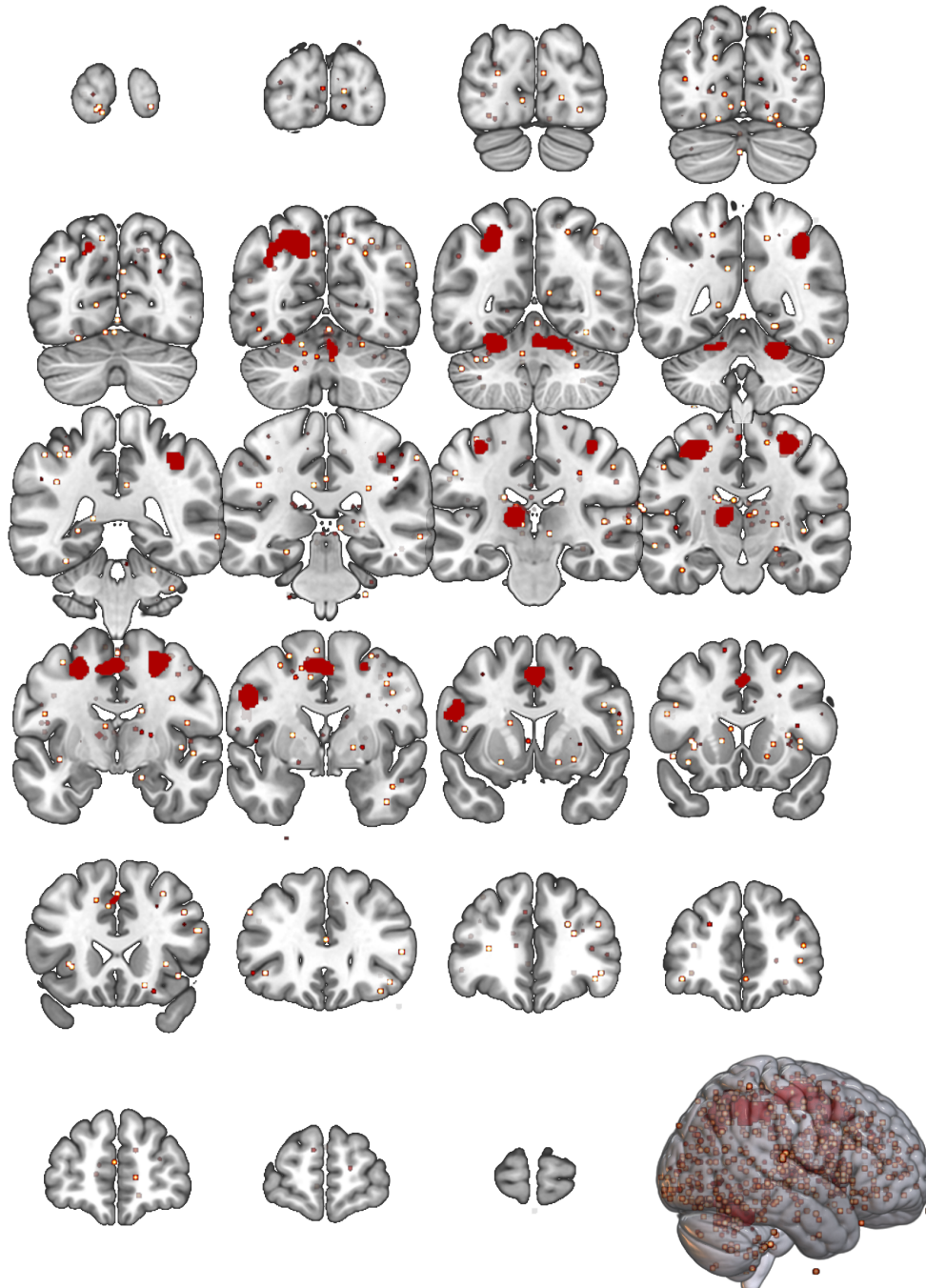


Figure 2. – *Foci and converging clusters for the meta-analysis of all studies regardless of the task and the baseline. N of experiments = 48, N of participants = 640, N of foci = 1074.*

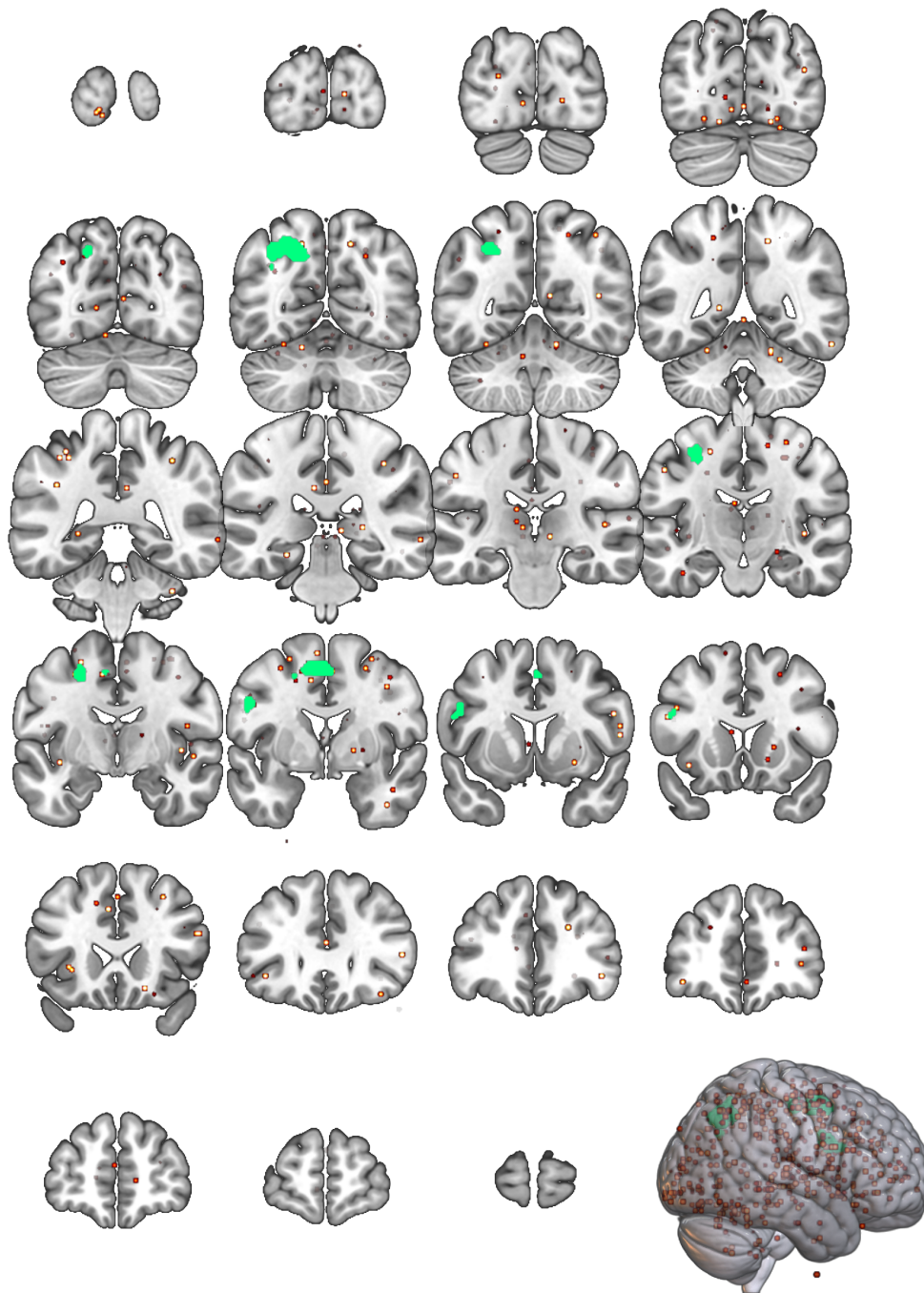


Figure 3. – *Foci and converging clusters for the meta-analysis focusing on the serial reaction time task. N of experiments = 29, N of participants = 382, N of foci = 600.*

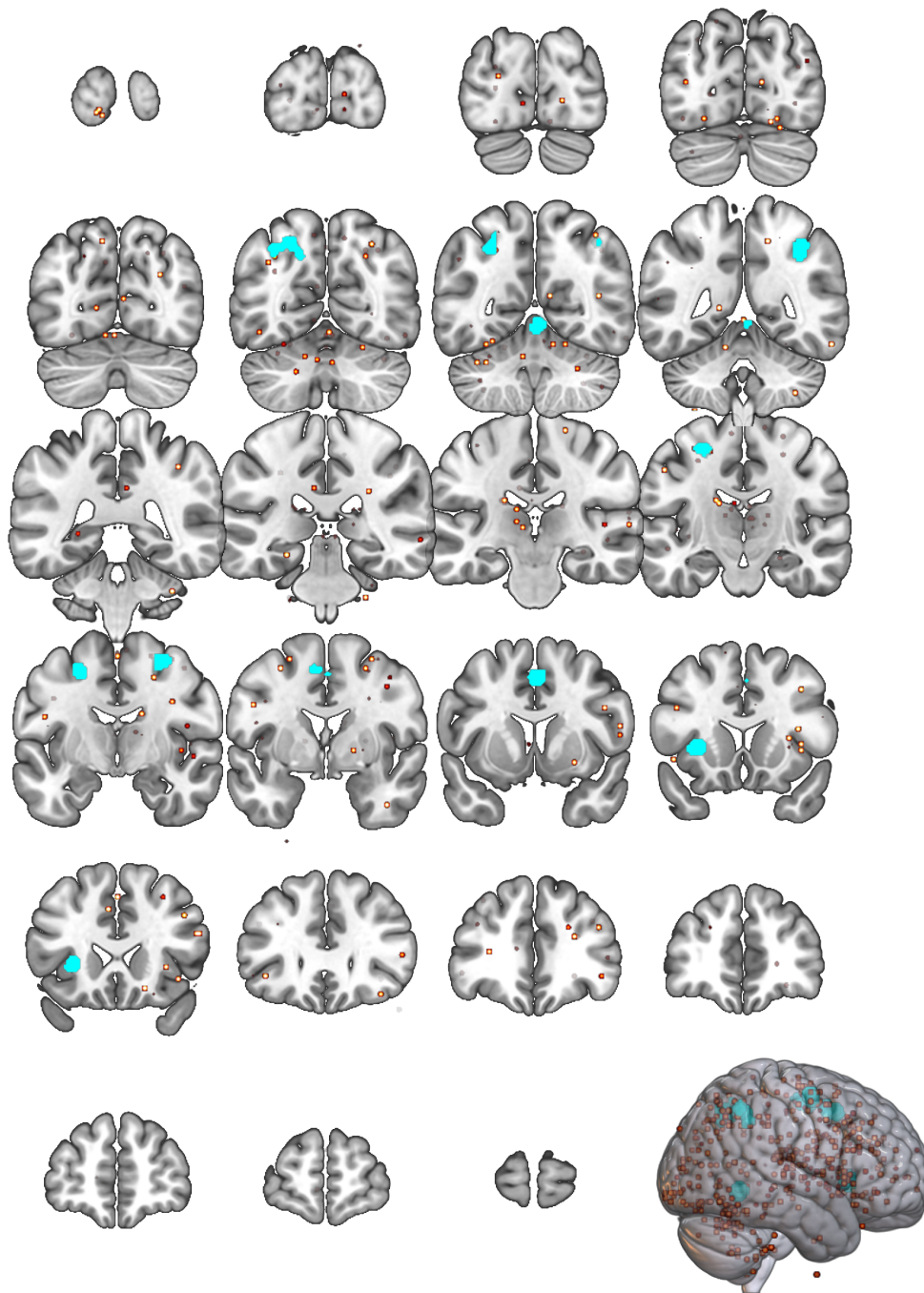


Figure 4. – *Foci and converging clusters for the meta-analysis focusing on studies that accounted for motor execution in the baseline. N of experiments = 25, N of participants = 372, N of foci = 541.*

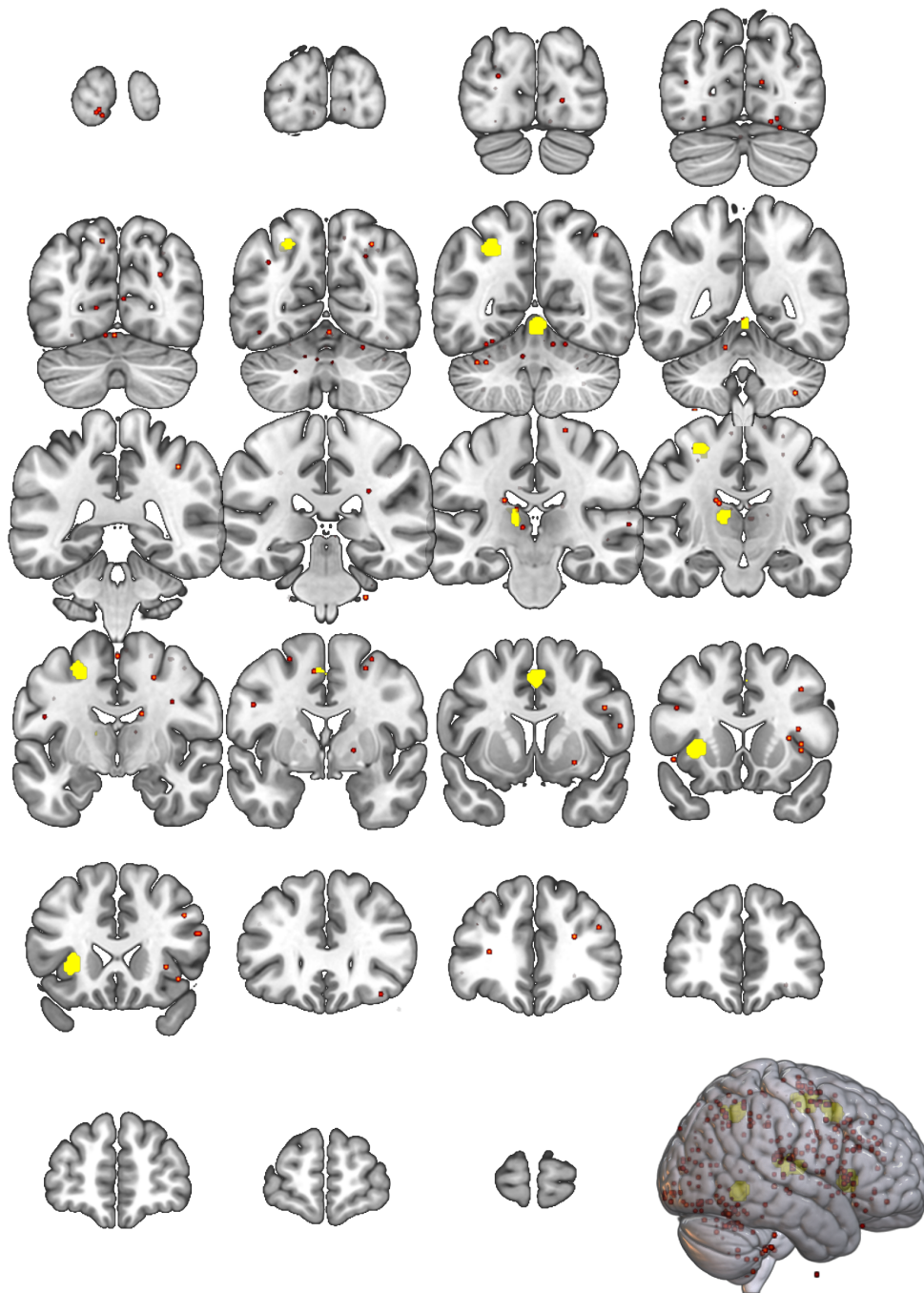


Figure 5. – *Foci and converging clusters for the meta-analysis focusing on studies that used known or previously practiced sequences as the baseline. N of experiments = 11, N of participants = 172, N of foci = 359.*

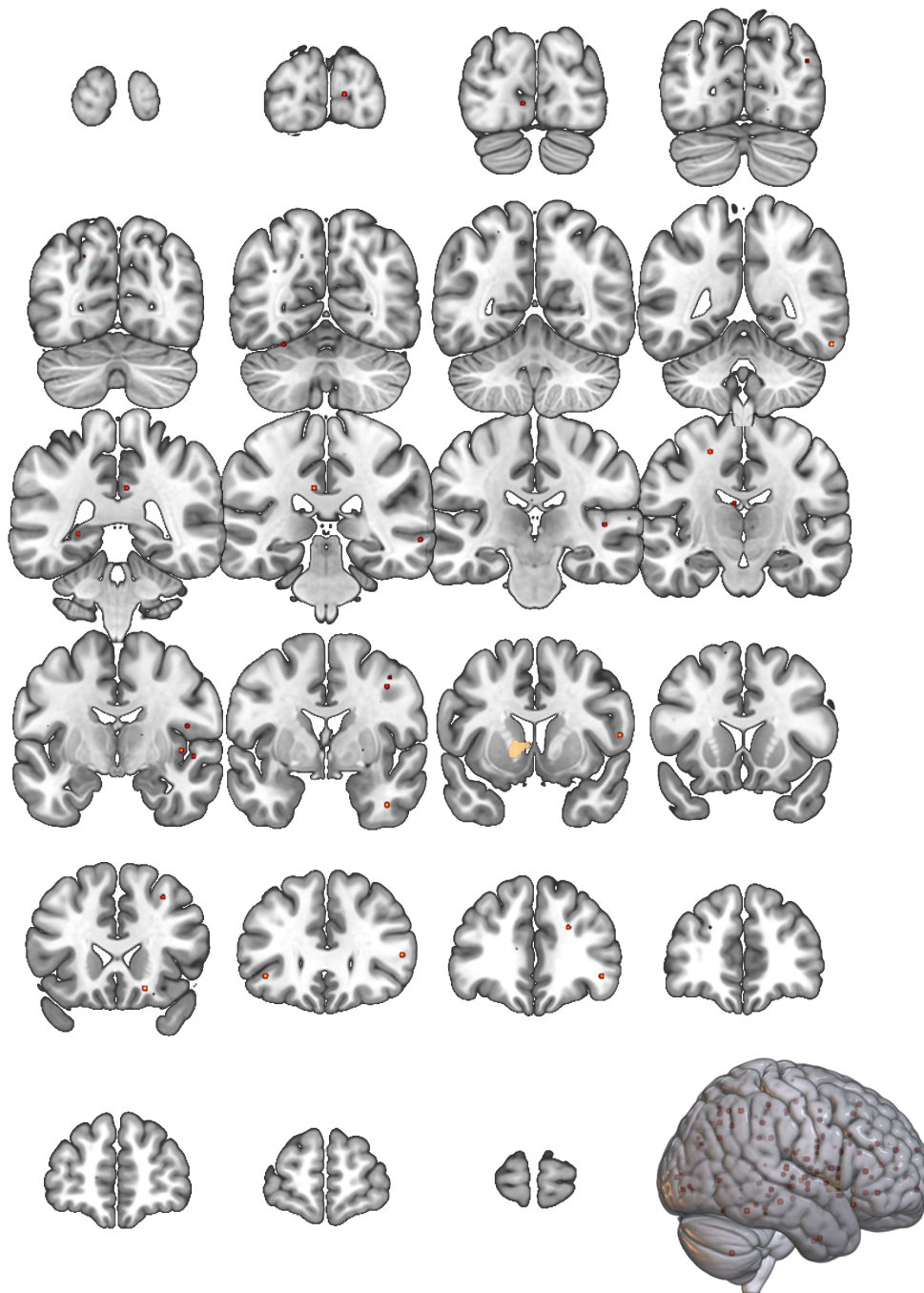


Figure 6. – *Foci and converging clusters for the meta-analysis focusing on studies that used random sequences as the baseline. N of experiments = 12, N of participants = 166, N of foci = 137.*

C. Learning-related functional connectivity changes between the right crus I and the left precuneus

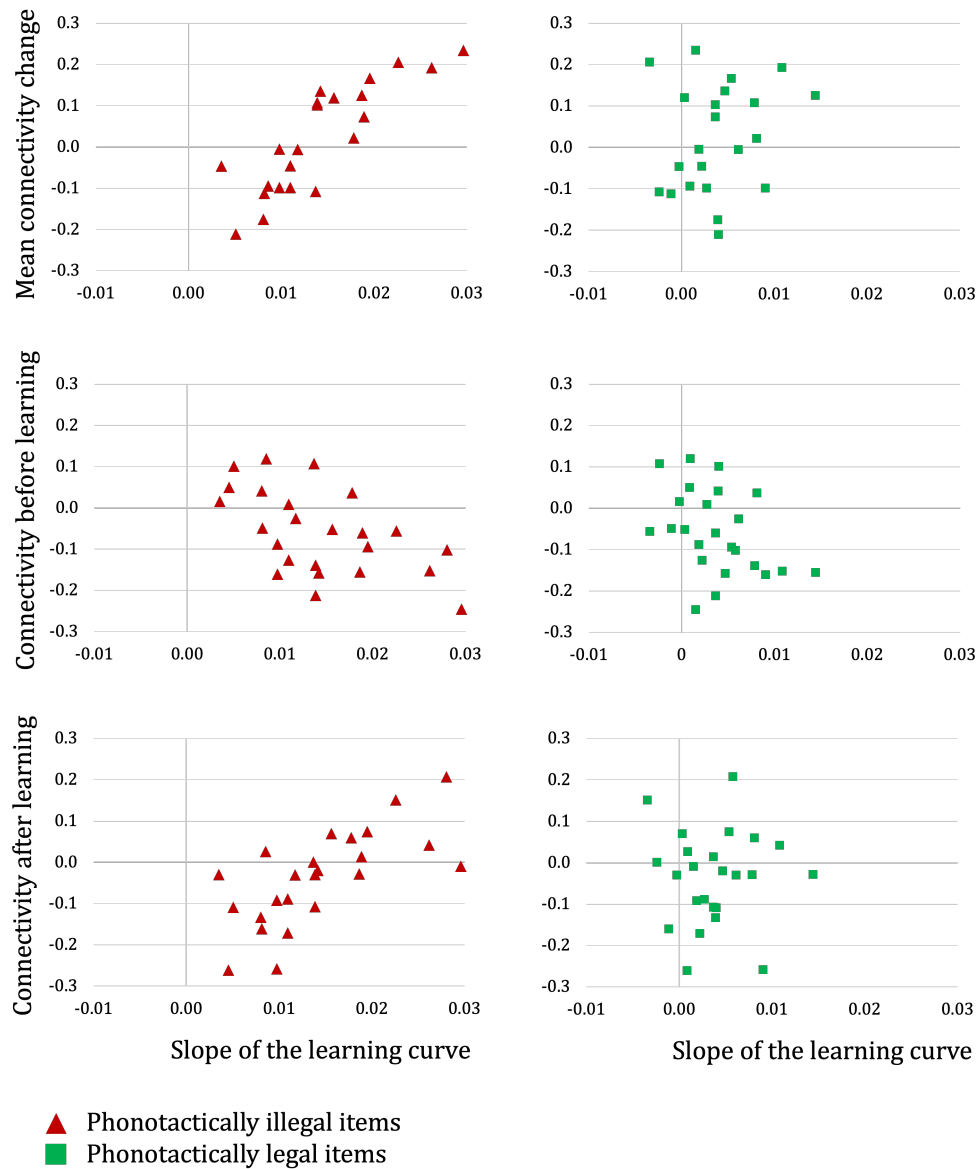


Figure 7. – Learning-related functional connectivity changes between the seed in the right cerebellar crus I (38, -64, -30) and the target in the left precuneus (-16, -70, 36).

Groups



Ignouchk