Université Sorbonne Nouvelle - CNRS
École Doctorale 622 : Sciences du Langage
UMR7018 / Laboratoire de Phonétique et Phonologie

Thèse

en vue de l'obtention du grade de docteur en phonétique

Presentée par

Daria D'Alessandro

Interspeaker variations in V-to-V coarticulation: effects of Motor Speech Disorders, age, speech tempo changes, and boundary type.

Sous la direction de

Mme Cécile Fougeron

Soutenue à Paris le 15 décembre 2022 devant le jury composé de :

Pascal Perrier (Rapporteur) Professeur, Grenoble-INP, GIPSA-lab

Douglas H. Whalen (Rapporteur) Professeur, CUNY, Haskins Laboratories

Véronique Delvaux (Examinatrice) Chercheur qualifié FNRS, UMONS

Anne Hermes (Examinatrice) Chargée de recherche CNRS, LPP

Cécile Fougeron (Directrice de thèse) Directeur de recherche CNRS, LPP

Variations individuelles dans la coarticulation anticipatoire V-à-V : effets des Troubles Moteurs de la Parole, de l'age, de changements de tempo et du type de frontière.

La coarticulation anticipatoire se réfère à l'anticipation des mouvements articulatoires pour la réalisation de cibles de parole à venir et peut être considérée comme un indice de planification. Dans quatre études, la coarticulation anticipatoire V-à-V est étudiée dans différents Troubles Moteurs de la Parole, i.e. Apraxie de la Parole et Dysarthrie associée à la SLA, la maladie de Wilson, et la maladie de Parkinson (et comparée à la coarticulation C-à-V), chez des adultes âgés de 20 à 93 ans, et dans un groupe restreint de locutrices dans différentes conditions de parole : tempo lent, rapide et normal, dans un mot, à travers une frontière de mot et de proposition relative.

Les résultats montrent une réduction de la coarticulation V-à-V dans l'Apraxie de la Parole et la Dysarthrie, qui pourrait être expliquée par des déficits spécifiques à ces pathologies. Une réduction non-linéaire de la coarticulation avec l'âge semble liée à un ralentissement du débit jusqu'à 70 ans, alors qu'une relation directe n'est pas trouvée pour les locuteurs plus âgés. Les différences inter-individuelles de coarticulation en réponse aux changements de tempo suggèrent que la relation entre la coarticulation et le débit articulatoire est spécifique au locuteur. Des variations inter-individuelles de coarticulation sont trouvées aussi en fonction du type de frontière et ne sont que partiellement expliquées par le phrasé prosodique.

Ces résultats sont discutés selon deux axes, l'un traitant de la taille des unités de planification motrice dans la parole, et l'autre discutant comment peuvent êtres modélisées les variations de coarticulation en fonction du locuteur et de la population.

Mot clés : coarticulation, AoS, Dysarthrie, âge, planification motrice, débit

Interspeaker variations in V-to-V coarticulation: effects of Motor Speech Disorders, age, speech tempo changes, and boundary type.

Anticipatory coarticulation refers to the anticipation of articulatory movements for the achievement of forthcoming speech goals and can be considered an index of planning in speech. In four investigations, anticipatory V-to-V coarticulation is investigated in different Motor Speech Disorders, i.e. AoS and Dysarthria associated with ALS, Wilson Disease, and Parkinson's Disease (and compared to C-to-V coarticulation), in adults spanning 20 to 93 years old, and in a small set of speakers in different conditions: at a slow, fast, and habitual speech tempo, and within a word, across word boundary, and across clause boundary.

The results show a reduction of V-to-V coarticulation in AoS and Dysarthria, which could be accounted for by disorder-specific impairment. A non-linear reduction of coarticulation is found with age. If this reduction can be seen in relation to a slowing of speech in speakers aged 20 to 70, with age-specific patterns of covariation, a relationship between these two factors is not found after 70 y.o.a. Individual responses to changes in speech tempo suggest that the relationship between coarticulation and articulation rate is speaker-specific. Interspeaker variations in V-to-V coarticulation are found also depending on boundary type and are only partially explained by specificities in the prosodic phrasing.

These results are discussed along two axes, one discussing the size of the planning units in speech, i.e. units over which the movements for the achievement of a string of speech targets are planned, and the other discussing how speaker-specific and population-specific variations in coarticulation can be modeled in a coproduction account of coarticulation.

Keywords: coarticulation, AoS, Dysarthria, age, motor planning, rate

A mia nonna Jolanda,

il cui ricordo mi ha accompagnata fin qui

Acknowledgments

I will start these acknowledgments by thanking my supervisor, Cécile Fougeron: thank you for including me in this project, for your guidance, your always useful, sharp, and smart remarks, the discussions, the critique, and the encouragement, until the very last days of writing.

I would like to express my gratitude to Pascal Perrier and Douglas Whalen for having accepted to review my work, and to Véronique Delvaux and Anne Hermes for having accepted to be part of my jury. I am truly honored to have you participating in my defense.

This work would have not been possible without the MoSpeeDi project (FNS CRSII5_173711), the related projects, MonPaGe, Speech'N'Co, ChaSpeePro, and the people formerly or currently on them: Marina Laganaro, whose feedback is always precious; Leonardo Lancia and Ivana Didirkova, mentors and friends, who designed and recorded the speech tempo changes; the people who worked on the data collection, starting from Michaela Pernon and Angélina Bourbon, with whom I have had the chance to work in several occasions; Fanny Guitard-Ivent, who helped me with the MARS modeling and answered to all my statistics questions; Alice Yildiz, who helped with the perception experiment and the prosodic annotation, and all the PhD students and postdocs who collaborated on the MoSpeeDi project in particular.

I would have never thought of embarking on this journey without Caterina Petrone, without whom I would not even know what experimental research is, Jelena Krivokapić, and Simone Falk.

I would like to thank some important students and former students of the LPP: Amelia Pettirossi, who has been incredibly supportive all these years, and without whom I would have most probably never finished (and would have a much more boring life); Angélina Bourbon, for the most empathetic support, the conversations on speech planning and so much more; Amélie: for her friendship inside and outside of the lab. Many thanks also to the other students and permanents of the LPP who gave me the chance to work in an amazing environment, and who (especially in the last month of writing) have encouraged me, talked to me when I was stressed, helped me in my laptop misfortunes, given me advice or just stopped by to ask me how I was doing: Bowei, Clémence, Nicolas, Naomi, Anne, Fanny, Jinyu, the LPP support team Roland, Angélique, Gwen, who saved me from bureaucracy a million times, Jacqueline Vaissière, Annie Rialland, and all the other colleagues and former and current PhD students with whom I had conferences, lunches, conversations, and parties. I would like to include in this group also Francesca Carbone, who shared from afar a similar journey, that we started together.

Marta and Maria: you are my oldest friends, you are family to me and although we are spread across Europe, I always felt your support and friendship, and not least the material support in this submission.

Writing these acknowledgments, I realized that, near or far, I am surrounded by a lot of people, and I consider myself really lucky. I will likely thank them in person: Jay, my friends in Naples, Sergio, Mirb, Marika, and the others; Anna, Ale, Giovy and Simo; the friends of the MDI, with whom I shared two years in Paris; my roommate Theresa, who encouraged me in the last months and helped with the references. Bianca, for having found us a home, the endless deep conversations, and the colors. Juliana, for the support, the talks, and the corrections until the end. My brother Ferdinando for the continuous music exchanges who have enriched my work playlists.

I thank my parents for having given me the chance to pursue higher education by financially supporting my university studies.

Finally, my city, for providing the colors for the graphs: your sky, buildings, stores.

Daria

[ɛl fɛ dezɔʁmɛ paʁti də la sɛkt fonetik dit mɛʁsi puʁ sə ʁesi də sɔ̃paʁkuʁ inisjatik]

Table of contents

List of Figures

List of tables

1 Introduction: what and why

"Speaking is coarticulating gestures"

Farnetani and Recasens, 2010

The term coarticulation refers to a process underlying the articulation of speech sounds and at the same time to the output of this process. In speech, the movements realized by the articulators to produce segments overlap and interact with each other. As a result, the acoustic and articulatory characteristics of a segment are influenced by the characteristics of adjacent and non-adjacent ones (Farnetani & Recasens, 2010; Volenec 2015; Recasens, 2018).

Coarticulation, and especially anticipatory coarticulation, i.e. the influence of a segment on a preceding one, has long served to investigate the processes of planning the phonetic code and its execution. Indeed, if a speech goal is anticipated and visible on a preceding one, the movements for producing it have begun at the same time as the preceding one or during its production. The question that has been long addressed is whether coarticulation reflects the planning of the movements necessary to achieve the subsequent speech goals within a speech plan, or it is just the natural consequence of the execution of successive speech plans in a rapid and fluent way, leading to some unplanned overlap. In other words, whether the overlap between segments is controlled, or uncontrolled. In this respect, here the notion of planning of studies on motor control is adopted. Movement planning entails all the processes of preparation of movements that occur before the initiation of the movement (Grimme, Fuchs, Perrier, & Schöner, 2011). This notion has been borrowed for speech in models of speech production (e.g. DIVA: Guenther, 1994; Guenther, Hampson, & Johnson, 1998; and GEPPETO: Perrier, 2014) and in previous studies on anticipatory coarticulation (e.g. Ma, Perrier, & Dang, 2015; Noiray, Wieling, Abakarova, Rubertus, & Tiede, 2019, which will be cited again in the next sections).

Arguments in favor of a planned anticipatory coarticulation have been provided by a series of experimental studies, which often investigated the anticipation of a vowel in a preceding one across one or more consonants, a phenomenon commonly known as vowel-to-vowel (from here on V-to-V) coarticulation. Since the seminal paper of Öhman (1966), who observed that, in V_1CV_2 sequences, the formant transition toward V_2 began during the production of V_1, V-to-V coarticulation has been one of the privileged means to investigate the scope of anticipatory coarticulation in acoustics. Early arguments showing that anticipatory coarticulation is "largely planned" have been provided by Whalen in his well-known 1990 paper (Whalen, 1990). In his experimental paradigm, speakers had to utter disyllabic $V_1C_2V_2$ or trisyllabic $əC_1V_1C_2V_2$ sequences where, alternately, C_2 or V_2 are either displayed from the beginning (known condition) or displayed right after the speakers started to utter the word (unknown condition). In the unknown condition, even though speakers had virtually the time to anticipate the following consonant or vowel during the execution of V_1, coarticulatory effects appeared later and were way smaller than the ones found in the known condition. Although there was an effect of the unknown condition on both the anticipation of the vowel and the consonant, the anticipation of V_2 was particularly affected, with a small amount of coarticulation appearing at the very end of V_1. This result suggests that the utterance needs to be planned as a whole for extensive anticipatory coarticulation to appear and thus that anticipatory coarticulation is largely controlled for. The planned nature of V-to-V anticipatory coarticulation is also supported by studies that show that coarticulatory patterns are language-specific. Cross-linguistic differences in coarticulation have been shown, *inter alia,* by Beddor, Harnsberger, and Lindemann (2002) for English and Shona, by Mok (2010) for Thai and English, and by Ma, Perrier, & Dang (2015) for French and Mandarin Chinese (for a review of earlier studies see Manuel, 1999). If coarticulation is language-specific, it means that during acquisition speakers learn also the amount of overlap that is allowed in the language for that specific context, discouraging a view of coarticulation as mechanical.

Another question that has been raised concerning the nature of coarticulation is whether carryover coarticulation, i.e. the effect of a segment on a following one, has to be accounted for by the same mechanism as anticipatory coarticulation. Indeed, some accounts of coarticulation argue that, if anticipatory coarticulation is an index of planning, carryover coarticulation is not planned, but depends more on biomechanical constraints, such as the inertia of the organs of the speech apparatus (e.g. Recasens, 1984). Some studies have highlighted differences between anticipatory and carryover coarticulatory patterns. The hypothesis that carryover coarticulation is not controlled is supported by studies that show a more variable extent and a greater sensitivity to articulatory constraints for carryover than for anticipatory coarticulation (Recasens 1984, 2002, 2015). However, some arguments have been adduced to support the hypothesis that carryover coarticulation could also be controlled. For example, V-to-V carryover coarticulation has been reported to present language-specific patterns (Beddor *et al.*, 2002; Mok, 2010); and second language learners have been shown to adjust their degree of C-to-V coarticulation to approach that of the target language (Oh, 2008). Therefore, it is still debated if the process of coarticulation has to be distinguished into anticipatory or carryover forms.

In this thesis, we will primarily rely on the investigation of anticipatory extrasyllabic V-to-V coarticulation, which allows addressing the subject of speech planning on units larger than the syllable. The speech actions aimed at the production of segments belonging to the same syllable have been proposed to present a particular cohesion (Browman & Goldstein, 1988) and this statement is supported by many studies on intrasyllabic coarticulation (Farnetani & Recasens, 2010). However, if anticipatory V - to-V coarticulation is planned, the extent and/or scope of the coarticulatory influence would reflect the size of the unit on which the movements for the production of a speech target are anticipated, here a unit encompassing two syllables. Evidence of language-specific patterns in coarticulation can indeed be interpreted as evidence of cross-linguistic differences in the units of speech planning. This is precisely the argument used by Ma, Perrier, and Dang (2015) to suggest that the syllable plays a different role in planning in French and Mandarin Chinese. In their investigation, they examined anticipatory V-to-V and V-to-C coarticulation in V_1CV_2 sequences in both languages, showing that, if both speakers of French and Mandarin Chinese exhibited anticipatory V-to-C intrasyllabic coarticulation, only French speakers exhibited anticipatory extrasyllabic V-to-V coarticulation. The lack of coarticulation beyond the syllable level in Mandarin Chinese suggests that the syllable has a stronger influence on speech planning in this language than in French.

The notion of V-to-V coarticulation as a planned and controlled process carries some implications and raises some questions. Indeed, albeit the precise scope of V-to-V coarticulation is still unclear, evidence has been adduced for coarticulation to span over a long stretch of segments and to extend beyond the word level, suggesting that coarticulation is controlled over a large speech plan (e.g. Abry & Lallouache, 1995; Grosvald, 2009). However, the scope and the degree of anticipatory coarticulation varies according to speaker-dependent factors such as speakers' identity (Abry & Lallouache, 1995; Grosvald, 2009; van den Heuvel, Cranen, & Rietveld, 1996; Robert, Dautcourt, Laprie, & Bonneau, 2005; Zellou, 2017; Guitard-Ivent, 2018a) or age (e.g. Barbier, 2016; Barbier *et al.*, 2020; Noiray, Abakarova, Rubertus, Krüger, & Tiede, 2018), but also according to production-dependent factors, such as speech rate (e.g. Hertrich & Ackermann, 1999; Recasens, 2015) or prosody (e.g. Cho, 2004; Guitard-Ivent, Turco, & Fougeron 2021).

What are the implications of these variations in coarticulation for speech planning? What is the size of a unit over which coarticulation is controlled? How coarticulation is regulated within such a unit? Does a change in coarticulation corresponds to a change in the size of these planning units?

In this dissertation, I attempt at tackling these questions by investigating interspeaker variations of acoustic cues of V-to-V extrasyllabic coarticulation in French. In particular, interspeaker variations are analyzed according to population, particularly according to Motor Speech Disorders and age in adult speakers; and according to speech rate and prosodic boundary across multiple repetitions in a selected set of speakers.

Throughout all studies, the linearity of the relationship between coarticulation and features linked to the temporal organization of speech such as vowel length, vowel-to-vowel lag, or articulation rate is also explored. Indeed, a change in the temporal organization of speech, which can go along with other factors of variations, could lead to - and at the same time reflect - a change in the degree of overlap between segments, and thus a change in coarticulation degree.

2 Theoretical framework on coarticulation and research questions

2.1 Theories of coarticulation

Over the years, several theoretical accounts of anticipatory coarticulation have been proposed (see reviews in, e.g., Farnetani & Recasens, 2010; Mildner, 2018). Overall, two main categories of models can be identified in the literature, which differ in the way the process underlying coarticulation is explained: look-ahead and coproduction models. In the look-ahead category, besides the look-ahead model by Henke (1966) can be included the feature-spreading model (Daniloff & Hammarberg, 1973), the window model (Keating, 1990), and models based on optimal planning such as GEPPETO (Perrier & Ma, 2008; Perrier, 2014; Winkler, Ma, & Perrier, 2010) and DIVA (Guenther 1994, 1995; Guenther *et al.,* 1998; Guenther, Ghosh, Nieto-Castanon, & Tourville, 2006). Coarticulation is considered the result of a look-ahead process: the movements for the production of a speech goal can be started in advance as long as the articulatory movements for the production of the preceding ones allow it. The account of coarticulation as coproduction has been elaborated over the works of different researchers, such as Öhman (1966, 1967), Fowler (1980), Bell-Berti and Harris (1981), and in the Articulatory Phonology framework (Browman & Goldstein, 1986, 1988, 1990) and the Task Dynamics model (Saltzman & Munhall, 1989; Fowler & Saltzman, 1993). Coproduction models consider coarticulation as the result of the coproduction or overlap of spatiotemporally defined speech goals. These two ways of considering coarticulation have different implications as for the scope of coarticulation. Indeed, in a look-ahead model, the scope of coarticulation depends uniquely on the preceding context, while in a coproduction model speech goals have each their timing, therefore the scope of coarticulation is tendentially fixed.

In the next paragraphs, the main concepts and accounts of the look-ahead and the coproduction theories of coarticulation will be presented. In particular, I will focus on how coarticulation is accounted for in the DIVA model (Guenter, 1995) and in the Articulatory Phonology/Task Dynamics model (Browman & Goldstein, 1990, Fowler & Saltzman, 1993)

2.1.1 Coarticulation as a look-ahead process

2.1.1.1 The look-ahead model

The look-ahead model of coarticulation, originally proposed by Henke (1966), but similar to the feature-spreading model by Daniloff and Hammarberg (1973), considers that the articulatory target for a segment is specified in terms of binary phonological features, where unspecified features are 0. The movement for a feature of a later segment can be anticipated in all the preceding segments that are unspecified for this feature, with a scanning look-ahead mechanism. As it was, this model has raised some perplexities. For instance, the notion that a segment can be anticipated in a preceding one only if this one is unspecified for a feature was challenged by data on V-to-V coarticulation, which showed that anticipation can occur in a vowel specified for a contradictory feature (i.e. in early studies, Berenguel & Cowan, 1974; Öhmann, 1966). However, with some modifications, mainly in the way speech goals are defined, the basic concept of this model has been adopted by later accounts of coarticulation.

2.1.1.2 The window model

In her window model of coarticulation, Keating (1990) proposed that, for each physical dimension, segments do not have one possible value, but a range of possible spatial values. The target for each segment is therefore represented by the entire range of values, which defines the contextual variability allowed for that segment, and it is called a window. Segments associated with large windows allow for great contextual variability, while segments associated with small windows allow for little variability. For a given dimension, a sequence of segments corresponds to a sequence of windows, which are connected by interpolation functions called paths. Most of the path falls inside a segment, while some part falls into narrow transition zones between the segments. Therefore, windows are ranges in which a path is allowed to fall. The choice of the path between two windows responds to the principle of minimal effort, in terms of displacement or peak velocity. Keating's model combines the principle of look-ahead models with the principle of minimal articulatory effort, which will be further developed in models of speech production based on optimal planning.

2.1.1.3 Models of optimal planning: GEPPETO and DIVA

The concept of optimal planning arises in models of motor planning of limb movements (see reviews in Grimme *et al.,* 2011; Perrier, 2012) and has been then applied to models of speech productions such as the GEPPETO model (Perrier & Ma, 2008; Winker *et al.,* 2010; Perrier, 2014) and the DIVA model (Guenther, 1994, 1995; Guenther *et al.,* 1998; Guenther *et al.,* 2006). The basic assumption is that the central nervous system would use internal representations of the speech production apparatus to plan the articulatory movements aimed at the production of a speech target in a way that will entail the minimal effort for the articulators (Winkler *et al.,* 2010). Therefore, coarticulation would be the result of a process of planning that would lead the speaker to minimize articulatory effort by decreasing the displacement of the articulators when possible. Similarly to the window model, in the GEPPETO and DIVA models the speech goals correspond to ranges/windows that include all the variability allowed in the production of a target. However, instead of a range of values, targets are multi-dimensional.

2.1.1.3.1 The GEPPETO model

In the GEPPETO model (GEstures shaped by the Physics and by a PErceptually oriented Targets Optimization, e.g. Perrier & Ma 2008; Perrier 2014; Winker *et al.,* 2010), speech goals are formant targets in the acoustic space, while motor targets are associated with the acoustic speech goals. Each target is described as a 3D ellipsoid in the acoustic space of the first three spectral maxima, F1, F2, F3. The canonical realization of the target corresponds to the center of the ellipsoid, while the possible variations are determined by the standard deviation in the three dimensions. The motor commands for the production of subsequent speech goals are selected thanks to an optimal planning, where the minimum of effort is defined as the minimization of the distance between the speech targets. There are two constraints to this minimization of the distance, and therefore to contextual variability. The first one relates to perceptual accuracy requirements. In other words, optimization would lead to a high degree of coarticulation, as long as the perception of the desired speech target is not in danger (Barbier, 2016). The second one relates to the global level of force required to achieve the target: the force level must remain within a given range during the execution of the movement (Winkler *et al.*, 2010).

2.1.1.3.2 The DIVA model

In the DIVA model (Directions Into Velocity of Articulators, Gunther 1994, 1995; Guenther *et al.*, 1998; Guenther *et al.*, 2006) vocal tract targets for the production of each speech sound correspond to convex target regions defined within the orosensory space. Convex regions are multidimensional: for any two points in the region, all the points in the line that connects them are included in the region. Each dimension of the orosensory target thus defined specifies a range of acceptable positions for the production of that sound. The appropriate range of positions that the vocal organs can assume for the production of each phoneme is learned in the phase of speech development (that is the babbling phase in the model) by associating the orosensory target to the corresponding target phoneme, represented by a Speech Sound Map (SSM) cell. Each SSM cell codes a different speech sound. Therefore, each speech target is bimodal: the acoustic target corresponds to the SSM cell, while the orosensory target to the convex region target. The two are connected through a phonetic-to-orosensory mapping. The desired speech target defined by the convex target region is then produced thanks to an orosensory-to-articulatory mapping, through which the appropriate movements of the articulators are defined.

Coarticulation in the DIVA model (Guenther, 1995) is done by shrinking the convex region target for the coarticulated phoneme (anticipatory coarticulation), or by defining a specific trajectory between subsequent convex regions (carryover coarticulation). The anticipation of a speech target in a preceding one is done through the shrinking of the convex region target in order to obtain a reduced target that overlaps with the speech target that is anticipated on one or more dimensions. For each orosensory dimension, the baseline target for a phoneme corresponds to its habitual target range. If an upcoming phoneme can be anticipated on a specific dimension, the range reduces in size in the region of overlap, and the same procedure is done for the next target (Figure 1).

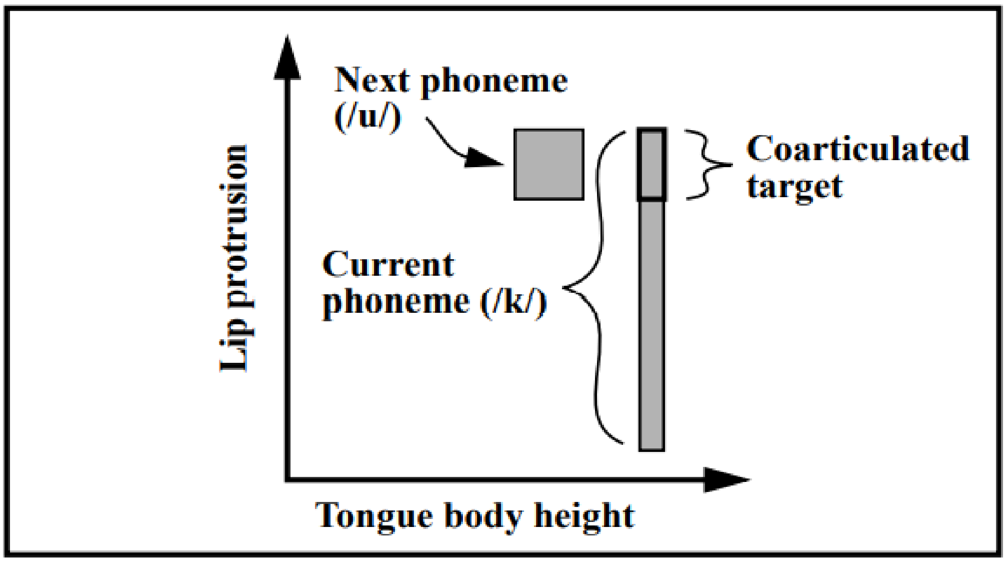

Figure 1. Representation of the anticipation of /u/ during the production of /k/ in the word "coo" in DIVA. The two convex region targets overlap in the dimension of lip protrusion. Therefore, the convex region target for /k/ is shrunk to include the overlapping portion in the lip protrusion dimension. From Guenther (1995).

Because the configuration used to produce a phoneme cannot extend beyond the limits of the convex region target, the amount of coarticulation allowed for each phoneme depends on its convex region. In sum, anticipatory coarticulation is planned through a look-ahead process in order to provide a more efficient sequence of articulatory movements.

Carryover coarticulation results from the moving from a target to the following one, thus does not stem from an explicit planning mechanism. Indeed, since the movements for a target start in the preceding one, the target configuration depends also on the configuration assumed by the preceding convex region target. In other words, the particular configuration assumed by the vocal tract when the movement for the production of a following target starts will determine the final configuration of the vocal tract for that target. If, along a orosensory dimension, the targets for two phonemes overlap, that specific articulator does not need to be moved and can stay in the position it had already assumed for the first target. Since carryover coarticulation is a consequence of the dynamics of moving between subsequent speech targets, it is not planned as anticipatory coarticulation. However, it is not the product of inertia, because it results from explicit commands to the articulators. In Guenther (1995) an opposition is made between a pre-planned anticipatory coarticulation and a planned carryover coarticulation, in the sense that both are controlled, but in different ways.

2.1.2 Coarticulation as coproduction

2.1.2.1 Coarticulation as coproduction of vowels and consonants

In his acoustic investigation of 1966 on VCV coarticulation in three languages, Öhman observed that the formant pattern of the VC part of the utterance was influenced by the identity of the postconsonantal vowel. He hypothesized that the consonant was superimposed on a continuous movement between the vowels, which was the substrate of articulation. The coproduction of consonants and vowels would be made possible by the fact that they are produced by different articulatory systems that are fairly independent of each other (Öhman 1966, 1967). The idea that consonants and vowels are inherently different is adopted by later coproduction models (e.g. Fowler, 1980; Browman & Goldstein, 1990).

2.1.2.2 Intrinsic timing view of speech production

Fowler (1980) postulated that speech goals are four-dimensional entities defined by the coordinative structures of muscles that contribute to their realization. Coarticulation stems from the overlap of these continuous four-dimensional segments that present their intrinsic timing. Therefore, the contextualization of a speech goal does not result from its modification as in look-ahead theories, but merely from its coproduction with adjacent segments who also have an inherent duration. The concept of intrinsic timing is further developed by Bell-Berti and Harris (1981). They postulate that coarticulation arises because the movements for the production of a segment start earlier than its acoustic onset and finish later than its acoustic offset. For each articulator, the time of anticipation is fixed and independent of the preceding context, if there is no articulatory conflict (while, in case of articulatory conflict, the fixed time can be modified and the movement will start later). Therefore, coarticulation is time-locked.

2.1.2.3 Articulatory Phonology and Task Dynamics

The fundamental assumption of the Articulatory Phonology approach (Browman & Goldstein, 1986, 1988, 1990, 1992) is that the basic phonological units are articulatory gestures, which are linguistically significant events that unfold during speech production. Gestures consist in coordinated movements, and specifically, in the formation and release of constrictions in the vocal tract, which are aimed at the achievement of a significant speech goal. Gestures thus defined have been implemented in the Task Dynamics model (Saltzman & Kelso, 1987; Saltzman & Munhall, 1989), in which gestures are characterized by using dynamical equations. Gestures are organized in larger structures called gestural scores, which specify the appropriate gestures for the production of a given utterance, and the values of the dynamic parameters of the selected gestures. Inside a gestural score, phasing rules coordinate gestures that are associated with each other. Association is done between gestures that contribute to the production of the same segment and between gestures to produce different segments (from here on, intergestural coordination). Finally, inside these gestural scores, gestures are organized into articulatory tiers: gestures that are articulatory independent are on different tiers. For example, as for oral gestures, there is a tier for the lips, one for the tongue body and one for the tongue tip. In Browman & Goldstein (1990) vowels and consonants are specified to be on different tiers.

How coarticulation works in this framework is specified in Browman and Goldstein (1990) and especially in Fowler and Saltzman (1993). Coarticulation results from the overlap of underlying invariant gestures aimed at the production of different segments. Gestures are spatiotemporally defined, and present activation waves that are smoothly shaped. That is, each gesture has a phase of gradual implementation, a phase in which it is maximally active, and a relaxation phase, so the influence of each gesture on the vocal tract waxes and wanes gradually. The phase in which the gesture is maximally active corresponds to the actual phase of production, while the span of anticipatory coarticulation corresponds to the phase of implementation of the gesture, and the span of carryover coarticulation corresponds to the phase of relaxation of the gesture (Figure 2).

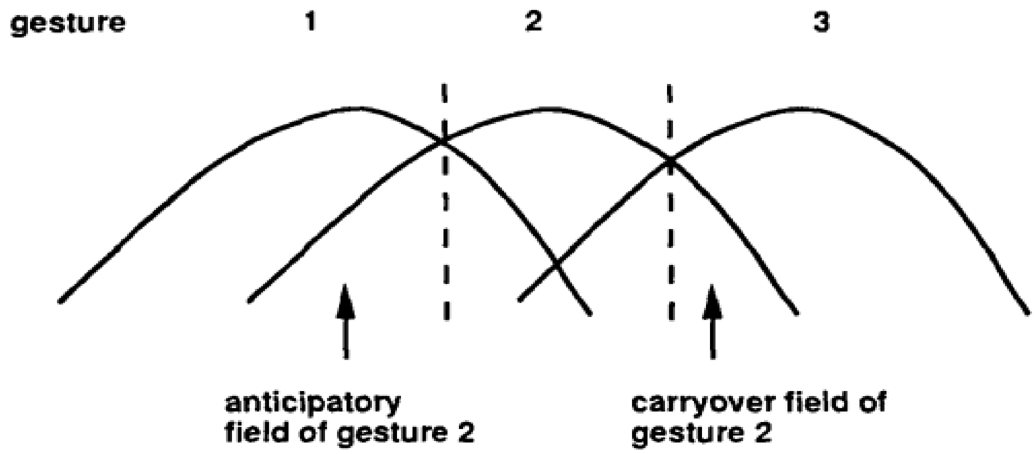

Figure 2. Representation of the activation waves of three overlapping gestures. The dotted lines indicate the fields for anticipatory and carryover coarticulation. From Fowler and Salzman (1993).

Since the goal of a gesture must be invariantly achieved, this coproduction is context-sensitive. That is, in the production of a sequence of gestures, the overlap between one gesture and a nearby one will involve either the articulators that are not been used for the production of the intended gesture, or the articulators that will not interfere with the achievement of the gestural goal (Fowler & Salzman, 1993). In the first case, overlap happens between gestures that do not use the same articulators and therefore are on different tiers: for example, between a labial consonant (tier for the lip gesture) and a vowel (tier of the tongue body/tongue tip gesture). In the second case, there is a partial overlap of gestures on the same tier, which leads to a blending of the characteristics of the two gestures: for example, between two vowels (Browman & Goldstein, 1990). In Task Dynamics, each gesture has a degree of blending strength, which is inversely correlated to sonority, so that open vowels are the weakest. When there is coproduction between two gestures that share the same articulator, the output of the blending will depend on the blending strength of the gestures involved. If the two gestures have similar strength, i.e. vowels, the blending results in an averaging of the characteristics of the two vowels. If the two gestures have different blending strength, i.e. a velar consonant and a vowel, the stronger gesture suppresses the weakest one (Fowler & Saltzman, 1993).

2.2 Short notice on the notion of planned V-to-V coarticulation

2.2.1 Look-ahead vs coproduction perspective

In reviewing the main theories on coarticulation, I have focused on two perspectives that have been adopted in the literature, the look-ahead and the coproduction perspective. In particular, it is presented more in detail how coarticulation is addressed in the DIVA model (Guenther, 1995), for the look-ahead perspective, and the Task Dynamics model (Fowler & Saltzman, 1993) for the coproduction perspective. Anticipatory V-to-V coarticulation can be considered to be planned whether we adopt a look-ahead or a coproduction point of view, but there are some fundamental differences.

In a generalized look-ahead account such as DIVA, if there is a region of overlap between two convex region targets, the target for the production of the first phoneme can be modified in order to minimize the trajectory toward the target for the production of the second phoneme. If in a speech sequence V_2 is anticipated in V_1, this means that the convex region target of V_1 has shrunk in the direction of the target for V_2. Therefore, coarticulation is explicitly planned.

In a coproduction account such as Task Dynamics, since gestures are smoothly shaped, if in a speech sequence V_2 is anticipated in V_1, this means that the gesture for V_2 is in its implementation phase, so it is already active. In other words, the movement for V_2 has already started during the production of V_1, so there is a temporal overlap. Invariant speech gestures are phased with each other at the creation of a gestural score, that is, before production.

Three fundamental differences between these accounts can be listed:

1. **Spatial vs temporal.** In DIVA, coarticulation is seen in a spatial perspective, while in Task Dynamics coarticulation in seen in a temporal perspective.
2. **Necessarily vs non-necessarily planned.** In DIVA, coarticulation is explicitly planned, and has to be planned to be implemented. If the target for V_1 is not modified to anticipate V_2, coarticulation does not happen. In Articulatory Phonology/Task Dynamics, gestures that are phased with each other will display overlap. However, some overlap can be expected to occur also between gestures that are not phased with each other. This overlap is expected to be more variable, because the two gestures are not coordinated.
3. **Variable vs fixed scope.** For DIVA, the scope of coarticulation can be variable, since it can span over all the targets that share a region of overlap on a given dimension. In Task Dynamics, the scope of coarticulation is fixed, because gestures have their own intrinsic duration. In case of articulatory conflict, a gesture can start later, but the maximum scope of coarticulation will correspond to the time of implementation of the gesture.

2.2.2 Planning and planning unit in this dissertation

In this dissertation, I will move from a coproduction perspective. At the beginning of the introduction, a notion of planning borrowed from studies on motor control has been adopted, as the processes of preparation of movements that occur before the initiation of the movement (Grimme *et al.,* 2011). In V-to-V coarticulation, this means that the movements for the production of an upcoming vowel are anticipated in the production of the preceding vowel. From a coproduction perspective, in a planned V-to-V coarticulation, the overlap between the two vocalic gestures can be considered to be controlled. Therefore, the coordination or phasing between the two vowels would be somehow specified.

The notion of planning unit that will adopted in the rest of this dissertation relates to this definition of planning. That is, when speaking of a planning unit, it is meant the unit over which the movements for the production of a serie of speech targets are planned and over which coarticulation is specified. Specifically, over which the coarticulation between V_1 and V_2 is specified.

2.3 Research questions

In this dissertation, variations of anticipatory V-to-V coarticulation depending on population-specific factors, such as pathology and age, and on speaker according to speech condition, such as speech tempo and boundary type, are investigated.

Population-specific effects on coarticulation are investigated in two studies. In the first study **(MSD study**, chapter 5**)** anticipatory V-to-V coarticulation is examined in speakers affected by different types of Motor Speech Disorders, namely Apraxia of Speech and Dysarthria associated with Amyotrophic Lateral Sclerosis, Wilson Disease and Parkinson's Disease, and neurotypical controls, and it is compared to C-to-V carryover coarticulation. In the second study **(age study**, chapter 6**)** the relationship between the degree of V-to-V anticipatory coarticulation and articulation rate is investigated in a population aged 20 to 93 (this study is the subject of a published article, D'Alessandro & Fougeron, 2021).

In two other studies, the individual responses of five speakers to different speech conditions are investigated. In the third study **(tempo study**, chapter 7**)**, interspeaker variations of V-to-V coarticulation are examined across three different speech tempos, a self-paced comfortable one, slow, and fast. In the fourth study **(boundary study**, chapter 8**)**, inter and intraspeaker variations in V-to-V coarticulation are investigated across three different boundary types, syllable boundary within a word, word boundary and clause boundary, over 45 repetitions.

In its entirety, this dissertation aims at addressing two main questions.

The first main question of this dissertation is:

- To what extent V-to-V coarticulation covaries with articulation rate?

If coarticulation is considered as the overlap between spatiotemporally defined units, coarticulation and rate are expected to be the two faces of the same coin. Indeed, a change in coarticulation can be seen as a change in the temporal unfolding of the speech units and as a consequent change in rate. In the literature, this matter has been mainly addressed by studies that elicited either an increase in rate, either a decrease in rate, as part of the protocol. These studies, in particular the ones that analyzed V-to-V coarticulation, do not show a straightforward correlation between V-to-V coarticulation and speech rate, especially at slow rate. Moreover, some interspeaker variation emerges.

Here this question is addressed in two ways. On one hand, the relationship between coarticulation and rate is investigated in the speech of speakers who naturally present differences in their speaking rate, that is, speakers affected by Motor Speech Disorders compared to neurotypical speakers, and neurotypical speakers of varying age, in order to investigate whether rate and coarticulation are related in a "non-artificial" setting. Simply put, it will allow answering to the question: do speakers who speak naturally slower coarticulate less than speakers who speak naturally faster (and viceversa)?

On the other hand, the relationship between coarticulation and rate is investigated in the individual responses of a small set of speakers to a speech tempo manipulation, in oder to isolate the effect of rate on coarticulation and to explore interspeaker variations in the way speakers "manage" their coarticulation at a slow and at a fast rate. Moreover, a question raised by studies addressing the relationship between rate and kinematics is whether an increase and a decrease in speech rate rely on the same mechanism (see section 3.1.2). If previous studies on the effects of changes in rate on V-to-V coarticulation showed some speaker-dependent effects, in particular reporting no effects of rate on coarticulation for some speakers, they have mainly studied changes in rate in one direction, either up either down. Investigating the effects of a slowing of rate and of an acceleration of rate for the same speaker will allow addressing this question.

The second main research question of this dissertation is:

- What are the units of speech over which the movements for a string of speech targets are planned?

In a coproduction perspective, coarticulation can be planned/controlled or can be the result of an uncontrolled random overlap between speech gesture. The degree of coarticulation, and the stability of coarticulation could be an index of within-unit cohesion, thus indicating whether a particular coarticulation reflects the coordination between elements belonging to the same unit. In other words, the extent, the scope, and the stability of the coarticulatory influence would reflect the size of the unit on which the movements for a string speech targets are anticipated.

Coarticulation is investigated between two words and two clauses as opposed to between two syllables within a word in order to investigate what are the units over which coarticulation is planned in French. Studies that have investigated coarticulation across prosodic boundaries have generally compared coarticulation between words to coarticulation across boundaries of increasing strength. Therefore, it remains an open question whether V-to-V coarticulation within a word differ from V-to-V coarticulation between two words. Moreover, inter-speaker variations in coarticulation across prosodic boundaries has been previously found. It is investigated in particular: whether coarticulation within a word differ from coarticulation across two words, to test the status of the word as unit over which coarticulation is planned; and whether interspeaker variations in coarticulation are related to specificities of the prosodic phrasing, to investigate how units where coarticulation is planned interact with prosodic units, that is, units where other phonetic details and the rhythm of the utterance are organized.

Moreover, the units of planning in speech could vary according to speaker or pathology. Indeed, a reduction of coarticulation in AoS has been attributed to an impairment in speech planning. V-to-V extrasyllabic coarticulation is compared to CV intrasyllabic coarticulation in pathological speakers to investigate whether AoS speakers plan on smaller units than neurotypical speakers.

More specific goals and hypotheses will be presented at the beginning of the chapters dedicated to each study. However, the results will be discussed in their entirety in the discussion chapter.

In the next chapter, the main findings related to the effects of rate changes and prosody on coarticulation, and to population-specific effects on coarticulation, will be reviewed.

3 Background: factors of variation in coarticulation

3.1 Effects of rate on coarticulation

Changes in speech rate and changes in coarticulation degree can be expected to be correlated. If coarticulation is seen as the overlap between speech units spatiotemporally defined (i.e. coproduction framework), that is, presenting their own duration, an increase in overlap would make gestures “slide” more into one another and thus would reduce the duration of the sequence, while a decrease in overlap would produce a lengthening of the sequence duration. In other words, a change in coarticulation degree could be seen as a change in the temporal unfolding of the speech gestures and as a consequent change in rate (following e.g. Browman & Goldstein, 1990). Nevertheless, even though intuitively we could expect coarticulation to linearly change with speech rate, the literature shows that the relationship between rate and coarticulation is not that linear. If a tendency toward an increase of coarticulation at fast rate and a decrease of coarticulation at slow rate is found for intrasyllabic vowel-to-consonant coarticulation, results on V-to-V coarticulation are not straightforward, especially at slow rate.

The next sections will review the main findings on the effects of an increase in rate and of a decrease in rate on CV and V-to-V coarticulation, and on the effects of rate changes on the kinematics of speech movements.

3.1.1 Coarticulation across speech rate manipulations

Studies eliciting an intentional change in speech rate as part of the protocol have generally focused on rate increases, showing a general increase in coarticulation. However, the effects are more consistent for the anticipation of a vowel in a preceding consonant, in CV sequences, or for the overlap between two consonants in CC sequences (e.g., Gay 1978; Engstrand, 1988; Agwuele, Sussman, & Lindblom, 2008, Hardcastle, 1985; Byrd & Tan, 1996) than for the anticipation of a vowel in a previous one (e.g. Matthies, Perrier, Perkell, & Zandipour, 2001; Recasens, 2015; D'Alessandro, Bourbon, & Fougeron, 2020).

An increase in intrasyllabic CV coarticulation at fast rate has been found in acoustic studies on English in which coarticulation has been measured on F2, often with locus equations (linear regressions of the frequency of F2 sampled at vowel onset, on the frequency of F2 sampled at vowel nucleus). Gay (1978) found that, in CV sequences, F2 frequencies at vowel midpoint were closer to vowel onset frequencies at fast rate than normal rate. In graded-rate tasks, where participants were asked to progressively speed up their speaking rate, starting from their habitual rate, Tjaden and Weismer (1998) and Agwuele, Sussman, and Lindblom (2008) obtained similar results on CV sequences, showing a gradual increase in coarticulation, again measured on F2 of the vowel, with an increase of rate. The same pattern is showed in an x-ray study on Swedish by Engstrang (1988). He analyzed tongue movements in /ipi/, /apa/ and /upu/ sequences at normal and fast rate, reporting that, if at normal rate there was a relaxation of the tongue-related movement during the consonant, this was absent at fast rate, due to an increase in coarticulation with the flanking vowels. An increase in overlap at fast rate has also been reported for consonant clusters in English (Hardcastle, 1985; Byrd & Tan, 1996). In particular, Byrd and Tan (1996) showed that the overlap between the two consonants in different CC clusters gradually increased in four progressively faster speech rates, defined "normal", "medium", "faster", and "fastest".

A tendency toward an increase of coarticulation at fast rate is shown also in studies on V-to-V anticipatory coarticulation. However, in this case, the effects of rate appear less striking. In an articulatory study, Matthies, Perrier, Perkell, and Zandipour (2001) examined the anticipation of V_2 in V_1 in /iC_nu/ sequences in English, where C_n was a number of consonants varying from 1 to 3, at a normal and a fast rate. In their results, they reported an increase in labial, but not lingual coarticulation at fast rate with respect to normal rate. Some inter-speaker differences in coarticulation were equally reported. Recasens (2015) found a slight (barely significant) increase in lingual anticipatory coarticulation between /a/ and /i/, as measured by F2, at a fast rate in CVC sequences in Catalan. In a study on the effects of fast rate on sentence repetition as a function of age group in French (D'Alessandro *et al.,* 2020) we found a speaker-dependent effect of rate on coarticulation. Indeed there was an increase of anticipatory V-to-V coarticulation in younger speakers (<40 y.o.a.), but not in older speakers (>68 y.o.a.) who increased repetition rate without significantly changing coarticulation degree (I will return to this results in the discussion).

In general, fewer studies have looked at the effects of a slowing of rate on coarticulation. Acoustic investigations on local coarticulation showed a tendency toward a decrease of local CV or VC coarticulation at slow rate, but the results are less systematic than the ones reported for the fast rate. In two studies employing a graded rate task, Tjaden and Weismer (1998) and Weismer and Berry (2013) found a gradual decrease of anticipatory CV coarticulation with the progressive rate slowing. However, in Weismer and Berry (2013) these results concerned three speakers out of four: one speaker did not change her coarticulation degree. A decrease in anticipatory extrasyllabic VC, but not anticipatory intrasyllabic CV coarticulation has been reported in Tjaden and Wilding (2005), in a study comparing coarticulation in neurotypical and pathological speakers at a normal and a slow speech rate. They analyzed, in V_1CV_2 sequences formed by the English article /ə/ + a monosyllabic word, the influence of C on V_1 and of V_2 on C, reporting an effect of rate slowing for the first, but not for the second type of coarticulation.

The results reported by acoustic studies investigating V-to-V coarticulation at slow rate are even less consistent, showing no change in coarticulation, a decrease, or yet an increase of coarticulation. Hertrich and Ackermann (1995) analyzed the effect of a slowing of rate on anticipatory and carryover V-to-V coarticulation in productions of

/gətVtə/ pseudowords in German. The results depended on coarticulatory direction and speaker: if all six speakers showed a linear reduction of carryover coarticulation at the slowing of rate, anticipatory coarticulation was either unchanged, or increased at slow rate, depending on the speaker. In the aforementioned study by Tjaden and Wilding (2005), they examined also V-to-V anticipatory coarticulation on the same V_1CV_2 sequences /əCV_2/ where V_2=/i, ɑ, u/, reporting vowel-dependent results. Coarticulation was reduced at slow rate when V_2 was /i/ but increased when V_2 was the back vowels /ɑ/ and /u/.

Overall, the literature reports a linear increase in intrasyllabic coarticulation with an increase in speech rate, and a less systematic tendency toward a decrease in intrasyllabic coarticulation at slow rate. More complex is picture painted for V-to-V coarticulation. A fast rate, the results show that an increase in rate can correspond, but does not necessarily correspond, to an increase in coarticulation degree. At slow rate, the results points in very different directions. For both fast and slow speech, various factors seem to be at play to influence the effect of rate on V-to-V coarticulation, such as speakers' identity (Matthies *et al.*, 2001; Hertrich & Ackermann, 1995) and age (D'Alessandro *et al.*, 2020). The different results showed for CV and V-to-V coarticulation could be related to the different domains of coarticulation, that is, within a syllable as opposed to between two syllables. However, the inconsistency of the results on V-to-V coarticulation can be also considered in the light of the results on the effects of rate on the kinematics of speech movements.

3.1.2 Variety of articulatory strategies to modify speech rate

Observation of speaker-specific kinematic responses to speech rate manipulations shows that speakers can achieve a change in speech rate by using different articulatory strategies (review in Berry, 2011). Indeed, two ways in which a change in speech rate can be achieved are by modifying the overlap between gestures or by modifying articulatory displacement. In particular, an increase in rate could be achieved through an increase in overlap (as seen in the previous section) or a decrease in displacement (e.g. Kent & Moll, 1972; Flege, 1988) which would translate in vowel undershoot (Lindblom, 1963; Moon & Lindblom, 1994; Fourakis, 1991). These two strategies could equally result in a change in coarticulation patterns. However, other strategies have been reported for the control of speech rate, such as a change in velocity, with an increase of peak velocity at fast rate and a decrease at slow rate (e.g. Kuhen & Moll, 1976; Van Son & Pols, 1992; Adams, Weismer & Kent, 1993; Hertrich & Ackermann, 2000; McClean, 2000) or a change in both peak velocity and displacement, in a directly or, more often, inversely correlated fashion (e.g. Munhall, Ostry & Parush, 1985; McClean & Tasko, 2003). Ostry and Munhall (1985) analyzed with ultrasound tongue dorsum movements in the production of CV syllables where C=/k, g/ and V=/a, u, o/ by three speakers at a slow and a fast rate. At fast rate, two of the subjects reduced movement amplitude, maintaining peak velocity unchanged. The third subject increased peak velocity without presenting changes in movement amplitude. If these two behaviors appear as the two main mechanisms to control speech rate, a variety of strategies can result by their interaction, as shown by the results of Goozée, Lapointe, and Murdoch (2003), who examined with EMA repetitions of /ta/ and /ka/ syllables by eight speakers at a moderate rate and at a fast rate ("as fast as possible"). At fast rate, seven speakers showed a reduction in the distance travelled by the tongue. Among these seven speakers, four exhibited no change in velocity, while two showed a decrease in velocity, and one an increase in velocity. One participant showed no change in displacement, but a marginal increase in velocity. An increase in peak velocity without changes in movement amplitude would lead the speaker to reach the target in less time, without modifying the target.

Another aspect to be considered is that the strategies to increase and decrease rate could not mirror each other. McClean (2000) examined upper and lower lip, tongue and jaw movements in the repetition of a sentence at slow, normal and fast speech rate by nine speakers. He reported overall decreased velocity for slow speech, but either no change in velocity, or a decrease or an increase in velocity for fast speech. The difference between these conditions is emphasized also in studies that showed asymmetrical velocity profiles at slow speech. Adam, Weismer & Kent (1993) analyzed the velocity profiles of tongue tip closing and opening gestures in /tap/ and /tad/ (in the sentence *tap a tad above*) across different slow and fast speech rates, showing different profiles for fast and slow speech: if velocity profiles at fast rate were symmetrical and characterized by one peak, at slow rate they were asymmetrical and characterized by multiple peaks. The same pattern is reported by Perkell, Zandipour, Matthies, and Lane (2002): in a comparison between the production of sentences in slow, normal, fast and clear conditions, they showed that, at slow speech, the patterns of movements were different than the ones found for the other conditions. The slow condition was characterized by having the same displacement as the normal condition but with multiple velocity peaks, which suggested less effort, and less smoothness, in the production of a slow rate with respect to a normal or fast rate. The authors interpret these data as evidence that slow speech is more unnatural for speakers, hence the different movement pattern.

Taken together, these results suggest that the relationship between rate, intergestural coordination and kinematics is rather complex, and this complexity has to be taken into account in considering the mixed patterns of coarticulation observed with changes in rate.

3.2 Effects of prosody on coarticulation

The phonetic realization of vowels and consonants is known to change when they are produced in certain positions in a prosodic domain, such as the initial and final positions, and when they are produced in prominent syllables. At prosodic boundaries and under prominence, segments are lengthened and present a more extreme articulation (*inter alia,* Fougeron, 1999; Cho, 2011). The magnitude of the changes that segments undergo when produced in proximity to a prosodic boundary is considered to be proportional to the strength of the prosodic boundary, that is, the degree of prosodic disjunction between adjacent prosodic units (Cho, 2011; Cho, 2016). Studies investigating the effects of prosody on coarticulation showed that coarticulation degree is modulated by boundary proximity and prominence, but the results are not systematic. Across prosodic boundary between two words, there is a tendency toward a decrease in coarticulation with the increase of boundary strength, but some individual variation has been reported.

The matter of how coarticulation vary as a function of prosodic position can be addressed by two angles. Indeed, prosodic constituents act as units of organization of speech, over which some aspects of speech production, such as F0 contours and rhythm are planned. At the same time, prosodic boundaries and prominence are characterized by a change in the temporal fabric of the utterance, in that they slow the temporal unfolding of gestures, that become longer and increase in magnitude. Changes in coarticulation between sounds at different positions in the prosodic domains can be related to these two linked aspects. Indeed, it could be asked whether the domain over which the coordination between speech gestures is organized, that is, a unit over which coarticulation is planned, coincides with prosodic constituents, and to which ones. If changes in coarticulation at the edges of prosodic boundaries or across prosodic boundaries would mark the edges of units over which the movements for the articulation of a string of speech sounds are planned, one could wonder whether the effects of prominence and of boundaries on coarticulation have to be accounted by the same mechanism, that is, a localized change in the temporal organization of speech.

In the next sections, it will be first presented how the effects of prosody on coarticulation have been modeled in a coproduction account, through the action of a prosodic gesture and a modulation gesture, then the main results on coarticulation under prominence, in domain initial and domain final position of a prosodic domain, and across a prosodic boundary will be reviewed.

3.2.1 Modeling the effect of prosody on coarticulation: the π-gesture and the μ-gesture

Byrd and Saltzman (2003), within the Articulatory Phonology/Task Dynamics framework, proposed that prosodic boundaries are instantiated by a π-gesture or prosodic gesture. A π-gesture has no independent realization like the constriction gestures, but acts trangesturally, altering the time flow of all constriction gestures that are concurrently active. Specifically, it acts slowing down and retarding the activations of co-active gestures, which become temporally longer and less overlapped. This will also cause the gesture to be less undershoot and thus show larger spatial magnitude. It is important to note that π-gestures differ only in strength of activation, that is, there are not different types of prosodic gesture. The strength of the π-gesture depends on boundary strength, so higher boundaries, that is, boundaries that encode greater disjunction, have stronger activation. Moreover, the activation strength of the π-gesture is maximum at phrasal edges, so the effects "wear out" the further gestures are from the boundary (Byrd & Saltzman, 2003; Krivokapić, 2020; Byrd & Krivokapić, 2021). Therefore, this account would predict a decrease of coarticulation at the edge of a prosodic domain and across a prosodic boundary, which would be stronger for consonant and vowels closer to the boundary. Coarticulation would decrease at the increase of boundary strength.

To model also the effect of prominence, and not only of boundaries, on constriction gestures, an extension of the π-gesture model has been proposed, where prominence is instantiated by a clock slowing gesture, called modulation gesture or μ-gesture (Saltzman *et al.*, 2008). The μ-gesture also acts on the constriction gestures slowing their temporal unfolding: under prominence, gestures become longer and larger, and these effects can extend over time. Therefore, the effects of a prosodic boundary and of prominence on the constriction gestures would be similar, and there would be a decrease of gestural overlap under prominence. However, the mechanism underlying the π-gesture and the μ-gesture is different. Indeed, the π-gesture is superimposed to the constriction gestures and is incorporated when the gestural score is already specified. On the other hand, the μ-gesture is implemented with the constriction gesture, so it modulates the gestural score at its creation. Another difference is that for the modulation gesture, a temporal modulation gesture μ_T gesture and a spatial modulation μ_S gesture are proposed. The μ_T and μ_S gestures would act respectively on the temporal and the spatial parameters of the coactive constriction gestures (Saltzman, Nam, Krivokapić, & Goldstein, 2008). If the μ_T gesture effects would be similar to the ones of the π-gesture, the introduction of a μ_S gesture would account for effects of prominence that are different from the effects of prosodic boundaries.

3.2.2 Coarticulation in prominent syllables

Several studies have investigated the effect of prominence on coarticulation, showing a general tendency toward a reduction of coarticulation under prominence. This has been reported for both coarticulation between two vowels or between a consonant and a vowel.

The effect of prominence on V-to-V anticipatory and carryover coarticulation has been analyzed in different languages. Overall, stressed vowels have been found to be more resistant to coarticulation than unstressed ones, regardless of the directions of the effects observed. A reduction of anticipatory and carryover coarticulation in stressed vowels has been showed in English (Fowler, 1981; Magen, 1984; Beddor *et al.*, 2002; Cho, 2004), Italian (Farnetani, 1990), Greek (Nicolaidis, 1999), Cantonese and Mandarin Chinese (Mok, 2013) and Catalan (Recasens, 2015).

However, some language-dependent, context-dependent and speaker-dependent variations are found. For instance, in her electropalatographic investigation, Nicolaidis (1999) showed that, in VCV sequences were V was /i/ or /a/ and C was /p, t, s/, there was less coarticulation in stressed vowels when the intervocalic consonant was /s/ or /p/, but no effect of stress was found when there was an intervocalic /t/. Moreover, the magnitude of the effect depended on speaker for the /p/ and /s/ context. Conklin and Dmtrieva (2020) found an effect of stress on carryover, but not anticipatory V-to-V coarticulation, in Spanish. Cross-languages differences in the effect of prominence has been reported, for example, by Beddor, Harnsberger, & Lindemann (2002), who showed an effect of prominence on coarticulation in English, but not in Shona, where stressed and unstressed vowels exhibited a similar degree of coarticulation.

The effect of prominence on consonant to vowel coarticulation has been investigated in a series of acoustic studies by Cho and colleagues (Cho, 2017; Joo, Jang, Kim, Cho, & Cutler, 2019; Jang, Kim, & Cho, 2018; Li, Kim, & Cho, 2020). Overall, these studies reported a decrease of vowel nasalization in CVN words (where N=nasal consonant) in accented vowels under focus in American English, Australian English, Korean and Mandarin Chinese, respectively.

3.2.3 Coarticulation at the edges of prosodic domains

Studies addressing intergestural coordination at the edges of prosodic domains, and namely in domain initial or domain final position, show either a tendency toward a decrease in overlap in domain initial position and an increase in domain final position, either no effect of position on overlap degree.

In domain initial position, changes in coordination have been reported by some studies for CV, VC or CC sequences, but the results are not consistent. A reduction of CV intrasyllabic carryover coarticulation in Intonational Phrase (IP) initial position has been found in American and Australian English, Korean and Mandarin Chinese in the aforementioned studies by Cho and colleagues (Cho, 2017; Joo, Jang, Kim, Cho, & Cutler, 2019; Jang, Kim, & Cho, 2018). They examined #NVC words (NV words for Mandarin), where # represents a word or IP boundary, reporting a decrease in vowel nasalization when the word was IP initial. In French, Guitard-Ivent, Turco, and Fougeron (2021) found different results depending on coarticulation type and analyzed sequence. The authors investigated, in word-medial vs IP initial position, anticipatory VC coarticulation in tautosyllabic VC and heterosyllabic V.C sequences, V-to-V coarticulation in V_1CV_2 sequences, and carryover coarticulation in CV sequences. Local coarticulation was analyzed as the effect of consonant place of articulation, alveolar or uvular, on F1 and F2 of the vowel, while V-to-V coarticulation was analyzed as the effect of V_2 eight on V_1. If coarticulation was reduced for #VC and #V.C sequences in IP initial position, with respect to word-medial position, no effect of position was found for either anticipatory V-to-V or carryover CV coarticulation. A similar result on carryover CV coarticulation in French has been reported by Meynadier (2002), who carried out an electropalatographic investigation of #CV sequences (C=/t, k, l/), where # represents a boundary varying from syllable, to Accentual Phrase (AP), to IP boundary, finding no change in coarticulation depending on boundary type. In American English, Byrd (2000) analyzed with EMA the intergestural timing in the sequence #mi, where # was either a word, a minor or a major prosodic boundary, showing speaker-dependent results. Indeed, only one speaker out of three increased intergestural timing, and thus reduced overlap, at the increasing of prosodic boundary strength.

Other studies have examined consonant overlap in CC clusters, reporting mixed results. Bombien, Mooshammer, Hoole, Künhert, and Schneeberg (2006) investigated with electropalatography /kl/ clusters in German in word initial position and in initial position of minor and major prosodic boundaries, reporting less overlap between the consonants after a prosodic boundary than in word initial position. On the other hand, Byrd and Choi (2010) examined CC clusters (/sp, sk, kl/) in American English in word, intermediate phrase (ip) , IP and Utterance initial position, reporting no difference in overlap degree depending on position.

Some of the studies mentioned here addressed also the effects of domain final position on coarticulation. In this position, changes in coarticulation, when attested, appear to go in the opposite direction as the ones found in domain initial position, and thus toward an increase in coarticulation. In his investigation, Meynadier (2004) reported an increase of VC anticipatory coarticulation in French at the increasing of prosodic boundary strength. Cho (2017), Jang, Kim, and Cho (2018) and Joo, Jang, Kim, Cho, and Cutler (2019) showed an increase in the anticipatory nasalization of the vowel in CVN# sequences followed by IP boundary with respect to word boundary, in American English, Korean and Australian English. On the other hand, the results of Li, Kim, and Cho (2019) reported no effects of IP final position on vowel nasalization in CVN# sequences in Mandarin.

3.2.4 Coarticulation across prosodic boundaries

Between sounds on the two sides of a prosodic boundary, coarticulation has been shown to overall reduce at the increasing of boundary strength. Particularly relevant for the present investigation is the finding that there is less anticipation of a following vowel in a preceding one when the two vowels are separated by a prosodic boundary. Cho (2004), in an articulatory study, reported a progressive decrease of V-to-V anticipatory coarticulation in $CV_1\#CV_2$ English sequences where V=/i, a/ at the increasing of the # boundary from word, to intermediate phrase (ip), to Intonational Phrase (IP) boundary. The domain final target vowel /a/ was less coarticulated with a postboundary /i/ across an ip than a word boundary and across an IP than an ip boundary. Carryover coarticulation also decreased across a higher boundary with respect to word boundary, with less influence of the preboundary /i/ on the postboundary /a/. The same pattern is shown for CV and CC intergestural timing, with an increase of the time-lag between gestures, and therefore a decrease of overlap. In the aforementioned articulatory study by Byrd (2000) she measured in the sequence mə#mi the time from the target achievement for the preboundary /m/ to the peak velocity for the postboundary /i/, and found that the time lag gradually increased at the increasing of the boundary strength from word, to minor to major prosodic boundary. An increasing of this time lag would entail a decreasing of overlap between the preboundary and the postboundary syllables. Meynadier (2004) found a gradual increase in the lag between the offset of the C_1 and the onset of the vowel in $C_1\#C_2V$

sequences at the increase of the boundary strength (word, AP, IP). For C#C sequences, a gradual decrease in the overlap between the two consonant gestures at the increase of boundary strength has been reported by Byrd and Choi (2010) for /sp, sk, kl/ clusters, and by Meynadier, for /tk, lk/, but not /kt, lk/ clusters, suggesting that the effect of boundary could vary for segments of different identity (and in different languages).

If these studies overall agree in indicating a decrease in coarticulation across a prosodic boundary, some questions remain open. Indeed, they have generally compared coarticulation across word boundary with coarticulation across boundaries of increasing strength, leaving open the question whether coarticulation between two syllables in a word is different from coarticulation across two syllables separated by a word boundary. Although this question is addressed in Hardcastle (1985) who investigated the production of /kl/ clusters across syllable boundary, word boundary and boundaries of increasing strength, he reported no consistent effects of a word boundary on the degree of overlap (this study, which reported a good amount of individual variation in the effect of boundary on overlap, is rediscussed in the next section). Moreover, one could wonder if the same results obtained by Cho (2004), who investigated the effect of a prosodic boundary on V-to-V coarticulation in English, could be obtained in French, a language where boundaries are usually marked by pitch accents, thus adding the effect of prominence. To these questions it can be added the matter whether the effect of a prosodic boundary on coarticulation are consistent across speakers. Indeed, as it will be reviewed in the next section, some studies showed interspeaker variations.

3.2.5 Interspeaker variations in coarticulation across prosodic boundaries

Some studies reported individual variations in the patterns of coarticulation across different boundary types, which could be related to possible differences in the prosodic renditions of the boundary conditions by the speakers. Tabain (2003) showed a speaker-specific effect on the degree of anticipatory V#C coarticulation in French across word, AP, IP and Utterance boundaries. Indeed, some speakers did not show a difference in coarticulation degree between boundaries higher than the word level, and some speakers did not show a difference between the word and the AP level. Interspeaker variation is reported also by Harcastle (1985) in his electropalatographic investigation on k#l sequences in English across syllable, word, clause and sentence boundary produced at normal and fast rate. While the overlap tended to decrease for all speakers when the two consonants were separated by a clause or sentence boundary, this pattern is not shown by all speakers. Moreover, at fast rate, some speakers produced a greater overlap between consonants across clauses than across a word or syllable boundaries, suggesting that for these speakers the constraints relative to the production of successive sounds in a shorter time have to some extent "overridden" eventual timing differences related to boundary type.

Individual variation in between words coarticulation has been explicitly addressed by the investigations of Abry & Lallouache (1995, 1996) that have led to the elaboration of the Modèle d'Expansion du Mouvement (Movement Expansion Model). The MEM has been elaborated to account for the anticipation of labial protrusion in French across an increasing number of consonants (Abry & Lallouache, 1995; Abry, Lallouache, & Cathiard, 1996) and tested also for English (Noiray, Cathiard, Ménard, & Abry, 2011). Abry & Lallouache (1995) recorded with a lip shape tracker the production of [iC_ny] sequences by four French speakers, where C_n corresponded to a number of consonants from one to five. The sentence with no intervening consonants between the vowels was "*Ces deux scies ultèrent*" ([iy] sequence), while the sentence with the maximum number of consonants was "*Ces deux Sixtes sculptèrent*" ([ikstsky] sequence). They looked at the evolution of the rounding movement during the consonant interval, showing that the movement time linearly increased with increasing consonant number. The rounding movement began during the production of /i/ in the zeroconsonant condition, and progressively later with increasing number of consonants. Crucially, they showed that the increase in movement time increased following a speaker-specific slope. Speakers' specific behavior could fall between the time-locked and the look-ahead behavior. In other words, the reduction of coarticulation at the increasing of distance between the vowels was speaker-specific. If, in this case, the increasing of the distance between vowels was to attribute to the increase in the number of consonants, and there was no change in boundary type, which was kept a word boundary, these results show that speakers could differently manage coarticulation between two words.

3.3 Population-specific effects on coarticulation

3.3.1 Age-varying population

Changes in coarticulation with age have been studied primarily in relation to childhood development. While there is no agreement in the literature as to whether coarticulation increases or rather decreases with age in children, coarticulation patterns have been reported to progressively approach adult patterns taken as the reference. This "evolution" in coarticulation has been seen in relation to an evolution in the size of the units of planning in speech with development. However, speech continues to change during adulthood. In particular, one well known change in speech with age is a slowing of rate, especially attested for older adults. A variation of coarticulation in adulthood with age could relate to this slowing of speech.

The next sections will review the main findings on coarticulation in children and how they have been interpreted in relation to the units of planning in speech. Then, the attested changes speech undergoes with aging that could impact coarticulation will be presented.

3.3.1.1 Coarticulation in childhood development

Anticipatory coarticulation patterns has been shown to change during childhood. This has been reported for both intrasyllabic and extrasyllabic effects, although the direction of these changes is not clear. Studies on V-to-V coarticulation in children have reported either a greater or a lesser degree of coarticulation in children than adults. Greater coarticulation in children than adults, and a progressive decrease of coarticulation with children increasing age, is found by Rubertus and Noiray (2018) and Noiray, Abakarova, Rubertus, Krüger, and Tiede (2018) who examined anticipatory V-to-V coarticulation, and VC coarticulation, in four cohorts of German children aged 3 to 7 years old, compared to young adults. In particular, if children from 3 to 5 presented similar patterns of coarticulation, the 7 years old group showed a decrease of coarticulation in the direction of adult-like patterns. Similar results, with more V-to-V coarticulation in children than adults, are reported by Nijland *et al.*, (2002) for Dutch children ages 4 to 6, and by Boucher (2007) for American English children aged 3 to 5. Conversely, Barbier (2016; Barbier *et al.*, 2020) found less coarticulation in French children aged 4 to 10 when compared to young adults, and more variability of coarticulation in children. Moreover, they also analyzed sequence duration, showing longer durations for children, which the authors see as potentially related to the lesser coarticulation showed in children. A progressive increase of coarticulation in children with increasing age was also previously reported in American English by Repp (1986) who compared two children aged 5 and 9, and Hodge (1989) who compared 3, 5 and 9 years old children.

Similar inconsistencies are found in studies on CV anticipatory coarticulation. Zharkova, Hewlett, and Hardcastle (2012), and Zharkova (2017) found less intrasyllabic coarticulation in children aged 6 to 9 than adults, and in children aged 5 compared to 13 years old adolescents, in Scottish English. Conversely, Nittrouer, Studdert-Kennedy, and Neely (1996) reported greater coarticulation in English children aged 3 to 6 than adults, and the aforementioned study by Noiray, Abakarova, Rubertus, Krüger, and Tiede (2018) reported similar results for German children aged 3 to 7.

The interest in examining changes in anticipatory coarticulation in childhood development lies in what these changes can tell us about the planning units in speech. Indeed, depending on the direction of these changes toward more or less coarticulation in adults, different hypotheses have been made regarding the maturation of the speech production system (see review in Noiray *et al.,* 2018). The finding of a greater intra and extrasyllabic coarticulation in children and its consequent decrease with age (e.g. Nijland *et al.,* 2002; Rubertus & Noiray, 2018), has served as evidence to affirm that children organize their speech over larger units than adults, such as words, and are lexically driven. With age, they would acquire more articulatory precision, leading to different patterns of coordination and to greater independence of successive speech elements. The opposite perspective is suggested by the finding of a lesser intra and extrasyllabic coarticulation in children and its consequent increase with age (e.g. Barbier *et al.,* 2020; Zharkova *et al.,* 2012). Based on these results, the organization of speech in children would be segmentally-driven. With age, children would increase in gestural cohesion within and between units of speech. A third hypothesis (Noiray *et al.,* 2019) start from the observation of consonant-dependent patterns of V-to-V coarticulation in children to suggest that both a more holistic and a more segmental organization of speech can be found in children depending on the gestural demands of the segments to produce. Rather than a change in unit size, the variation in coarticulatory patterns would attest a tuning of gestural coordination with age, until adult-like patterns are reached.

These hypotheses bring also different assumption regarding the units of organization of speech planning and production in adults. Indeed, despite the controversial results and the ensuing dispute, the results on childhood development show how coarticulation is not stable during the lifespan, and coordination within and between syllables evolves throughout childhood, which coarticulatory patterns progressively approaching adult ones with increasing age. In that respect, the adult coarticulation pattern to which children's productions are compared is intended as 'the' reference for a mature speech production system where coarticulation is assumed to be stable. That said, this reference is very often that of quite young adults (for example, under 30 y.o.a. in Noiray *et al.*, 2018; Barbier *et al.*, 2020; around 30-40 y.o.a. in Zharkova *et al.*, 2012), as is the case of many other studies on which our knowledge of speech production is based. However, speech patterns in adulthood do not stay the same: variations in the spatial and temporal organization of speech are observed. In the next paragraph these variations are addressed, and especially the well attested decrease in speech rate with age, which could be related to a change in coarticulation patterns.

3.3.1.2 Changes in speech with age in adulthood

There are several reasons for speech to evolve during adulthood. These can be linked to physiological or cognitive changes accompanying natural aging, but also to many other changes conditioned by speech usage and life experience. Several age-related speech changes have been documented in the literature (see, *inter alia*, Fougeron, Guitard-Ivent, & Delvaux, 2021). Here we will focus in particular on the changes reported on aspects linked to the temporal organization of speech.

Many cross-sectional studies that have explored the effects of aging have shown a deceleration of speech rate. At the sentence level, a slowing of speech in older speakers has been reported in both spontaneous and read speech as a decrease of articulation rate or an increase in sentence duration. For instance, an increase in sentence duration in spontaneous speech with age has been showed by Horton, Spieler, and Shriberg (2010) for speakers aged 20 to 67, while a decrease in articulation rate in both spontaneous and read speech has been reported by Ramig (1983) for speakers aged 65 to 75 with respect to younger groups. Slower articulation rate in read speech has been reported also by Jacewicz, Fox, O'Neill, and Salmons (2009) for speakers aged 51 to 65, and by Bourbon and Hermes (2020) for speakers aged 68 to 88, with respect to younger groups. At the segmental level, longer vowels in older speakers have been reported by Albuquerque, Oliverira, Teixeira, Sa-Couto, and Figueiredo (2019), who showed, in groups of speakers aged 35-49, 50-64, 65-79 and >80, a progressive increase in vowel duration with increasing age of the groups. Longer consonant duration has been instead reported for example by Morris and Brown, for speakers over 75 of age. The results reported by Mücke, Thies, Mertens, and Hermes (2020) showed that this difference in segmental duration between older and younger speakers is accompanied by differences in the kinematic profiles of tongue movements. Indeed, they reported slower and more asymmetrical tongue movements, with prolonged deceleration phases, in four speakers aged 70 to 79, compared to four speakers aged 20 to 29.

This slow speech in older speakers has been related to the overall slowing of body movements with age. Finger movements and handwriting have also been shown to be slower in older than in younger adults (respectively, Caçola, Roberson, & Gabbard, 2013; Bilodeau-Mercure *et al.,* 2015; Rosenblum, Engel-Yeger, & Fogel, 2013). For instace, Rosenblum *et al.,* (2013) studied handwriting performances in four groups of speakers aged 31-45, 4-60, 61-75, and >76, reporting an increase of on-paper and in-air time for older participants. That is, older participants showed slower handwriting movements and longer times between strokes. Hirai, Tanaka, Koshino, & Yajima (1991) investigated orofacial movements non-related to speech production, showing an effect of age. They used ultrasound to examine the execution of tongue movements timed with a metronome set at different frequencies by younger (mean age 27) and older (mean age 66) participants, reporting slower movement durations at each metronome frequency for the older than for the younger group.

The majority of the literature focus on comparisons between age groups of varied age, so it is difficult to individuate a timeline of these changes. Notwithstanding, some investigations seem to indicate that they are not specific to old age. For example, the results of Bilodeau-Mercure *et al.* (2015) support the idea that the slowing down of finger movements occurs quite early (already for the 37–54 years old group). For speech, changes occurring early in adulthood have also been reported. Jacewicz, Fox, and Wei (2010) found an increase in speech rate until the late 40s, and then a decrease for older speakers. Conversely, Fougeron, Guitard-Ivent, and Delvaux (2021) documented, on a large sample of speakers, a continuous decrease in speech rate from 20 to 93 y.o.a., with a sharper slowing down after the mid-50s.

One aspect that has to be considered, regarding a slowing of speech and the production of longer vowels in older speakers, is that these longer vowels would leave the speaker more time to reach the articulatory targets. Therefore, a hyperarticulation of vowel target could be expected with age. Some results in the literature point in this direction. Fletcher, McAuliffe, Lansford, and Liss (2015), in a cross-sectional analysis on speakers aged 65 to 90, reported more peripheral vowel targets for speakers who exhibited longer vowel duration. In their longitudinal study, Gahl and Bayeen (2019) showed that from 20 to 50 years of age, vowels for the same speakers tend to get more peripheral, leading to an expansion of the vowel acoustic space. However, this was found for both long and short vowels, so it does not seem to depend of the time speaker have to reach the articulatory target. This change from more reduced to more peripheral realizations of vowel targets, regardless of vowel duration, suggests that the kinematic organization of speech may also change according to speakers' age. Middle-aged speakers could present a bias toward hyperarticulation, that would lead them to increase movement displacement and potentially increase velocity when peripheral targets need to be reached in a reduced time, as is the case for short vowels.

Other studies looking for age-related changes in vowel articulation in much older speakers have shown quite inconsistent results. Analyses of formant frequency shifts in groups of speakers of different age have reported vowel-dependent and sex-dependent results, that do not allow to paint a clear picture (e.g. Xue & Hao, 2013; Rastatter & Jacques, 1990; Torre & Barlow, 2009). For example, Eichoorn, Kent, Austin, and Vorperian (2018) compared vowel formants in female and male English speakers across three age groups, 20-30 years of age, 40-60 y.o.a. and 70-92 y.o.a., showing higher F1 of /æ/ and /u/ and higher F2 of /u/ in middle aged and older women compared to young women, but only higher F2 of /u/ for older men compared to young men.

On the other hand, some studies suggest rather a tendency toward a reduction of vowel targets at old age, even in presence of longer segmental durations (e.g. Liss, Weismer, & Rosenbek, 1990; Mücke *et al.*, 2020). Albuquerque *et al.* (2019, 2020) and Oliveira *et al.* (2021) analyzed a large pool of Portuguese speakers reporting a tendency toward vowel centralization for old male speakers, and a lowering of F1 and F2 for old female speakers. In a follow-up study, Albuquerque *et al.* (2021) analyzed ultrasound data for two young and two old female speakers, showing that the lowering of the first two formants previously reported was compatible with a reduction of vowel space.

A change in kinematics with age, with either more hyperarticulated or more hypoarticulated vowel targets, could be another factor, besides rate or in connection with rate, that can affect V-to-V coarticulation patterns in middle-aged or older speakers. The relationship between changes in kinematics with age and changes in rate is still unclear, while the effects of these changes on coarticulation are yet to be explored.

3.3.2 Motor Speech Disorders

Motor Speech Disorders (from now on MSD), which include Dysarthrias and Apraxia of Speech (AoS), result from neurological impairment of planning, programming and execution of speech. In other words, MSDs affect different stages of speech encoding that come after the linguistic and phonological encoding, which is affected in aphasia. Dysarthrias are associated with impairment at the level of motor programming or execution, while Apraxia of Speech is associated with impairment at the level of planning. An impairment in coarticulation in dysarthria could be attributed to a smaller range of movements and slower or imprecise movements, while an impairment in anticipatory coarticulation in AoS could be attributed to a deficit in planning. However, if the literature on AoS is clearer in indicating an impairment in anticipatory coarticulation, studies on dysarthria present inconsistent results and often analyze speakers affected by Parkinson's Disease. These speakers, affected by hypokinetic dysarthria, do not present some characteristics shared by other dysarthria types and by AoS, such as slowing of speech, which might be related to coarticulation. In the next paragraphs, dysarthria and Apraxia of Speech will be introduced and the principal results on coarticulation will be summarized.

3.3.2.1 Dysarthrias

Dysarthria identifies with a group of neurologic speech disorders that result from disturbances in the control of the speech musculature due to damages to the central or peripheral nervous system (Darley, Aronson, & Brown, 1969a). It can be categorized into different types, spastic, flaccid, mixed spastic-flaccid, ataxic, hypokinetic, hyperkinetic in chorea, and hyperkinetic in dystonia. Each is characterized by distinguishable auditory perceptual characteristics and different underlying pathologies, following the neurophysiological classification of Darley *et al.,* (1969a, 1969b, 1975), who described each type of dysarthria with specific clusters of perceptual features. Speakers affected by dysarthrias can show modified range, direction, coordination, force, speed, tonus of speech movements and impaired movement patterns (Duffy 2012). Dysarthria globally affects the whole speech production system. It involves impairments of the respiratory, laryngeal and articulatory systems of speech production, with the nature of impairment varying depending the on severity and dysarthria type (Kent, Kent, Duffy, & Weismer, 1998).

The impairment in Dysarthria is considered to be situated either at the level of speech execution, or at the level of motor programming or execution depending on dysarthria type. For motor programming, I will refer to the definition of van der Merve (1997). Van der Merwe (1997, 2021) suggests, as levels of speech encoding, a stage of linguistic-symbolic planning, i.e. phonological planning, which is impaired in aphasia, a stage of motor planning, which is impaired in apraxia (see below), a stage of motor programming and a stage of execution, which are be impaired in dysarthrias. In the motor programming stage, motor programs for the muscles and articulator of the speech production system are selected and sequenced. Moreover, these muscle-specific programs are specified in terms of “muscle tone, rate, direction and range of movements”. During the execution phase, these programs are transformed in automatic motor adjustments.

In the study that will be presented in Chapter 5, it will be specifically addressed mixed spastic/flaccid dysarthria associated with Amyotrophic Lateral Sclerosis (ALS), mixed hypo/hyperkinetic dysarthria associated with Wilson Disease, and hypokinetic dysarthria associated with Parkinson’s Disease. Therefore, the speech characteristics of these dysarthria types, as described in Duffy (2O12) and Enderby (2013), will be briefly introduced.

a) **Flaccid dysarthria**. Flaccid dysarthria can be considered a disorder of speech execution. It can affect a single or more muscle groups, so there are different subtypes associated with distinct speech abnormalities. However, the principal speech characteristic is muscle weakness and reduced muscle tone, which affects speed, range and accuracy of movements. Speakers can present abnormality in lip movements, reduced movements of the tongue, reduced phonation time and reduced intelligibility.
b) **Spastic dysarthria**. Spastic dysarthria can also be considered a disorder of neuromuscular execution. The principal speech characteristics are muscle weakness and muscle hypertonicity, that is, excessive muscle tone. Speakers can present spasticity, slow movements reduced in range and force.
c) **Hypokinetic dysarthria**. Hypokinetic dysarthria can be considered a disorder of motor control or motor programming. It affects the preparation, switching and maintaining of the motor programs. Speech characteristics are difficulties

in initiating movements, movements reduced in amplitude, range and rigidity. Individual speech movements are slow, but repetitive movements can be fast.

d) **Hyperkinetic dysarthria**. Hyperkinetic dysarthria can be considered a disorder of motor control or motor programming. It causes abnormal, rapid or slow, regular or unpredictable involuntary movements that disturb the rhythm and rate of speech.

3.3.2.2 Apraxia of Speech

Apraxia of Speech (AoS) is an impairment of volitional speech production in absence of motor or execution deficits (Darley 1975). Commonly reported symptoms of AoS are phonemic errors such as perceived phoneme omissions, substitutions, additions and exchanges, phonetic distortions, prosodic abnormalities, difficulty initiating speech, groping, and effortfullness. Lengthening of segments, silent intervals between syllables and difficulty in transitioning between segments create the impression of syllable segregation, resulting in a “syllabified” or “segmented” speech (Mc Neil, Robin, & Schmidt, 1997; Kent & Rosenbek, 1983; Ziegler, 2008; Duffy, 2012). One of the characteristics of apraxia is variability, which leads to inconsistency of errors across repetitions: an erroneous production of a word can be followed by a production of the same word without any errors, which could lead to misleading evidence in small population samples (Ziegler, Aicher, & Staiger, 2012).

Apraxia of Speech is to be distinguished from aphasia on one side, and from dysarthria on the other. Indeed, AoS impairment is considered to be situated at a stage after phonological encoding, and at a stage prior to motor programming and execution. In other words, a person affected by Apraxia of Speech completes the linguistic processing, including the semantic and grammatical formulation of the message and the retrieval of the abstract phonological representation, but they are not able to translate it into the motor commands that guide the articulators (e.g. Croot, 2002; Ziegler, 2008). If the literature agrees that AoS affects a stage of speech encoding roughly in between a purely phonological and a purely executive stage, several proposals have been made over the years on how to better describe or model this stage.

Mc Neil *et al.* (1997) define it as a phonetic-motoric disorder affecting the transformation of the intact phonological representation of an utterance into the learned kinematic parameters for the intended movements. In Van der Merwe (1997), this concept is specified by situating AoS impairment at the stage, in speech encoding, of motor planning, where abstract phonemes are transformed into specified or contextualized motor goals. From a cognitive perspective, Whiteside and Varley (1998a) identify the process that is disrupted in AoS as retrieval of syllable- or word-sized verbo-motor plans, at the level of phonetic encoding. The authors move from Level's idea that motor plans for high-frequency syllables are stored in a mental syllabary (Levelt & Wheeldon, 1994; Levelt, Roelofs, & Meyer, 1999). Apraxic speakers would be unable to access these patterns, so that, at every production, the motor plans for syllables and words have to be indirectly assembled from phoneme-sized units. More recently, a similar view comes from the contextualization of AoS in the DIVA/GODIVA model. In this framework, it is considered as an impairment in the temporal buffering of the phonological plans for an utterance and in the choice and execution of the correct motor plans. Specifically, AoS would damage the speech sound map, impacting the generation of motor commands for the syllables to produce (Miller & Gunther 2021).

Different accounts have stem from a Task Dynamics perspective. Kelso and Tuller (1981) affirmed that apraxia can be characterized as a disorder in which the "meaning of events" is disrupted. Meaning that apraxic patients would be unable to contextualize speech movements and tune them into functional coalitions according to a specific speech task. Elaborating on this perspective, Ziegler (2009, 2020) proposed that in AoS the cohesiveness of intra and intergestural coordination is broken, and these speakers have to assemble at all times the individual gestural components of each speech goal.

Overall, some common ideas can be individuated in the literature, despite the differences in terminology and in the specifics. AoS can be considered an impairment at an encoding stage where the assemblage and contextualization of speech elements in cohesive speech units, and therefore also their coordination, happens. This encoding stage can be considered as phonetic planning or motor planning, as opposed to a stage of motor programming or execution of these motor programs that is impaired in Dysarthria.

AoS and Dysarthria, although can be traced to different pathomechanisms, present some common characteristics of altered speech, with the degree of overlap between the two disorders varying according to severity and type of dysarthria. Besides phonetic distortions and abnormal prosody, of particular importance for the present investigation is that both AoS and Dysarthria entail longer segmental duration and a slowing of speech, at the exception of hypokinetic dysarthria, which can lead to an increase in speech rate.

3.3.2.3 Coarticulation patterns in Motor Speech Disorders

3.3.2.3.1 Coarticulation in Dysarthrias

Studies on V-to-V coarticulation in Dysarthria are scarce and reported either unimpaired coarticulation either inconsistent patterns for dysarthric speakers compared to control speakers. Dogil and Meyer (1998) investigated V-to-V coarticulation in a German word, in one speaker affected by flaccid dysarthria and two speakers affected by AoS, compared to a control speaker, showing unaffected coarticulation for the dysarthric speaker. If the production of only one speaker was investigated there, similar results are reported for a cohort of twelve speakers affected by Parkinson's Disease by Tjaden (2003). He examined the anticipation of V_2 in V_1 in sequences composed by the English article /ə/ followed by different monosyllabic words, finding no difference in coarticulation degree between Parkinson's Disease and control speakers. Different results depending on coarticulatory direction are found in nine speakers affected by Cerebellar Ataxia by Hertrich and Ackermann (1999). They investigated anticipatory and carryover V-to-V coarticulation in threesyllabic /CəCVCə/ German pseudowords, showing unaffected carryover coarticulation, but inconsistent patterns of anticipatory coarticulation due to the high degree of trial-to-trial variability shown by the dysarthric speakers.

In the same study, Hertrich and Ackermann (1999) analyzed also anticipatory CV and carryover VC coarticulation, showing a reduction of anticipatory coarticulation, but unaffected carryover coarticulation for ataxic speakers. The dysarthric speakers also showed longer durations with respect to controls, suggesting that some of the reduction in anticipatory CV coarticulation could be accounted for by this slowing of speech, while carryover coarticulation would not be affected. As for anticipatory V-to-V coarticulation, the variable pattern found for ataxic speakers does not allow to draw conclusions in this respect.

Other studies on local coarticulation target mainly Parkinson's Disease speakers, presenting mixed results. Iraci (2017) investigated the effect of geminate consonants on the duration of the preceding vowel in Italian in five speakers affected by Parkinson's Disease, showing no expected compensatory shortening for these speakers, which can be interpreted as a lack of coarticulation. A reduction of coarticulation in Parkinson's Disease is reported also by Martel-Savageau and Tjaden (2017), while Tjaden (2000) presented opposite results. Indeed, Martel-Savageau & Tjaden (2017) found less coarticulation in nine Canadian French speakers affected by Parkinson's Disease than in control speakers. They also examined articulation rate, without finding a difference in articulation rate between the two groups. On the other hand, slightly more coarticulation in nine English speakers affected by Parkinson's Disease than in control speakers is shown in Tjaden (2000). Despite shorter segmental durations were found for the Parkinson's Disease group, a further comparison with the control speakers showed that the fast rate of the dysarthric speakers could not explain the increase in coarticulation. In the direction of an increased coarticulation in

Parkinson's Disease can be interpreted also the results of Roland *et al.* (2016). They showed that, in the production of VCV glides in Belgian French (e.g. /aja/) by speakers affected by Parkinson's Disease, the two vowel targets were closer to the target of the intervocalic C than control speakers, probably as a strategy to limit displacement.

Taken together, these studies do not paint a clear picture of coarticulation patterns in dysarthria. The comparison is complicated by the different languages investigated and the fact that the studies focus on the comparison between a control group and a group of speakers affected by one type of dysarthria, which is often hypokinetic dysarthria associated with Parkinson's Disease.

3.3.2.3.2 Coarticulation in Apraxia of Speech

Studies on coarticulation in Apraxia of Speech focus mostly on anticipatory CV and V-to-V coarticulation, reporting an overall reduction of coarticulation with respect to control speakers. Among the first to show this reduction, Ziegler & Von Cramon (1985, 1986a) investigated anticipatory V-to-V and CV coarticulation in one speaker affected by AoS compared to a group of control speakers. The authors analyzed, in /CəCVCə/ german pseudowords, the anticipation of V2 in V1 /ə/ and in the consonant burst, showing a lack of anticipation for the AoS speakers, together with long silent intervals before the consonant burst. In a follow-up study, the results on CV coarticulation were replicated for 7 apraxic speakers, who showed less coarticulation than the control group (Ziegler & Von Cramon, 1986b). A reduction of anticipatory V-to-V coarticulation in AoS in German is also shown by Dogil and Meyer (1998) in the production of two apraxic speakers. Similar results are reported for English by Tuller and Story (1988) and Whiteside and Varley (1998b). Tuller and Story (1988) showed, in sequences composed of two English words, reduced anticipatory CV and carryover VC coarticulation in three apraxic speakers; while Whiteside and Varley (1998b) reported, in CV or CVC pseudowords preceded by the article /ə/ a reduction of V-to-V and CV anticipatory coarticulation in one apraxic speaker compared to a control speaker. A reduction of CV anticipatory coarticulation in one English speaker affected by Apraxia is also shown by Southwood, Dagenais, Sutphin, and Garcia (1997) who analyzed acoustic, perceptual and EPG data of productions of CVC words at slow, normal and fast speech. With respect to the control speaker, the apraxic speaker showed delayed coarticulation, with the effect being stronger at fast rate. A further decrease of coarticulation at fast rate could suggest more difficult in planning the next segment when the speaker has less time to do it.

Other studies have reported speaker-dependent or item-dependent patterns of coarticulation in AoS. Speaker-dependent results are reported by Katz, Machetanz, and Schölne (1990), who investigated CV and V-to-V anticipatory coarticulation in German words in two apraxic and two control speakers. Of the two apraxic speakers, one showed anticipatory coarticulation, while the other showed no anticipation in either the preceding consonant or vowel. Bartle-Meyer and Murdoch (2010) analyzed with EMA and EPG the anticipation of C_2 in the tongue tip movement at the release of C_1 in the second syllable of the English words *sergeant* and *scarlet,* in three AoS speakers compared to controls. Two of the apraxic speakers showed less anticipation in the word *sergeant*, by showing greater tongue tip displacement, while no difference with control speakers was shown in the production of *scarlet.* The authors attribute this result to the presence of the cluster: the production of the cluster could be more difficult to plan for apraxic speakers, who would therefore show a reduction of anticipation. Inconsistent patterns were found for the third speaker.

The results on AoS seem more consistent than the ones on dysarthric speakers, indicating almost systematically a reduction of anticipatory coarticulation in these speakers. However, these studies often focus on the examination of little cohorts of apraxic speakers, composed by one to three speakers, and target English or German. Finally, with the exception of the study by Dogil and Meyer (1998) presented in both sections, who show unimpaired V-to-V coarticulation in dysarthria, but impaired coarticulation in apraxia, the two Motor Speech Disorders does not seem to be often compared.

4 General methods

This chapter presents some general elements of the methodology, and in particular the methodology shared by more studies. It will be presented: 1) the target items on which coarticulation is investigated; 2) the corpus used in the **MSD** (chapter 5) and **age** (chapter 6) studies; 3) the procedures used for the annotation of the target items and the formant extractions; 3) the measures of V-to-V coarticulation; 4) the measures of articulation rate; 5) some generalities on Linear Mixed Models, which were used to analyze the data in **MSD, tempo** and **boundary** studies (respectively, chapter 5, chapter 7, and chapter 8).

Other methodology relative to the single studies, including the participants' pool for each study, will be presented in the method section of the dedicated chapters.

4.1 Target items

4.1.1 V-to-V coarticulation

V_2-to-V_1 coarticulation was analyzed as the influence of **V_2 /i/** on **V_1 /a/** in **/papi/** sequences while **V_1 /a/** in **/papa/** sequences was used as control context.

These two sequences were chosen for the following reasons:

- The effect of V_2 height on a preceding low vowel has been investigated in several studies on lingual coarticulation. Low vowels tend to present a weak resistance to coarticulation; in these sequences, lingual coarticulation of /a/ with /i/ is further facilitated by the nature of the intervocalic consonant /p/, which does not require tongue dorsum activation (e.g. Recasens, Pallarès, & Fontdevila, 1997).

- It allowed having meaningful words in the three studies in which within word coarticulation is investigated, where these target sequences correspond to meaningful French words *papa,* "dad" and *papi* "grandpa".

A clarification on V-to-V coarticulation and vowel harmony in French:

> Anticipatory V-to-V coarticulation affecting vowel height in French has been phonologized in a process of vowel harmony (albeit optional). However, this process affects mainly the mid vowels: the mid vowels in word internal position tend to be influenced by the height of the following vowel (e.g. the first vowel will be tendentially realized as mid-open in *aimait* [ɛmɛ], but as mid-close in *aimer* [eme] (Turco, Fougeron, & Audibert, 2016). The coarticulatory process that is investigated here does not fall in the category of vowel harmony.

4.1.2 C-to-V coarticulation (MSD study)

C-to-V coarticulation was analyzed as the influence of C **/ʃ/** on V_2 **/a/** in the monosyllabic word **/ʃa/**, *chat*, while V_2 /a/ in **/papa/** was used as control context. These words were chosen because they occurred the same number of times in the text; moreover, the effect of /ʃ/ on /a/ is somewhat comparable to the effect of /i/, therefore, the same measures has been used (see paragraph 5.2.3.1.2).

4.1.3 Panoramic view of the speech material for all studies

The sequences were embedded in sentences that were read by the speakers in a reading task. The corpus varied on the study. Since the speech material used for the **MSD study** and the **age study** was the same, it will be described in the next section, while the speech material for the **tempo** and **boundary** studies are described in the respective chapters.

Here a schematic description of the speech material for each study. At the end of each description, the sections were the detailed descriptions can be found.

MSD study (chapter 5). Target items were the French words ***papi*** "grandpa", ***papa*** "dad", and, ***chat*** "cat" embedded in eight sentences forming a short story. There were **6** occurrences of each target item in the text (described in section 4.3).

Age study (chapter 6). The target item was the French word ***papi*** embedded in eight sentences forming a short story (the same as the MSD study). There were **6** occurrences of the target item in the text.

Tempo study (chapter 7). Target items were the French words ***papi*** *and* ***papa*** embedded in two different sentences. Participant read the sentences at a comfortable self-paced tempo, at a slow tempo and a fast tempo. Each sentence was repeated 20 times in each condition, for a total of 20 occurrences of each item.

Boundary study (chapter 8). Target items were the bisyllabic sequences /papi/ and /papa/, embedded in meaningful sentences. The boundary between the two syllables was manipulated, from syllable boundary within a word (**within word condition**) to word boundary (**word boundary condition**) to clause boundary (**clause boundary condition**). 45 tokens of each target item were collected in each condition. In the **within word** condition, the target sequences were the words ***papi*** and ***papa***. In the **word boundary** condition, the target sequences were ***pa#pi***, in the phrase *papa pilote…* "dad pilots…", and ***pa#pa***, in the phrase *papa passe…* "dad passes…". In the clause boundary condition, the target sequences were ***pa##pi***, in the phrase *papa, pilote à…* "dad, pilot in…", and ***pa##pa***, in the phrase *papa, passant par…* "dad, passing by…". The target words and phrases were embedded in longer sentences.

4.2 Convention for the notation of target vowels

Throughout this dissertation, the following convention will be used to indicate the target vowels:

- **[a]i** for V_1 /a/ of *papi*
- **/i/** for V_2 of *papi*

- **[a]a** for V_1 /a/ of *papa*
- **ʃ[a]** for V_2 /a/ of *chat*
- **p[a]** for V_2 /a/ of *papa*

4.3 Speech Material for the MSD and the age study

The target words ***papi***, ***papa***, and ***chat*** were embedded in eight meaningful sentences forming a short story of about 188 words, which is part of the MonPaGe protocol (Laganaro *et al.,* 2021, Pernon *et al.,* 2020). It is a semi-automatized screening protocol specifically designed for the acoustic and perceptual assessment of the characteristics of the voice and the speech of patients who present signs of motor speech disorders. There were 6 occurrences of each target word in the text, given below in Table 1. English translation of the text is given in Appendix A.

Lundi, le chat, le loup et Papa vont à Bali. Les copains sont tout contents.
Mardi, Papi y va aussi. Il dit : "Je n'ai pas un sou ! Qui va prendre soin de moi ?" "Moi !" dit le chat, "moi !" dit le loup. "Vous ?", Papi réfléchit.
Mercredi, Papi dit : "Toi, le chat, tu es doux tu es chou, tu n'as pas de poux! Mais pas ce loup : il a une cape rouge et je n'aime pas ce gars-là!".
Jeudi, le chat et Papi se baladent à Bali. Papa glisse ! Aïe ! Ouille ! Son cou craque, son coulde claque, c'est la débâcle !
Vendredi, Papa a mal. Il pleure, il crie ! "Toi, Papi, aide-moi, trouve le nain !" "Un nain ? On n'en a jamais vu par ici ?!".
Samedi matin, le chat va voir son ami le loup et lui dit : "Aide-moi à soigner Papa!".
Samedi soir, le loup lui donne sa recette magique : "Coupe un oignon, cache-le sous la souche, et lorsque le lilas fleurira, Papa sera guéri!"

Dimanche, le chat tout doux, le loup magicien, Papa et Papi quittent Bali. Les copains sont tout contents.

Table 1. Text of the reading task used in the MSD and the age studies (MonPaGe protocol, Pernon et al., 2020). The target words papi, papa, and chat, are evidenced in bold.

4.4 Annotation of target items and formant extractions

4.4.1 Annotation of target items

MSD, Age, boundary studies. Targets words *papi*, *papa* and *chat* were manually segmented in phonemes using PRAAT (Boersma & Weenink, 2021).

Rate study. Recordings were semi-automatically annotated in sentences, words, and phonemes using EasyAlign (Goldman, 2011) for a previous study of the MoSpeeDi project (Didirkova, Lancia, & Fougeron 2020). The annotation of the target vowels of *papa* and *papi* were manually corrected by the author.

P*apa* and *papi* items: V_1, the intervocalic /p/, and V_2 were segmented. Since the duration of the first /p/ consonant was impossible to determine due to possible preceding pauses and a silent closure, the first /p/ was not taken.

4.4.2 Segmentation criteria

- For all items, vowel onset was defined by the beginning of voicing or of the second formant (whichever appeared first), while vowel offset was defined by the end of the second formant.

- In *papi*, the /i/ vowel was often partially devoiced at the beginning. Since formants could not be obtained for this part of the vowel, all annotation of /i/ included only the voiced part (Figure 3).

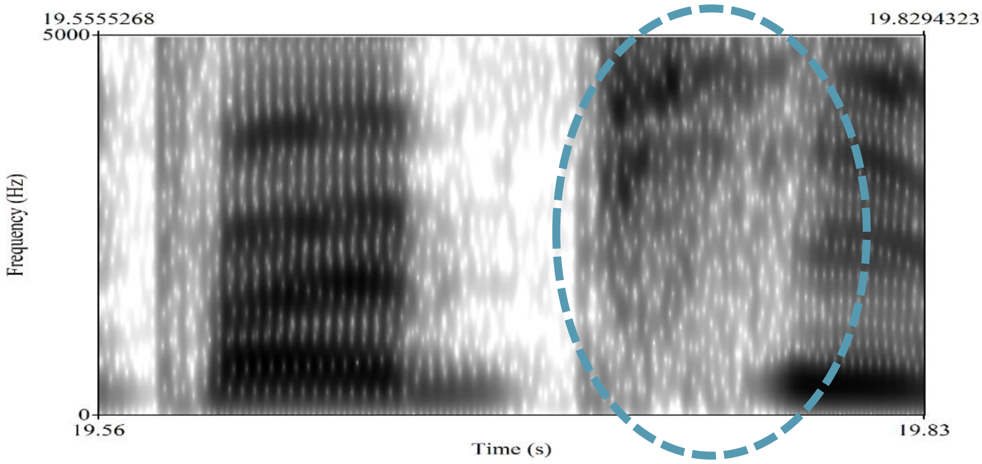

Figure 3. Example of a papi token, produced by a speaker partecipating in the tempo and boundary studies. Circled in blue, friction after the /p/ burst corresponding to the devoicing of /i/.

- For the *papi* and *papa* items, the duration of the intervocalic /p/ was defined as the interval between the two annotated vowels.

- For the chat items, the beginning of the /ʃ/ was defined as the beginning of the friction on the spectrogram, while the end coincided with the beginning on the vowel.

4.4.3 Formant extraction

The first two formants of the target vowels were automatically extracted thanks to a Praat script (Audibert, 2014, personal communication). The Burg algorithm of Praat (Boersma and Weenink, 2021) was used with the following settings: detection of 5 formants between 0 and 5kHz for males and between 0 and 5.5kHz for females on a 0.025 window length. Formant values were extracted every 10% of vowel duration. However, for the analyses, formant values taken in selected timepoints were used.

F1 and F2 were taken at:

- **50-60-70%** of **[a]i** and **[a]a** duration (*papa* and *papi*)
    - In this portion of the vowel, the effect of V_2 was strongest. Formant values were not taken closer to the edge because in the last portion of the vowel the effect of V_2 on F1 could be easily masked by the effect of the closure for the intervocalic consonant /p/;
- **30-40-50%** of **/i/** duration (*papi)*

    - This was the more stable part of the vowel that could be used as reference for the vowel target. Formants were not taken in the second half of the vowel because the post-vocalic context changed depending on the corpus;
- **20-30%** of **ʃ[a]** and **p[a]** duration (*chat* and *papa*)
    - In this portion of the vowel the effect of C was strongest. Carryover CV effects had a smaller scope than anticipatory V-to-V effects, thus this portion of the vowel was the more comparable to the one selected for the *papi* items as for coarticulation effects. We did not take formant values directly at vowel onset because the Praat formant detection for /ʃa/ was less reliable at this timepoint due to the transition from the fricative to the vowel.

4.4.4 Correction of formant values

Formant values were checked visually in search of outliers, and detection errors were manually corrected. Correction was carried out in two ways:

1. The automatic formant detection in Praat was manually reset by adjusting the number of detected formants, for example in case of erroneous detection of an additional formant, and formants were extracted manually with the *get formant* function of Praat;
2. The portion of the vowel where formants had to be obtained was selected and the *extract spectral slice* function was used to manually get the values of the first two formants.

4.4.5 Averaging of values and Bark transformation

Formant values obtained for the considered timepoints were averaged on a single value per vowel. This procedure was carried out to make up for detection errors that could have escaped manual correction.

In order to be able to pool male and female data in a single analysis the formant values in Hertz were transformed in Bark using the Traunmüller (1990) formula:

$$[Z_i = 26.81/(1+1960/F_i) - 0.53]$$

4.5 Measures of V-to-V coarticulation

Two measures were carried out to capture coarticulation. The next sections will present the measures and discuss the reasons of the measure choice.

4.5.1 Contextual difference index

The influence of /i/ on /a/ translates in a lowering of F1 and a rise in F2, thus [a]i tends to be less compact than [a]a (Figure 4). To capture this multidimensional effect, a composite measure of *F2-F1 compacity* is taken on the vowel.

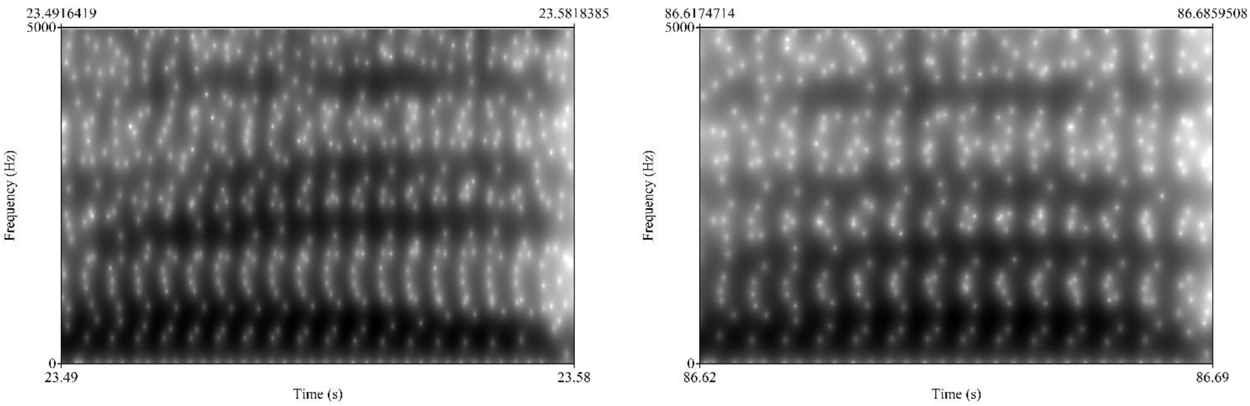

Figure 4. Spectrogram of [a]i (left) compared to [a]a (right). The compacity between the two formants (F2-F1) of /a/ appears to be reduced in papi compared to papa: F1 is lower and F2 is higher. Examples extracted by the production of a female speaker participating in the tempo and boundary studies.

Coarticulation is considered as the **contextual difference** between the *compacities* of [a]i and [a]a. The higher the *compacity* of [a]i, and thus the less compact with respect to the [a]a, the more there is coarticulation with /i//

4.5.2 Acoustic assimilation index

The measures of *F2-F1 compacity* of V_1 and V_2 in each *papi* token is used to measure the degree of ***acoustic assimilation*** of [a]i to /i/.

An *acoustic assimilation index* is made by calculating, token by token, the difference between the *compacity* of [a]i and /i/, divided for the *compacity* of /i/, for each *papi* item. The inclusion of F2-F1 of /i/ in the denominator is meant to relativize the difference between the two vowels on the specific realization of /i/ in the given token.

Acoustic assimilation index formula:

$$\frac{((F2 - F1/a/)\ -\ (F2 - F1/i/))}{(F2 - F1/i/)}$$

The index computes the degree to which the/a/ vowel assimilates to the spectral characteristics of the following /i/ within the same token. A higher value of this coarticulation index means that /a/ is spectrally more similar to /i/, and thus indicates more coarticulation; a lower value means that /a/ stays more spectrally distinct from /i/, and thus indicates less coarticulation (Figure 5).

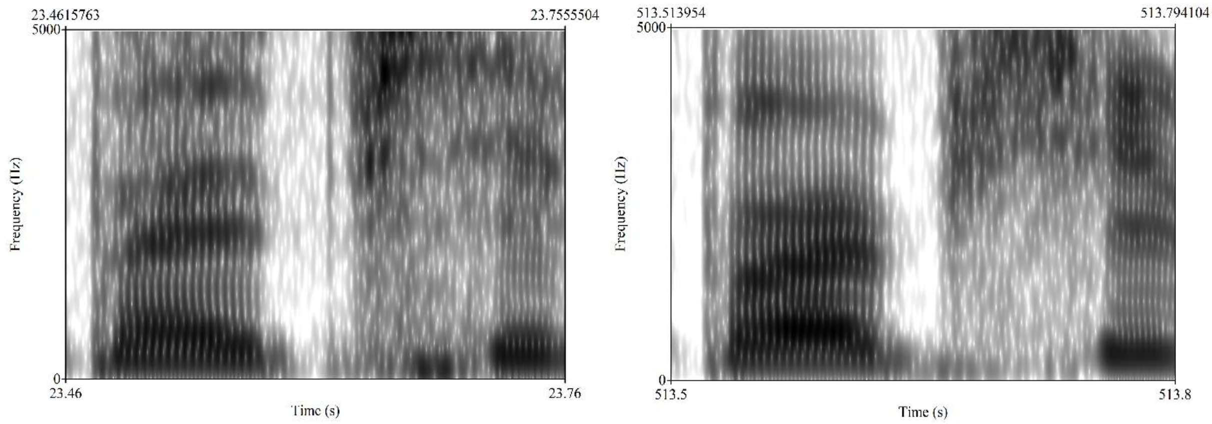

Figure 5. Exemple of a papi token with a high acoustic assimilation index (left) vs a papi with a lower acoustic assimilation index (right). Examples taken from the production of two speakers participating in the tempo and boundary studies.

4.5.3 Why two measures of coarticulation?

The two measures of coarticulation capture two different aspects of coarticulation and complement each other.

The measure of *compacity* is used as an *index of contextual difference,* capturing the contextualization of /a/ according to the following vowel. It allows testing whether in a given group, speaker or condition there is a significant difference between the two contexts and to what extent. This measure takes into account possible group-, speaker- or condition-specific realizations of [a]a. It is particularly useful to test for the presence of coarticulation in conditions where coarticulation could also not occur, for example when the two vowels are separated by a pause (in the clause boundary condition of the **boundary** study, or also in AoS speakers in the **MSD** study). This kind of measure, that is, the comparison between the formant values of a phoneme in a test and a control context is a method frequently used in acoustic studies on coarticulation (e.g. Recasens, 1989; Guitard-Ivent, 2018b).

The *acoustic assimilation index* captures to what degree [a]i is influenced by the following /i/ in each *papi* item. It allows testing the degree of coarticulation of [a]i with /i/ in each token, for each specific group, speaker and condition, without relying on an external reference. Moreover, this measure takes into consideration the specific realization of the trigger vowel /i/, which can have an influence on the formant values of the coarticulated [a]i. Since with this measure one value of coarticulation degree per token is obtained, this measure allows testing whether in one group or in one condition there is significantly more coarticulation than in another group or condition. Moreover, it allows correlations with other measures, such as speech rate.

A similar combination of two measures of coarticulation is used for instance by Cho (2004), but from an articulatory perspective.

4.6 Measures of articulation rate (MSD, age, tempo studies)

In the **MSD**, **age** and **tempo** studies, the relationship between coarticulation and the temporal organization of speech is investigated. In order to do so, different measures of articulation rate are taken depending on the study.

Here rationale behind each measure is briefly presented. A detailed reminder of each measure is in the method sections of each study.

- **MSD** study (chapter 5). The duration of V1 of the target item *papi* and the lag between the two vowels, from V1 offset to V2 onset, is taken.
    - ***Vowel duration.*** This measure is used as a proxy of articulation rate. It is chosen not to take a global measure of articulation rate in this study because AoS speakers could produce intersyllabic pauses. The presence of voiceless consonants in the target items and in the text indeed prevented identifying these pauses, and therefore calculating a measure of articulation rate for these speakers. Changes in rate have been shown to overall linearly affect vowel duration (Gay, 1981).
    - ***Vowel-to-Vowel lag***. This measure is taken as a more local measure of distance between the vowels, in order to test for an effect of longer lags between the vowels on the degree of overlap between them.
- **Age** study (chapter 6). An "external" measure of articulation rate per speaker calculated as part of the MonPage screening protocol (Laganaro *et al.,* 2021) was used in this study. It was calculated on the sentence *Melanie vend du lilas*, recorded as part of the protocol. A measure of speech rate was calculated as the number of expected phonemes over total sentence duration. A part of the production of the older speakers were checked in order to control for the presences of pauses in their production. Since few speakers inserted a pause between the subject and the verb, this measure is taken as a measure of *articulation rate*.
- **Tempo** study (chapter 7). A measure of *articulation rate* was calculated on the carrier sentences produced by the speakers as the number of expected phonemes over total sentence duration minus pauses duration. Articulation rate was taken in order to ensure that speakers had successfully complied the tempo task by decreasing and increasing their rate.

4.7 Statistical analyses: Linear Mixed Models

Data of the **MSD, tempo,** and **boundary** studies are analyzed through Linear Mixed Models built in the R environment (R Core Team, 2019) with the lme4 package (Bates, Maechler, Bolker, & Walker, 2014). The R^2 of each model is obtained with the r.squaredGLMM function of the MuMin package (Barton, 2009). The function returns two coefficients of determination, the marginal R^2, which evaluates how much variance in the data is explained by the fixed effects, and the conditional R^2, which evaluates how much variance in the data is explained by the fixed and the random effects combined. In the results, these will be presented as R^2m and R^2c. Multicollinearity is tested with the Variance Inflation Factor, through the vif function from the car package (Fox & Weisberg, 2019). The normality of residuals and homoskedasticity are tested visually with density curves, QQ plots and by displaying the residual values along the regression line (Winter, 2019). The main effects of each factor and of the interaction was tested by comparing the model with a certain factor or interaction with a model that lacks that particular factor by performing (by hand) the Likelihood ratio test as implemented in the anova() function. Post-hoc comparisons were carried out when needed using the lsmeans function (emmeans package). The coefficient estimate for each variable level and the significance of the difference between levels are obtained thanks to the post-hoc comparisons and given in the results. The significance threshold of p value is fixed at 0.05. The details of the fitted models are described in the statistical analysis section of each study.

5 Anticipatory V-to-V and carryover C-to-V coarticulation in different Motor Speech Disorders

5.1 Goals of the study

In this study, extrasyllabic V-to-V anticipatory coarticulation and intrasyllabic C-to-V carryover coarticulation are compared in the production of four groups of speakers affected by Apraxia of Speech (AoS) and by three different types of Dysarthria associated with Parkinson's Disease, Amyotrophic Lateral Sclerosis (ALS), Wilson Disease, *vs* the production of a group of neurotypical controls. Moreover, it is investigated whether differences in coarticulation degree between MSD speakers and control speakers can be accounted for by differences in segmental durations between the groups.

The first goal of this study is to investigate whether AoS speakers and Dysarthric speakers present different patterns of anticipatory V-to-V and carryover C-to-V coarticulation, that is, of two types of coarticulation that differ in direction, but also in the domain over which coarticulation is observed: the word vs the syllable.

Different predictions can be made regarding the coarticulatory patterns of AoS and dysarthric speakers. In AoS, extrasyllabic anticipatory V-to-V coarticulation could be expected to be impaired, while C-to-V intrasyllabic carryover coarticulation could be preserved. An impairment in V-to-V coarticulation for AoS speakers, which has been reported in the literature, can be considered to result from an impairment in motor planning and can account for another characteristic of AoS speech, that is, syllable segregation (longer duration of syllables and silent pauses between syllables). Conversely, carryover C-to-V coarticulation could be unaffected in AoS, if carryover and anticipatory coarticulation rely on different mechanisms, or if the coordination between speech elements is unaffected at the level of the syllable for AoS speakers.

Predictions are more difficult for dysarthric speakers. However, since the impairment for dysarthric speakers is not at the level of motor planning, but at the level of motor programming and execution, a certain degree of V-to-V coarticulation could be expected to be preserved in these speakers, while C-to-V carryover coarticulation could be expected to be more impacted.

Notwithstanding, AoS and Dysarthria, although different in terms of pathomechanism, are both characterized by a deficit in the temporal organization of speech. Slower movements are reported in AoS, and in flaccid and spastic dysarthria, while syllable segregation, with long silent intervals between syllables, is reported for AoS speakers. For this reason, a second goal of this study is to investigate whether differences in coarticulation patterns between healthy controls and MSD speakers can be accounted for by differences in the temporal organization of speech. To this purpose, measures of segmental durations were taken on the target items for V-to-V and C-to-V coarticulation in order to look for differences between the groups.

Moreover, to further test whether impairment in V-to-V coarticulation in speakers affected by MSD can be related to the temporal organization of their speech, regardless of underlying pathology, the relationship between the degree of V-to-V coarticulation and vowel duration, and between the degree of V-to-V coarticulation and the between-vowels distances (V-to-V lags) is examined in the MSD group, all pathologies confounded, and in the control group.

5.2 Methods

5.2.1 Participants

The productions of forty speakers affected by Motor Speech Disorders aged 24 to 93 (17 F, 23 M) and forty neurotypical speakers aged 25 to 83 (16 F, 24 M) were selected for this study.

5.2.1.1 MSD group

The forty MSD speakers were equally distributed in four groups according to pathology: post-stroke Apraxia of Speech (**AoS**), spastic-flaccid dysarthria associated with Amyotrophic Lateral Sclerosis (**D-ALS**), hypo- and hyperkinetic dysarthria associated with Wilson Disease (**D-Wl**) and hypokinetic dysarthria associated with Parkinson's Disease (**D-Park**). Each group was formed by 10 speakers. The patients were recruited in France (Pitié-Salpêtrière University Hospital and Lariboisière Hospital, Paris) and in Switzerland (Geneva University Hospitals). The recruitment was carried out in the context of two projects: MonPaGe, in collaboration with the universities of Geneva, Switzerland, and Mons, Belgium (Laganaro *et al.*, 2021) and MoSpeeDi (Projet Sinergia FNS-CRSII5_173711/1).

Mean age of the speakers in each group depended on the underlying pathology. It was overall higher in case of neurodegenerative disorder, therefore in the **D-Park** (*M*=71.5, SD=11) and **D-ALS** groups (*M*=74, SD=4.4), than in the **AoS** (*M*=52.5, SD=15.6) and **D-Wl** (*M*=35.5, SD=7.7).

All patients presented a level of severity that ranged from mild to moderate, except for one **AoS** patient who could be classified as severe. Severity was moderate for the **AoS** group and the **D-Wl** group, Mild to Moderate for the **D-ALS** group, and Mild for the **D-Park** group. The severity of the speech of the patients was assessed on a perceptual basis by expert Speech Language Therapists with the 'perceptual score' (PS) of the BECD (Auzou & Rolland-Monnoury, 2019), a composite score qualifying 'voice quality', 'phonetic realization', 'prosody', 'intelligibility', and 'naturalness of speech', each on a 4-point scale.

A summary of the characteristics of the MSD groups is given in Table 2 below.

	AoS	D-ALS	D-Wl	D-Park
Dysarthria Type		Mixed (Spastic & Flaccid)	Mixed (Hypo & Hyperkinetic)	Hypokinetic
Age	*M*=52.5, SD=15.6	*M*=74, SD=4.4	*M*=35.5, SD=7.7	*M*=71.5, SD=11
Sex	6 F, 4 M	7 F, 3 M	1 F, 9 M	3 F, 7 M
Perceptual Score/20	9 (Min 5, Max 15, SD=3.2)	7.8 (Min 4, Max 14, SD=3)	9.2 (Min 6, Max 12, SD=2)	6,7 (Min 3, Max 10, SD=1.9)
'Phonetic realization score/4	*M*=2.1, SD=0.7	*M*=1.6, SD=1	*M*=2.2, SD=0.6	*M*=0.8, SD=0.8
Severity	Moderate	Mild/Moderate	Moderate	Mild

Table 2. MSD study. Age, sex and severity of impairment for the speakers in the AoS (=Apraxia of Speech), D-ALS (=Amyotrophic Lateral Sclerosis), D-Wl (=Wilson Disease) and D-Park (Parkinson's Disease) groups.

5.2.1.2 Control group

The forty control speakers (**CTRL** group) were selected from the MonPaGe_HA (for Healthy Adults) database of spoken French (Fougeron, Delvaux, Menard, & Laganaro, 2018).

The MonPaGe_HA database contains the productions of speakers of various ages recruited through advertisements and among relatives in different French speaking locations: Paris (France), Geneva (Switzerland), Mons (Belgium), Montréal (Canada). All participants spoke French as their primary language (mother tongue and currently used language). Speakers over the age of 75 recorded for this database were screened for language and cognitive deficits (with the e-GeBAS, Chicherio, Genoud-Prachex, Assal, & Laganaro, 2019; or the MMSE, Folstein, Folstein & McHucg, 1975).

Productions of participants recruited in Paris and Geneva were selected for this study (mean age 53.5).

5.2.2 Speech material

Speech material for the reading task was the short text described in section 4.3.

Target words were *papi*, *papa* and *chat*. Six tokens of each target word occurred in the text. However, because of reading errors, some speakers produced *papa* instead of *papi* or viceversa while reading the text. Therefore, for some of the target words, there were less than six tokens per speaker:

- Controls: two speakers produced five *papi* and seven *papa*.
- MSD: seven speakers produced five *papi*, one speaker produced seven *papi*, and one speaker produced four *papi*. Five speakers produced seven *papa*, one speaker produced *five papa* and one speaker produced eight *papa*.

In total, 474 *papa* and 460 *papi* were analyzed.

5.2.3 Acoustic measures

5.2.3.1 Measures of coarticulation

5.2.3.1.1 V-to-V coarticulation

Coarticulation is measured as the contextual difference in *F2-F1 compacity* between [a]i of *papi* and [a]a of *papa,* and as the *acoustic assimilation* of [a]i to /i/ in *papi* as captured by the ***acoustic assimilation index***. The two measures are described in section 4.5.

5.2.3.1.2 C-to-V coarticulation

C-to-V carryover coarticulation is measured as the influence of C /ʃ/ on V_2 /a/ (**ʃ[a]**) in *chat*. V_2 /a/ preceded by /p/ (**p[a]**) in the word *papa* is used as control context. For this analysis, V_2 (instead of V_1) *papa* is taken because it was more comparable with **ʃ[a]**. Both vowels were followed by a varied segmental context and were susceptible to be accented.

Coarticulation is measured as the **contextual difference** between ʃ[a] and p[a]. Indeed, the influence of /ʃ/ on /a/ translates in a lowering of F1, and in a rise in F2, and thus in a less compact /a/ vowel. The higher the *compacity* of ʃ[a] with respect to p[a] tokens, the more there is coarticulation (Figure 6).

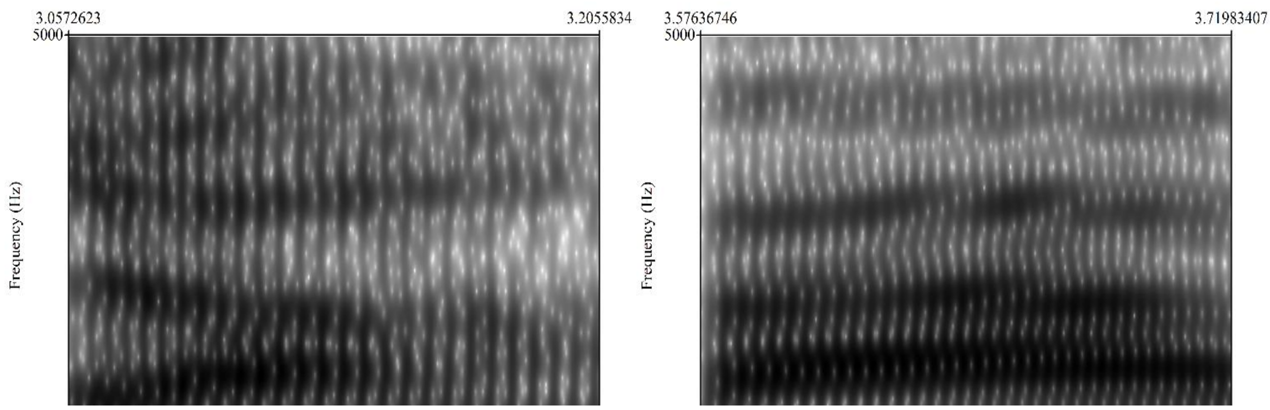

Figure 6. V_2 /a/ in chat (on the left) compared to V_2 /a/ in papa (on the right). At the beginning of /a/, F1 is lower and F2 higher in chat than papa.

5.2.3.2 Measures of segmental duration

Two measures of segmental duration were taken on each *papi* token:

- The duration of V1 [a]i in ms;
- The duration of the *V-to-V lag*, from the offset of V1 to the onset of V2, in ms;

V-to-V lags corresponded, for the majority of the speakers, to the duration of the intervocalic /p/. However, for AoS speakers, an intersyllabic pause was susceptible to be include. Indeed, AoS speech being characterized by syllable segregation, these speakers could present a pause between the two syllables of *papi*. Being /p/ a voiceless consonant, it was impossible to determine if a long silent interval before the burst for these speakers corresponded to a longer occlusion or to a silent pause after V_1 + /p/ occlusion.

- The duration of /a/ in *chat* was also taken in ms, to look for possible group differences in the production of C-to-V items.

5.2.4 Statistical analyses

The first analysis was run to test whether all groups present a contextual difference between the target vowels due to the influence of V_2 (V-to-V coarticulation) or C (Cto-V coarticulation), and whether the magnitude of this difference changes depending of the group.

To this purpose, two separate models were built, with ***compacity*** of [a]i and [a]a, and *compacity* of ʃ[a] and p[a], as dependent variables. Fixed effects were **context** (V_2 /a/ vs /i/, C /p/ vs /ʃ/) and **group** ("CTRL", "AoS", "D-Park", "D-ALS", "D-Wl") in interaction. Random intercepts for **sentence** and **speaker** and random slopes for **context**, in correlation with **speaker**, were modeled. Pairwise comparisons were carried out to observe the extent of the contextual difference for each group.

The fitted regressions were:

***F2-F1 compacity*~ Context*Group + (Context|Speaker) + (1|Sentence)**

To test whether the patients' groups show a lesser degree of V-to-V *acoustic assimilation* compared to healthy controls, a model with *acoustic assimilation index* as dependent variable and **group** as fixed effect was built. Random intercepts for **sentence** and **speaker** were modeled. No pairwise comparisons were run for this model. The CTRL group was set as the level of reference.

The fitted regression was:

***Acoustic assimilation index* ~ Group + (1|Speaker) + (1|Sentence)**

Finally, to test whether differences in V-to-V coarticulation degree could be accounted for by segmental duration differences, the effect of V_1 duration and V-to-V lag (in *papi*) in interaction, on *acoustic assimilation* was tested. Random intercepts for Speaker and Sentence were modeled. Two models were run, one for the MSD group all pathologies confounded, and one for the control group.

The fitted regressions were:

***Acoustic assimilation index* ~ V_1 duration (centered) * V-to-V lag (centered) + (1|Speaker) + (1|Sentence)**

In the next section, the results of these analyses will be presented in the same order as this section.

Before the results on the influence of the temporal cues on the *acoustic assimilation index*, differences in V_1 duration and V-to-V lag duration between the control group and the MSD groups will be presented and discussed.

5.3 Results

5.3.1 V-to-V coarticulation

5.3.1.1 Contextual difference index

Figure 7 displays *F2-F1 compacity* values for [a]i (in red) and [a]a (in yellow) for all groups. The higher the *compacity* values of [a]i are with respect to [a]a, the more there is coarticulation.

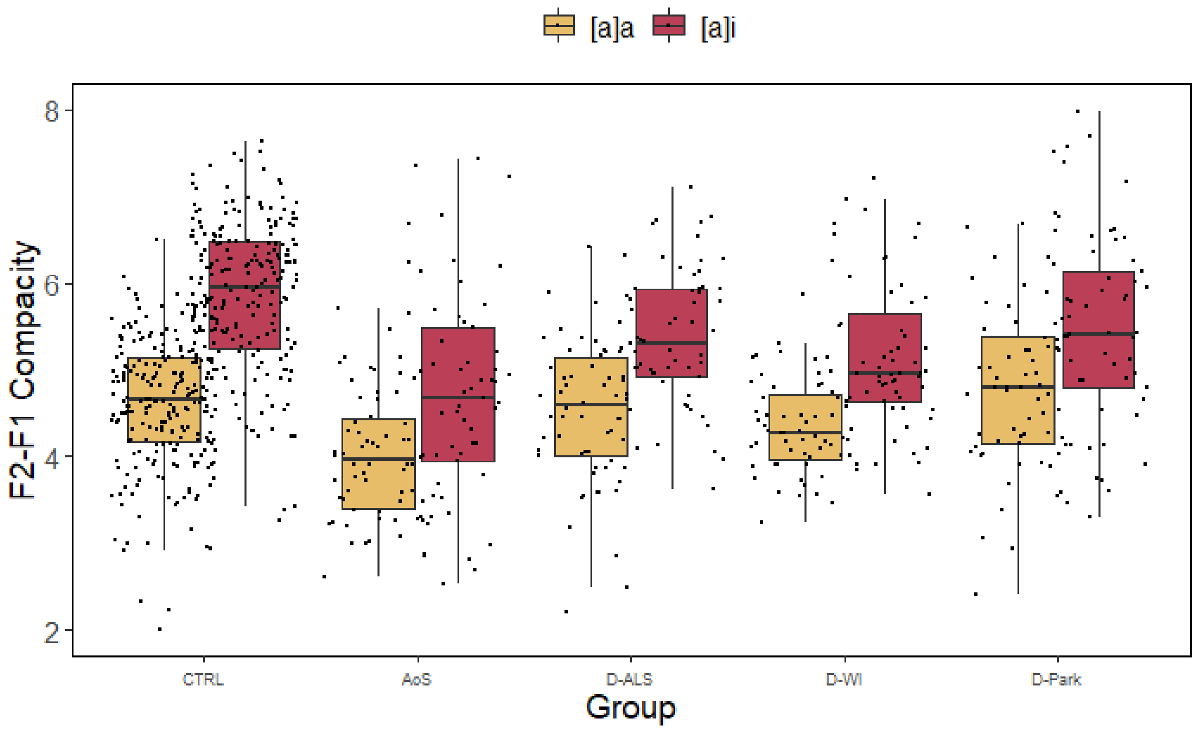

Figure 7. MSD study. F2-F1 compacity for [a]i (in red) vs [a]a (in yellow) for each group (CTRL=controls, AoS=Apraxia of Speech, D-ALS=Amyotrophic Lateral Sclerosis, D-Park=Parkinson's Disease). The higher [a]i compacities are with respect to [a]a, the more there is coarticulation.

The R^2 of the fitted model, a summary of the fixed effects and interactions, and the results of the pairwise comparisons are given in Table 3.

F2-F1 compacity is found to be affected by the nature of V_2 (/a/ vs /i/): *compacity* is higher for [a]i than for [a]a. However, the extent of the contextual difference depends on **group**.

The pairwise comparisons show that the contextual difference between [a]i and [a]a is significant for all groups. Therefore, it can be said that all groups present coarticulation. However, the extent of the contextual difference depends on the group (as showed by the coefficient estimates), with control group being estimated to present the greatest contextual difference. This difference reduces for the other groups to different extents, with AoS and D-Park speakers presenting the least contextual difference between the two contexts, and therefore less coarticulation.

***F2-F1 compacity~* Context*Group + (Context|Speaker) + (1|Sentence)** (R²m=0.30, R²c=0.79)			
Summary of the fixed effects		Pairwise comparisons [a]i vs [a]a	
Context	$\chi^2(1)$=41.8, *p*=<0.0001	CTRL	β=1.20, SE=0.1, *p*= <.0001
Group	$\chi^2(4)$= 9.5, *p*= 0.05	AoS	β=0.69, SE=0.2, *p*= 0.0002
Context*Group	χ^2 (4)=17.3, *p*=0.0017	D-ALS	β=0.84, SE=0.2, *p*= <.0001
		D-Wl	β=0.80, SE=0.2, *p*= <.0001
		D-Park	β=0.63, SE=0.2, *p*= 0.0005

Table 3. MSD study. Effect of context and group on F2-F1 compacity of V1 /a/. On the left: χ^2, degrees of freedom, and p values for the fixed effects. On the right: coefficient estimates, standard errors, and p values for the pairwise comparisons.

5.3.1.2 Assimilation index

Figure 8 displays the *acoustic assimilation index* for each group. The higher the *index* value, the greater the acoustic assimilation of [a]i to /i/, and thus, coarticulation.

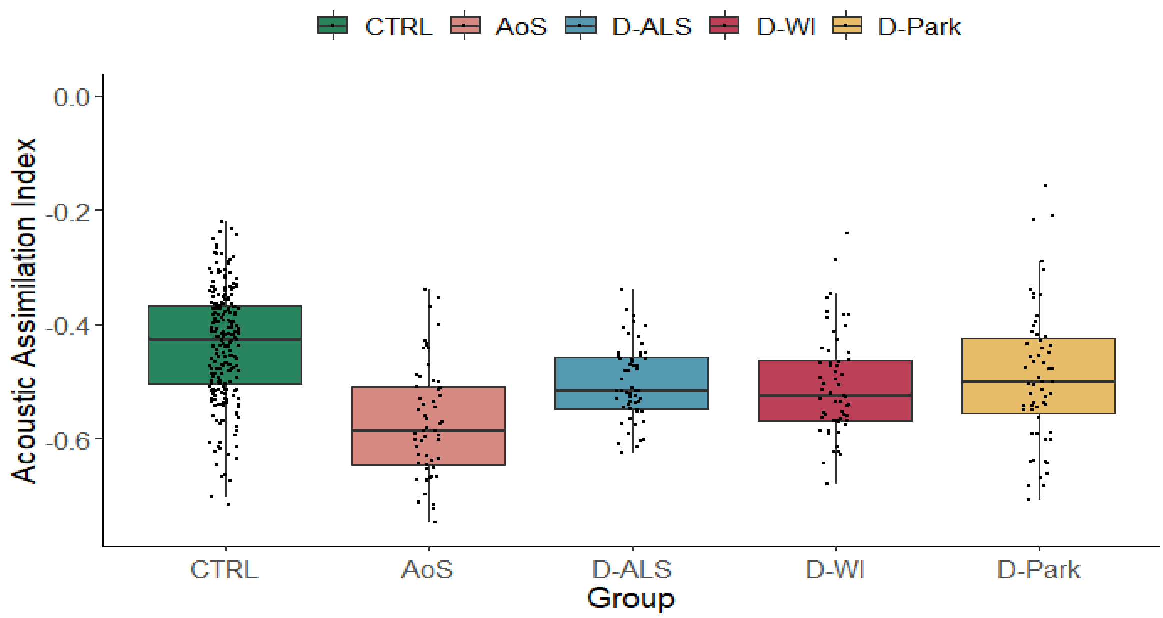

Figure 8. MSD study. *Acoustic assimilation index values per group (CTRL=controls, AoS=Apraxia of Speech, D-ALS=Amyotrophic Lateral Sclerosis, D-Park=Parkinson's Disease). The higher the index, the more the acoustic assimilation of [a]i to /i/.*

The R^2 of the fitted model, coefficient estimates, standard errors and p values for each group are given in Table 4.

Group has a significant effect on the *acoustic assimilation index*. The AoS, D-ALS and D-Wl groups significantly reduce the extent to which [a]i is assimilated to /i/ in *papi* with respect to the CTRL group. AoS speakers reduce the *acoustic assimilation* to the greatest extent, followed by the D-Wl speakers and by the D-ALS speakers. D-Park speakers do not significantly reduce the *acoustic assimilation* with respect to CTRL speakers.

***Acoustic assimilation index* ~ Group + (1|Speaker) + (1|Sentence)** (R2m=0.19, R^2c=0.77)
Reference level: CTRL group

AoS	β= -0.14, SE= 0.03, *p*= <.0001
D-ALS	β= -0.07, SE= 0.03, *p*= 0.025
D-Wl	β= -0.07, SE= 0.03, *p*= 0.017
D-Park	ns

Table 4. MSD study. Effect of group on the acoustic assimilation index. Coefficient estimates, standard errors, and p values for each MSD group (reference: CTRL group).

5.3.2 C-to-V coarticulation

5.3.2.1 Contextual difference index

Figure 9 displays *F2-F2 compacity* values of ʃ[a] (in red) and p[a] (in yellow) for all groups. The higher the *compacity* values of ʃ[a] are with respect to p[a], the more there is coarticulation.

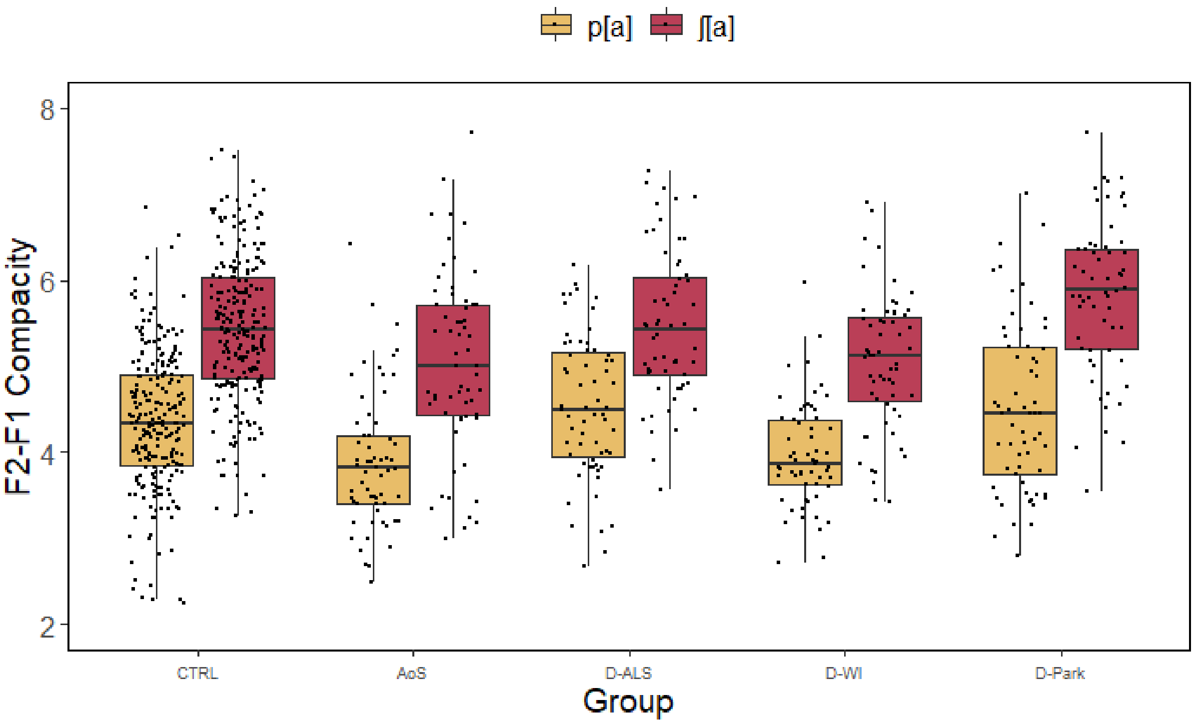

Figure 9. MSD study. F2-F1 compacity values of ʃ[a] (in red), vs p[a] (in yellow), for each group (CTRL=controls, AoS=Apraxia of Speech, D-ALS=Amyotrophic Lateral Sclerosis, D-Park=Parkinson's Disease). The higher the compacities of ʃ[a] with respect to p[a], the more there is coarticulation.

The R^2 of the fitted model, and a summary of the fixed effects and interactions is given in Table 5.

F2-F1 compacity of V2 /a/ is affected by the identity of the preceding consonant, but there is no interaction with **group**. *Compacity* is higher when /a/ is preceded by /ʃ/ than /p/, and this contextual difference is estimated to be, on average, of 1.14 Bark (SE=0.153, p=<.0001). The effect of the context does not significantly vary depending on group, meaning that all groups coarticulate to a similar extent. A significant effect of group is to be attributed to the production of overall more compact /a/ vowels for AoS speakers.

***F2-F1 compacity*~ Context*Group + (Context∣Speaker) + (1∣Sentence)** (R^2m= 0.33, R^2c=0.67)	
Context	$\chi^2(1)=25$, p=<.0001
Group	$\chi^2(4)=13.7$, p=0.008
Context*Group	ns

Table 5. MSD study. Effect of Context and Group on F2-F1 compacity of ʃ[a] and p[a]: χ2, degrees of freedom, and p values for the fixed effects.

5.3.3 Segmental durations

5.3.3.1 V-to-V coarticulation items

5.3.3.1.1 V_1 duration

In Figure 10, the duration of V_1 [a]i of each papi item is displayed for all groups.

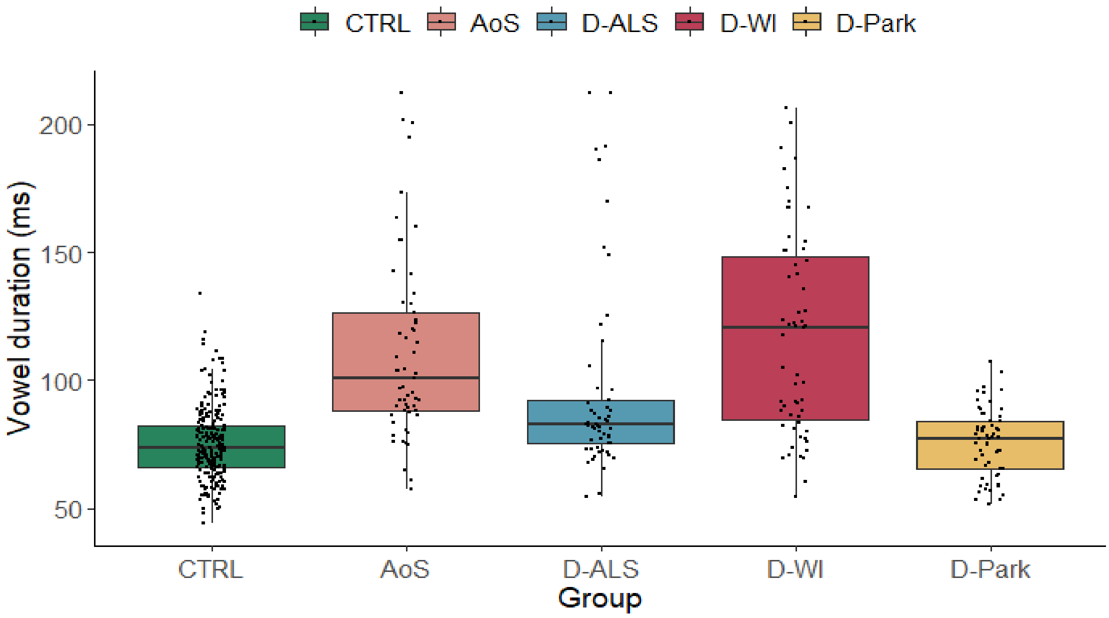

Figure 10. MSD study. Vowel duration for V_1 [a]i in papi items, for each group (CTRL=controls, AoS=Apraxia of Speech, D-ALS=Amyotrophic Lateral Sclerosis, D-Park=Parkinson's Disease).

It can be observed that V_1 duration varies depending on the group. With respect to the CTRL speakers, ALS, AoS and D-Wl speakers produce longer vowels. The longest vowels (*M*=117.6 ms) are exhibited by the D-Wl speakers (who also present the greatest variability in vowel duration across tokens, SD=39.5 ms). D-Park speakers' vowel duration does not differ from control speakers (*M*=75.7 ms, SD=13.4 ms).

5.3.3.1.2 V-to-V lags

In Figure 11, the duration of the lag between V_1 and V_2 in *papi* is displayed for each group.

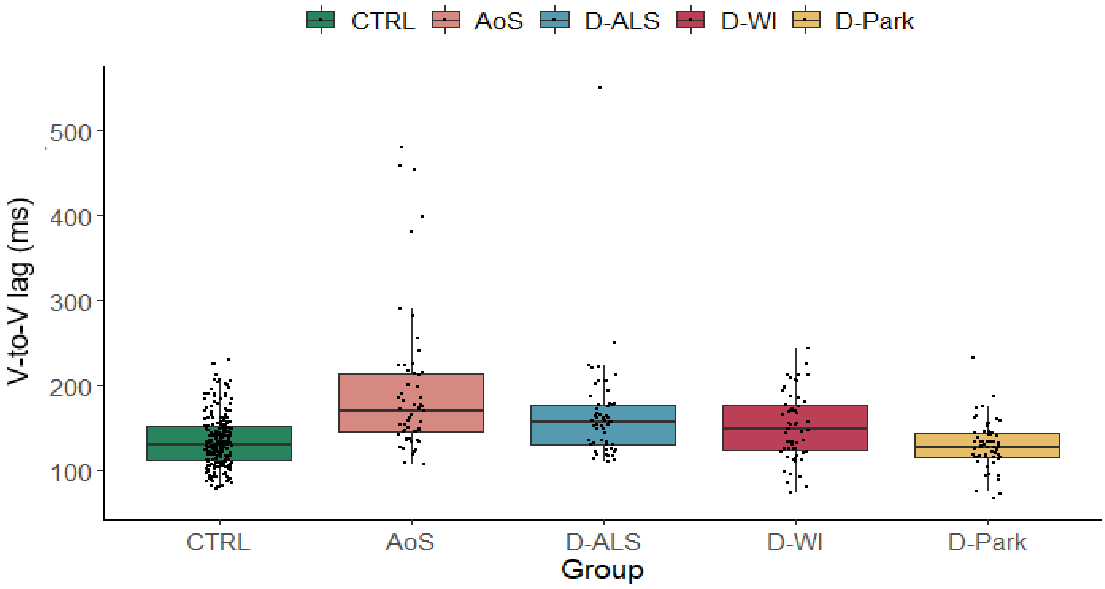

Figure 11. MSD study. V-to-V lag durations in papi for each group (CTRL=controls, AoS=Apraxia of Speech, D-ALS=Amyotrophic Lateral Sclerosis, D-Park=Parkinson's Disease).

It can be observed that, with respect to CTRL speakers, the AoS, D-ALS and D-Wl groups produce longer V-to-V lags. The longest lags are exhibited by the AoS speakers (*M*=194 ms, SD=86 ms), possibly for the insertion of silent pauses between the two syllables. The D-park group produce slightly shorter lags compared to the CTRL speakers.

5.3.3.2 C-to-V coarticulation items

5.3.3.2.1 V_2 duration

In Figure 12, the duration of ʃ[a] of each *chat* item is displayed for each group.

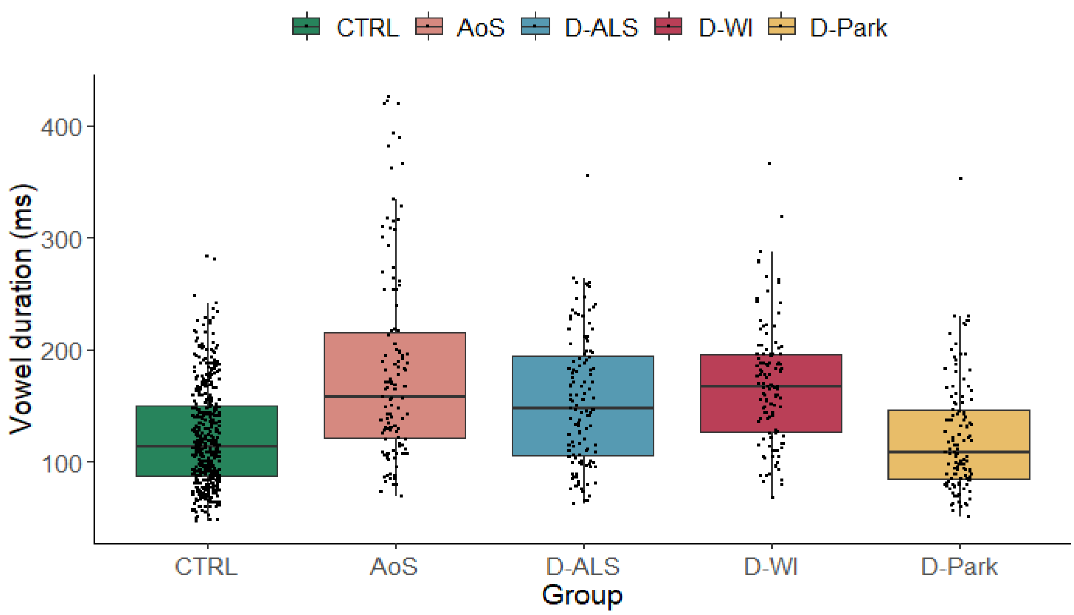

Figure 12. MSD study. V_2 /a/ duration in chat items for each group (CTRL=controls, AoS=Apraxia of Speech, D-ALS=Amyotrophic Lateral Sclerosis, D-Park=Parkinson's Disease).

It can be observed that, with respect to CTRL speakers, AoS, D-ALS and D-Wl speakers produce longer vowels. The longest vowel durations are exhibited by the AoS speakers (*M*= 183 ms, SD= 88.1), followed by D-Wl (*M*=169 ms, SD=53.4) and D-ALS (*M*=154 ms, SD=57.9). The mean vowel duration of D-Park speakers does not differ from CTRL speakers.

5.3.4 Effects of V_1 and V-to-V lag duration on the acoustic assimilation index

5.3.4.1 MSD all pathologies confounded

Figure 13 shows the relationship between V_1 duration (on the left) or V-to-V lag (on the right) of *papi* and the *acoustic assimilation index* for the MSD group, all pathologies confounded.

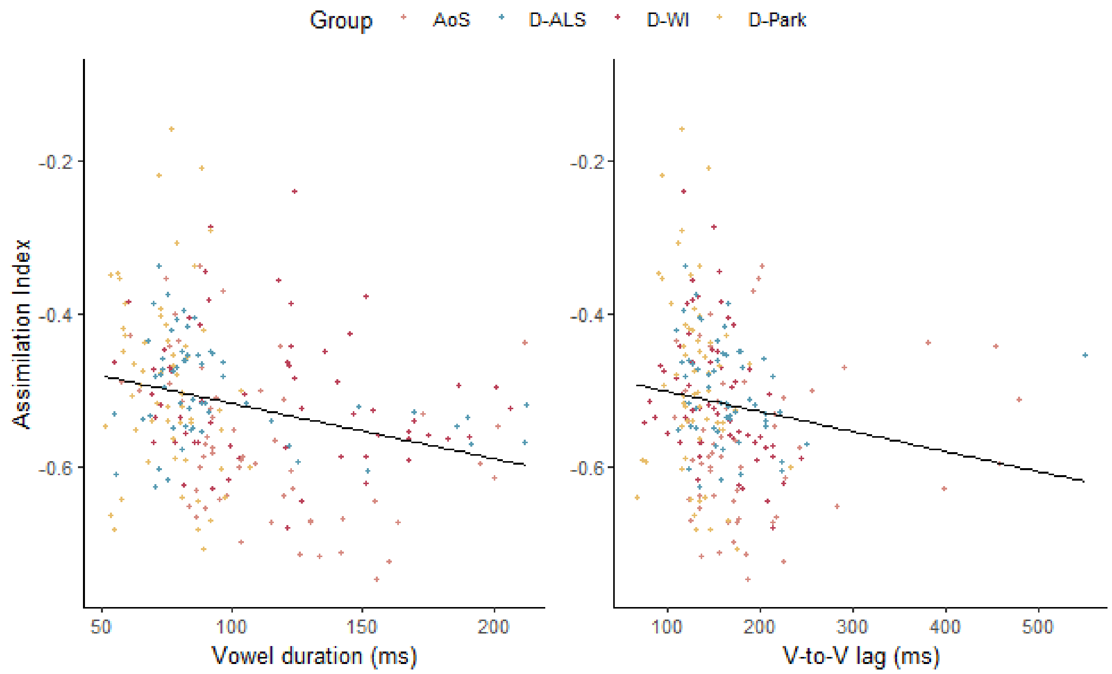

Figure 13. MSD study. Relationship between the segmental durations and the acoustic assimilation index in papi tokens for the MSD group, all pathologies confounded. On the left: relationship between V_1 duration (x axis) and the acoustic assimilation index (y axis). On the right: relationship between V-to-V lag duration (x axis) and the acoustic assimilation index (y axis).

The R^2 of the model and a summary of all the fixed effects and interactions are given in Table 6.

The *acoustic assimilation* of [a]i to /i/ in *papi* varies as a function of V_1 duration, but not V-to-V lag. An increase in vowel duration is accompanied by a reduction of *acoustic assimilation* (β= -0.00052, SE= 0.0002, p=0.032). The interaction between V_1 duration and V-to-V lag is not significant.

Acoustic assimilation index ~ **V duration*V-to-V lag + (1|Speaker) + (1|Sentence)** (R²m=0.032, R²c= 0.72)	
Vowel duration	$\chi^2(1)$=4.22, p=0.04
V-to-V lag	ns
V duration*V-to-V lag	ns

Table 6. MSD study. Effect of V duration and V-to-V lag on the acoustic assimilation index for the MSD group, all pathologies confounded: χ^2, degrees of freedom and p value for the significant fixed effect.

5.3.4.2 CTRL speakers

Figure 14 shows the relationship between V_1 duration (on the left) or V-to-V lag (on the right) and the *acoustic assimilation index* for the control group.

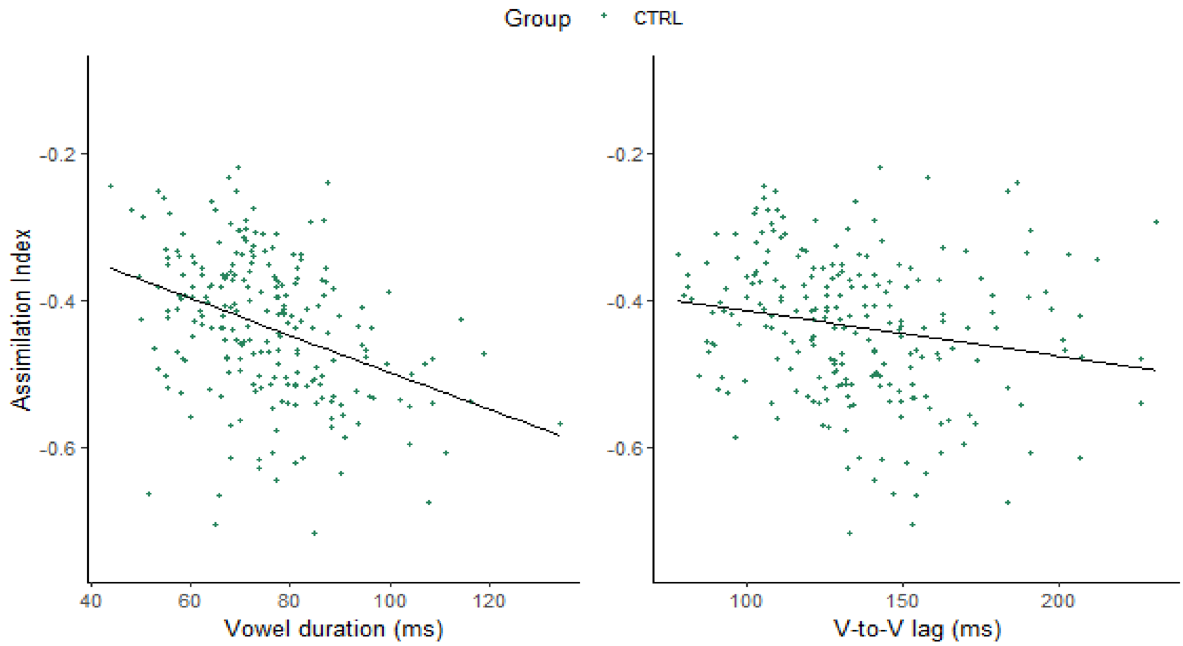

Figure 14. MSD study. Relationship between the segmental durations and the acoustic assimilation index in papi tokens for the CTRL group. On the left: relationship between V_1 duration (x axis) and the acoustic assimilation index (y axis). On the right: relationship between V-to-V lag duration (x axis) and the acoustic assimilation index (y axis).

The R^2 of the model and a summary of all the fixed effects and interactions are given in Table 7.

The *acoustic assimilation* of [a]i to /i/ in *papi* is marginally affected by vowel duration (*p=0.048)*. An increase in V_1 duration is accompanied by a reduction of the *acoustic assimilation index* (β= -0.00082 SE= 0.0004, *p*=0.046). V-to-V lag duration does not affect the degree of *acoustic assimilation.* The interaction between V_1 duration and V-to-V lag is not significant.

Acoustic assimilation index **~ V duration*V-to-V lag + (1ǀSpeaker) + (1ǀSentence)** (R^2m= 0.029, R^2c= 0.70)	
Vowel duration	$\chi^2(1)$=3.92, *p*=0.048
V-to-V lag	ns
V duration*V-to-V lag	ns

Table 7. MSD study. Effect of V duration and V-to-V lag on the acoustic assimilation index for the CTRL group: χ2, degrees of freedom, and p value for the significant fixed effect.

5.4 Summary of results

In this study, anticipatory V-to-V coarticulation is investigated in a group of ten speakers presenting Apraxia of Speech, and three groups of ten speakers each, presenting Dysarthria associated with Parkinson's Disease, ALS and Wilson Disease, compared to a group of forty neurotypical controls. Anticipatory coarticulation is measured as the contextual difference between the *F2-F1 compacity* of V_1 [a]i in *papi* and V_1 [a]a in *papa*; and as the *acoustic assimilation* of V_1 [a]i to V2 /i/ in *papi*. To compare patterns of V-to-V anticipatory coarticulation to patterns of local carryover coarticulation, C-to-V coarticulation is also examined, as the contextual difference between the *F2-F1 compacity* of V_2 ʃ[a] in *chat* and V_2 [a]a in *papa*. Segmental durations for the target items are taken in order to test for differences in the temporal organization of speech across groups. Moreover, the relationship between the degree of *acoustic assimilation*, V_1 duration and V-to-V lag duration in *papi* is analyzed separately for the MSD group and the CTRL group.

The main finding of this study is a reduction of V-to-V coarticulation in all MSD groups with respect to healthy controls. Indeed, the analysis on the *F2-F1 compacity* of V_1 /a/ of *papi* and *papa* shows coarticulation, measured as contextual difference, for all groups of speakers; but this difference is reduced for all the four MDS groups with respect to the controls. *Acoustic assimilation* of [a]i to /i/ in *papi* is significantly reduced for the AoS, ALS and Wilson Disease groups compared to the control group. The reduction of the *acoustic assimilation* in Parkinson's Disease speakers is not significant; however, this could be seen as the result of a hypoarticulation of the vowels, and does not have to be interpreted as these speakers presenting a degree of coarticulation comparable to control speakers (see the following section).

A second finding of this study is that, in the MSD group, the degree of *acoustic assimilation* is negatively correlated to vowel length, with a decrease of the assimilation degree at the increase of vowel duration. A descriptive analysis of V_1 durations shows that an increase in vowel duration is especially pronounced in Wilson Disease speakers, followed by AoS andALS. In Parkinson's Disease speakers, vowel duration is comparable to control speakers' (these results will be also commented in the next section). A decrease in *acoustic assimilation* at the increase in vowel length is also found for the healthy controls, but the effect is marginal. On the other hand, no group presents a relationship between the degree of *acoustic assimilation* and the distance between the two vowels, measured by the V-to-V lag.

Finally, C-to-V carryover coarticulation results unimpaired in the MSD groups. Indeed, the analysis on the *compacity* of V_2 /a/ of *chat* and *papa* shows coarticulation, measured as contextual difference, for all groups to a similar extent. Since the AoS speakers, the ALS and the Wilson Disease speakers present longer vowels, but no reduction of coarticulation, it is assumed that vowel length does not affect local carryover coarticulation.

5.5 On the results on V-to-V coarticulation in Parkinson's Disease

As seen in the results, V-to-V coarticulation, measured as the contextual difference between [a]i and [a]a, is found to be impaired in the D-Park group, but no significant difference is found in the *acoustic assimilation index* between control speakers and D-Park speakers. However, in this case the results on the *acoustic assimilation index* are misleading.

The acoustic assimilation index is sensible to capture not only the influence of /i/ on /a/ but also /a/ undershoot. Indeed, a lowering of the F1 of /a/ due to undershoot will result in more difference between F2 and F1, thus in a less compact vowel, and higher *F2-F1 compacity (*Figure 15):

Figure 15. Schematized effect of a lowering of F1 of /a/ on the F2-F1 compacity measure.

A lowering of F1 of /a/ due to an undershoot of the vowel in *papi* will thus lead to a /a/ vowel more spectrally similar to /i/, which.

What are the implications of this for the results on V-to-V coarticulation in MSD?

The Parkinson's Disease group is the group who showed the least reduction of the acoustic assimilation index with respect to the controls. The difference in the acoustic assimilation index between the D-Park and the CTRL group is not significant.

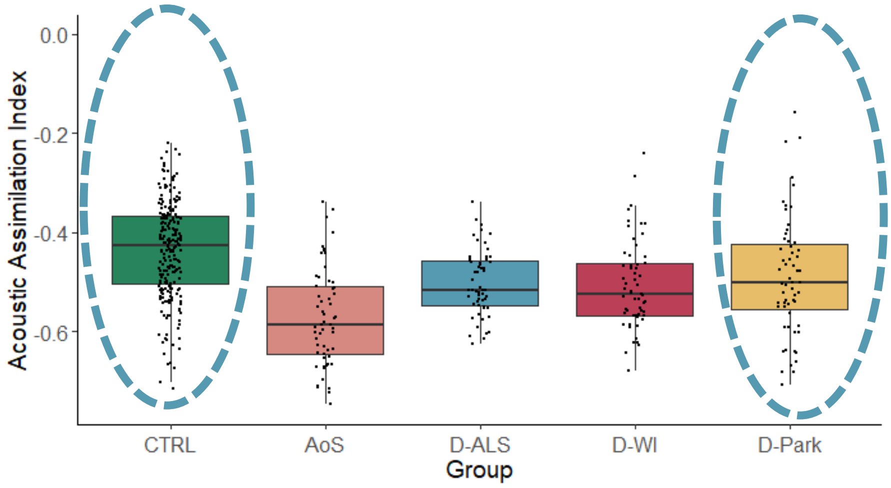

Figure 16. MSD study. Acoustic assimilation index for each group. In blue, the CTRL and D-Park groups are circled. The D-Park group did not show a significant reduction of the acoustic assimilation index with respect to the controls and was the MSD group with the highest acoustic assimilation index on average

However, the moderately high acoustic assimilation index showed by these speakers is not to attribute to a moderately high coarticulation degree, but to a lowering of F1 caused by undershoot. A similar effect of the reduction of movement displacement in Parkinsons' Disease on vowel acoustics, with more closed vowels in these speakers and in particular a lowering of F1 for the open vowel /a/ has been showed in French by Audibert & Fougeron (2012). If we look at *F2-F1 compacity* of [a]i and [a]a in CTRL and D-Park speakers, we can see that, with respect to the control speakers, D-Park speakers have tendentially higher *compacities* of [a]a that can be attributed to a lowering of F1 (and an overall reduction of the difference between the vowel /a/ in the two contexts, as showed by the analysis on *compacity*, Figure 17).

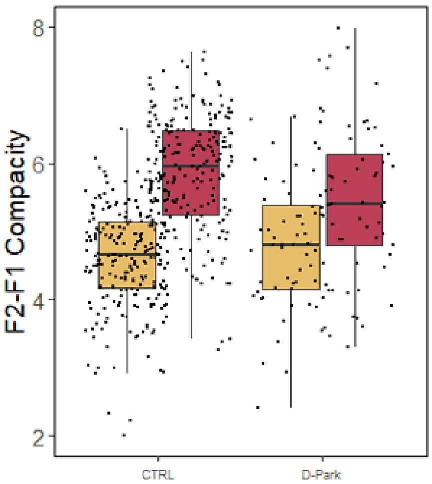

Figure 17. F2-F1 compacity for CTRL and Parkinson's Disease (D-Park) speakers. The compacity of [a]a tends to be higher in D-Park.

What is the consequence of this formant pattern of Parkinson's Disease speakers on the analysis on the relationship between vowel duration and coarticulation?

The results on the MSD group showed that V_1 duration of *papi* has an effect on coarticulation degree as measured by the *acoustic assimilation index*. This analysis was made on the MSD group alone, without the control speakers. Therefore, the speakers in this analysis who presented the shortest vowels and higher acoustic assimilation indices where the Parkinson's Disease speakers. However, we have seen that a high *acoustic assimilation index* for these speakers does not indicate a high coarticulation. This means that a reduction of coarticulation is found also in the MSD group who does not present a slowing in speech, and thus that a reduction in coarticulation in MSD speakers cannot be uniquely accounted for by longer segmental durations.

6 Changes in V-to-V coarticulation throughout adulthood

6.1 Goals of the study

In this study, anticipatory V-to-V coarticulation and its relationship with articulation rate are examined in 246 adults aged 20 to 93. Indeed, changes in coarticulation with age are attested in studies on childhood development, where children's patterns of coarticulation are compared to the patterns of coarticulation of young adults, considered as the reference for "adult" coarticulation. However, changes in speech during adulthood are attested in the literature. In particular, a slowing in speech rate and longer segmental durations have been reported with age. This has been mostly shown for groups of older speakers compared to groups of younger speakers.

The main goals of this study are to investigate whether coarticulation continues to change throughout life and whether a variation in coarticulation degree could stem from the attested slowing of speech with age.

A further goal of this study is to investigate whether a change in rate and V-to-V coarticulation with age, and the relationship between these factors, is linear. Indeed, the majority of the studies on the effects of age on speech have compared age groups of different compositions, making difficult to compare the findings of the literature. One question still open is whether it is possible to identify approximately a crucial age for changes in speech. The relationship between V-to-V coarticulation, articulation rate, and age as a continuous factor is analyzed with a non-linear regression with MARS modeling, in order to test for non-linearities in the changes in coarticulation and rate with age.

This study is the subject of a published article (D'Alessandro & Fougeron, 2021). The participant and results section of this chapter is reported as per the publication. The other sections were adapted to harmonize with the structure of this dissertation.

6.2 Methods

6.2.1 Participants

The productions of 246 healthy native French speakers (123 females and 123 males) spanning from 20 to 93 years of age were selected for this study. The distribution of speakers' age is illustrated in Figure 18. The recordings were selected from existing databases collected in the context of three related projects (the MonPage, MoSpeeDi, and Speech'N'Co projects).

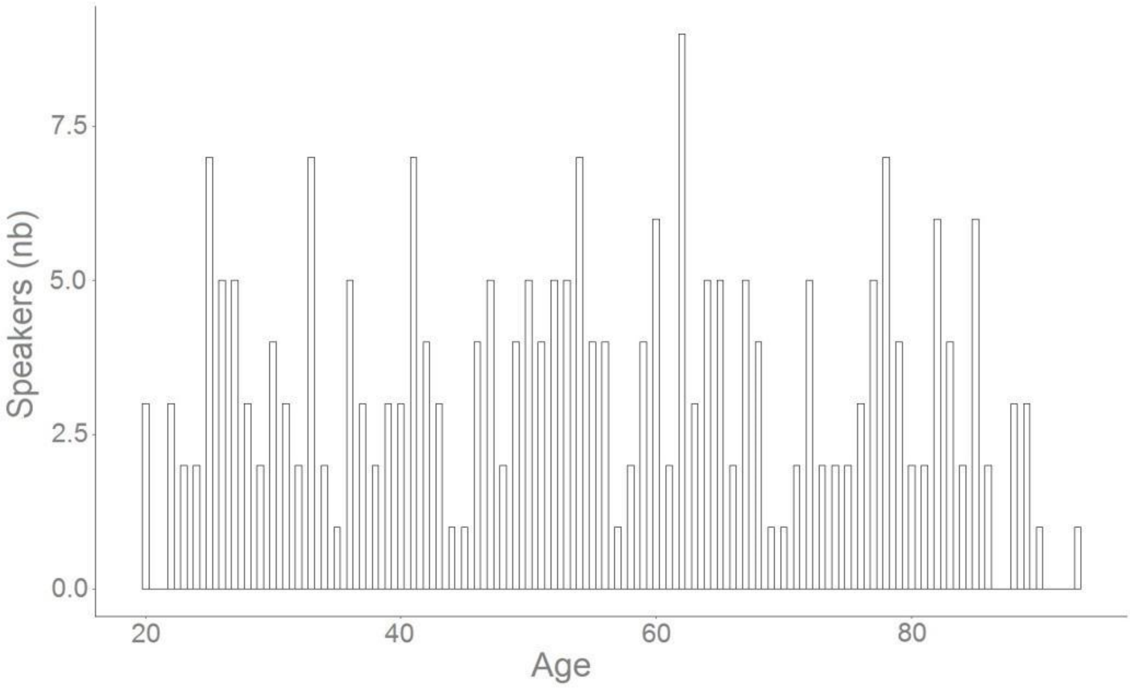

Figure 18. Age study. Distribution of the population according to the chronological age of the speakers.

Participants were recorded in three cities of different French-speaking countries: Paris, in France (42 females and 42 males), Geneva, in Switzerland (42 females and 42 males), and Mons, in Belgium (39 females and 39 males). They were all recruited from local communities in order to have a varied social and educational background in the population, but it was verified that recruitment was balanced across countries.

Regional diversity was also meant to introduce diversity in the population, but the inclusion of the participants was not strictly focused on well-defined regional varieties in each location. For example, speakers recorded in Geneva originated mainly from the larger Lemanic area; speakers recorded in Paris originated mainly from diverse regions within the northern half of France. All participants spoke French as their primary language (mother tongue and currently used language).

A subset of this data (127 speakers) was used in a pilot study (D'Alessandro and Fougeron, 2018) in order to test for confounds due to differences between French regional varieties. Dialectal differences were found in vowel duration, with participants from Belgium presenting longer vowels than both participants from France and Switzerland, but this regional property did not interact with age. Preliminary analyses on the variation of coarticulation according to age showed that the effect of age was similar in the three regional varieties. Therefore, speakers were pooled together for the present study.

6.2.2 Speech Material

Speech material for the reading task was the short text described in section 4.3.

V-to-V coarticulation was investigated on the target word *papi,* of which there were six occurrences in the text. Due to reading errors, 18 speakers produced five out of six /papi/ and four produced four out of six, thus a total of 1449 items were analyzed.

To measure articulation rate, speakers' production of the sentence *Melanie vend du lilas* ("Melanie sells lilac"), whose recording was part of the MonPaGe protocol, was used.

6.2.3 Acoustic measures

6.2.3.1 V-to-V coarticulation

V-to-V coarticulation is measured by the *acoustic assimilation index* as the acoustic assimilation of [a]i to /i/ in each *papi* token (described in section 4.5.2).

6.2.3.2 Articulation rate

As mentioned in the section 4.6 of the general methods, a measure of articulation rate per speaker was computed on the sentence *Melanie vend du lilas* as part of the MonPaGe screening protocol (Laganaro *et al*, 2021). The beginning and the end of the sentence were semi-automatically annotated in Praat in order to extract the total sentence duration. Speech rate was then calculated as the number of expected phonemes over sentence duration. For most speakers, it corresponds to a measure of articulation rate, but for few speakers who introduced a short pause after the subject of the sentence, it is a measure of speech rate (since the pause is included in the sentence duration).

6.2.4 Statistical analysis

6.2.4.1 MARS modeling

Statistical analyses were carried out using Multivariate Adaptive Regression Splines (MARS) models (Friedman 1991), built with the package earth (Milborrow, 2021) in the R software (R Core Team, 2021).

MARS modeling is an extension of linear models that can be used to model linear and non-linear relationships between variables. Unlike step functions, the nature of the non-linearity does not need to be assumed in advance. The model works on two steps. In the first one, the forward pass, the model looks for the first point in the data (knot) up to which a linear regression between x and y can be fitted with the smallest error, creating what is called a hinge function (a—0.x) or (0.x—a), where 'a' is the knot. It continues searching for these cutpoints until the end, creating a series of linear regressions. In the second step, the pruning one, the knots that do not contribute significantly to predictive accuracy are eliminated to avoid overfitting.

MARS automatically performs variables selection, excluding variables with no explanatory power (in case of collinearity) and assessing variable importance. Variable importance measures the impact of the prediction error as features are included (Friedman, 1991; Bohemke & Greenwell, 2019).

In earth, the models can be adjusted by manually specifying a series of parameters, such as the number of interactions between knots, on the basis of how many independent variables are included in the model, and the maximum number of knots, equally spaced, for which the model looks for in the forward pass (Milborrow, 2021). It is worth pointing out that, unlike linear models, the interaction in MARS do not indicate interaction between variables, but between knots found in the data for that variable.

For each model, as measures of performance estimate, the R^2 and the General R^2, which is the mean of the R^2 calculated for the different models earth created during the procedure, will be given.

6.2.4.2 Fitted models

First, in order to investigate the relationship between articulation rate and age (recall that rate is expected to slow down with aging), a MARS model was built with *articulation rate* as the dependent variable and **age** as the explanatory variable.

To test whether and how coarticulation covaries with age and speech rate, a MARS model was built with the *acoustic assimilation index* as dependent variable and **age** and **articulation rate**, as well as their interaction, as explanatory variables.

6.3 Results

6.3.1 Relationship between age and articulation rate

Figure 19 presents the distribution of the speakers according to their articulation rate and their chronological age. It clearly shows that speech rate decreases as speakers' age increases. This trend is continuous from 20 to 93 y.o.a. but a sharper slowing down occurs after middle age. The MARS model finds a knot at age 54, with a steeper decrease for speakers older than 54 ($\beta = -0.08$) than for speakers who are younger ($\beta = -0.04$). As expected though, speakers' age alone explains only a small portion of the variance in articulation rate in the population ($R^2 = 0.29$, $GVR^2 = 0.26$).

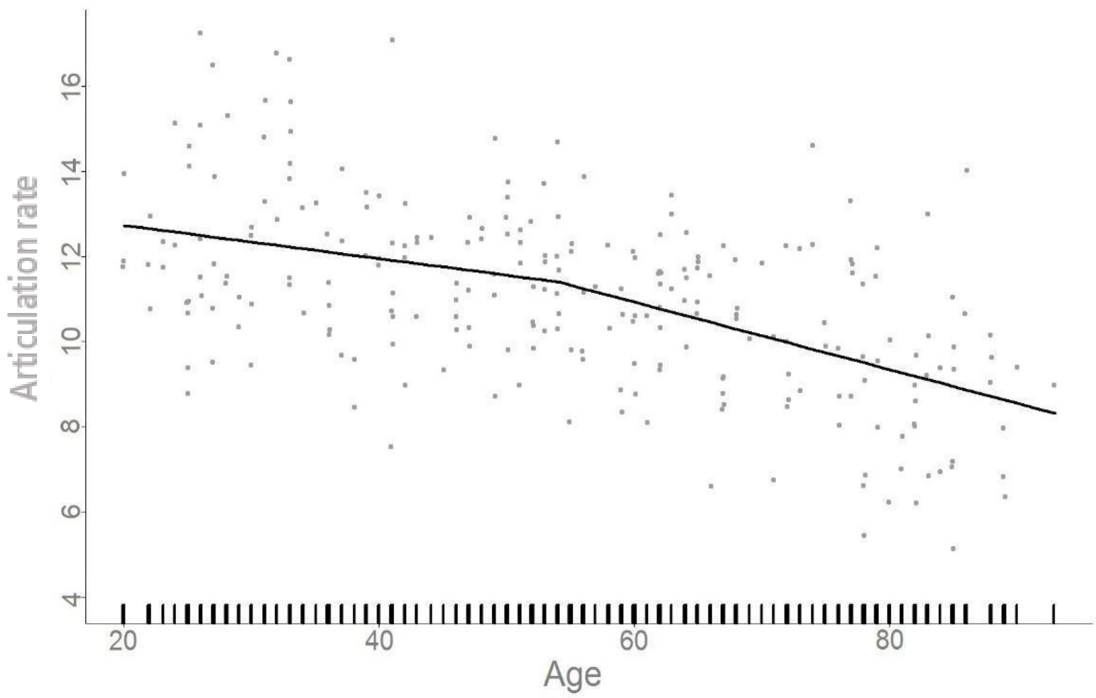

Figure 19 .Age study. Relationship between age (x-axis) and articulation rate in phoneme/s (y-axis) per speaker. The dark line represents the hinge functions of the MARS model, with a knot found at age 54.

6.3.2 Relationship between the acoustic assimilation index, articulation rate and age

In the second analysis, we tested how **age** and **articulation rate** predict coarticulation. The MARS model yielded the two covariates **age** and **articulation rate.** However, the two variables in the interaction explain the *acoustic assimilation index* only moderately ($R^2=0.18$, $GVR^2=0.16$). **Age** is given as the most important predictor, but the two factors are found to interact with each other in a complex but very interesting way. Before turning to the interaction between variables, it will described how each variable alone covaries with coarticulation.

Figure 20 presents the *acoustic assimilation indices* computed per token (5–6 per speaker) according to the speakers‘ age. It appears clearly that coarticulation reduces with an increase in age, but in a non-linear way. Coarticulation decreases smoothly up to a knot at 54 y.o.a. ($\beta = -0.003$) and then more abruptly after 70 y.o.a. ($\beta = -0.006$).

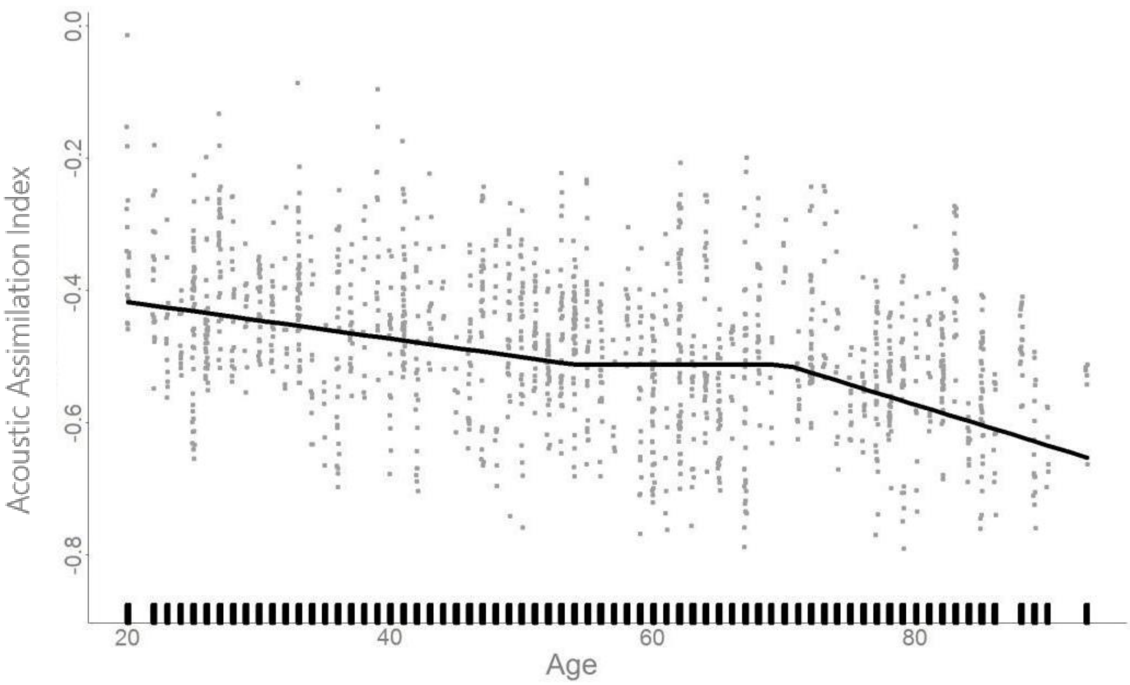

Figure 20. Age study. Relationship between age (x axis) and the acoustic assimilation index per token (y axis). The higher the index, the more there is coarticulation. The black line represents the hinge function of the MARS regression, that found two knots at age 54 and 70.

Between the knots found at age 54 and 70, there is a large dispersion of the *assimilation indices* with several interactions, which will be further discussed below.

Figure 21 presents the *assimilation indices* according to speakers' articulation rate. Again, the relationship is non-linear and much less continuous than the one found with age. Indeed, coarticulation is found to increase with articulation rate only for rates faster than 11.08 phoneme/s, where a knot is found ($\beta = 0.03$).

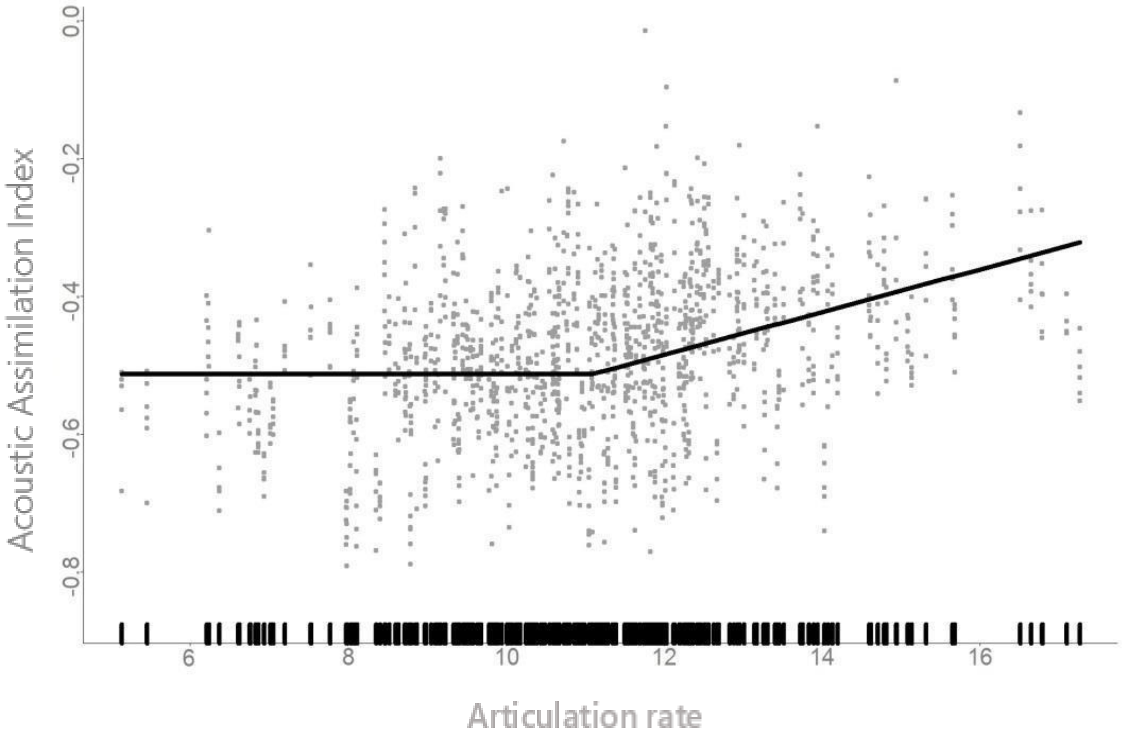

Figure 21. Age study. Relationship between articulation rate (x axis) and the acoustic assimilation index (y axis). The black line represents the hinge function of the MARS model, with a knot at 11.08 phoneme/s.

The interaction between the two predictors age and articulation rate and the *acoustic assimilation index* is illustrated in Figure 22.

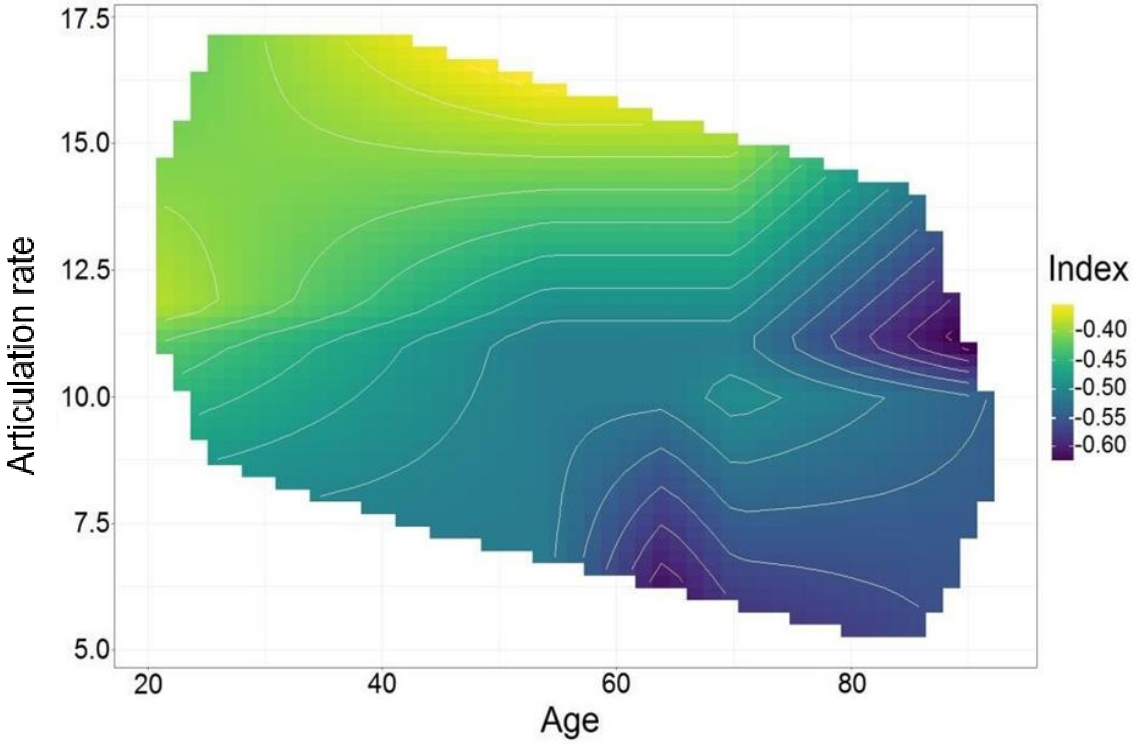

Figure 22. Age study. Interaction between age (x-axis) and articulation rate (y-axis) in the prediction of the acoustic assimilation index, in color: the more the acoustic assimilation index approaches 0 (lighter colors), the more there is coarticulation).

As in the previous figures, we see that until around the mid 50s speakers show a large amount of coarticulation (light colors and *acoustic assimilation index* approaching to 0), which slightly reduces (darker colors) for the speakers with the slowest rates (below about 11 ph/sec). For speakers above approximately 70 y.o.a., coarticulation is low (dark colors) and does not seem to depend much on speech rate. Indeed, we can see a small cluster of speakers in the middle right part of the figure which shows a very small coarticulation index and a rate around 11 ph/sec., while other speakers with a lower rate have a slightly bigger coarticulation index. Above 70 y.o.a., coarticulation seems to increase with rate only for the ones who speak the fastest (e.g., above 13 ph/s.). Nevertheless, for these older speakers, coarticulation is lower than that of younger speakers at the same rate. In the middle part of the figure, for speakers between 54 and 70 y.o.a., a wide range of articulation rates and a clear covariation between rate and coarticulation is found: the slower the speaker speaks, the less coarticulation is found. In particular, a very low coarticulation index is shown by speakers who present a rate lower than 10 ph/s.

6.4 Summary of results

This study investigates the relationship between anticipatory V-to-V coarticulation, articulation rate and speakers' age in a population spanning 20 to 93 years old. Coarticulation is investigated as the *acoustic assimilation* of V_1 [a]i to V_2 /i/ in *papi.* Since the relationship between coarticulation, rate and age does not have to be linear, it is explored through a non-linear MARS regression analysis, which allows individuating knots. These knots represent significant ages that act as turning points for the changes in coarticulation and rate and for how they interact. Though the precise ages individuated are dependent on the speakers' pool, some general tendencies can be drawn out, and can be useful in the light of studies that carry on age group comparisons.

As expected, speech is found to slow down with age. The results show a continuous decrease of speech rate with age, with a steeper decrease for speakers after middle age.

The main finding of this study is that coarticulation also varies with age in adult speech, and that this variation is not necessarily related to a variation in articulation rate. Indeed, the amount of anticipatory V-to-V coarticulation, measured as *acoustic assimilation*, is found to continuously decrease with increasing age. Moreover, the results show that the decrease in *acoustic assimilation* degree is not linear. Three stages can be roughly identified: coarticulation decreases gradually with age for speakers between 20 and middle-age, and then drops steeply with age for speakers older than 70. For speakers in between (50s-60s), a large diversity of assimilation indices is found, which could be either due to speakers' and/or tokens' specific patterns.

If coarticulation and speech rate tendentially covary, this relationship is not the same at all ages. Until the mid 50s, speakers present generally higher speech rates and there is a strong covariation of rate and coarticulation, with faster speakers (faster than a "threshold" rate found at about 12 ph/s.) coarticulating more than their slower peers, even though there are some exceptions to this covariation. From the mid 50s to 70 y.o.a, articulation rate overall lowers, and speakers present a wide range of coarticulation profiles, from low to high coarticulation degrees. For these speakers, coarticulation also covaries with rate, but the threshold rate over which coarticulation approaches that of younger speakers is set higher (around 14 ph/s). At rates lower than this threshold, middle-aged speakers always present overall less coarticulation than younger ones at the same rate. A more substantial change in the relationship between rate and coarticulation is observable for speakers older than 70. They show a globally lower coarticulation degree, rates lower than 14 ph/s. and, crucially, no variation in the degree of coarticulation according to rate, except for the (few) very fastest speakers.

7 Individual variations in V-to-V coarticulation according to changes in speech tempo

7.1 Goals of the study

In this study, individual variations in the patterns of V-to-V coarticulation of five speakers are investigated across three tempo conditions: a comfortable self-paced tempo, which serves as baseline for each speaker, a slow tempo and a fast tempo.

The main goal of this study is to further investigate the relationship between V-to-V coarticulation and rate. In the preceding study on age (**age study**, chapter 6) this matter has been investigated in a population that was varied in terms of articulation rate, and the results showed a fair amount of interspeaker variation in the relationship between coarticulation degree and rate. This study is aimed at isolating the effect of rate on coarticulation by eliciting a decrease and an increase in speech tempo in the production of speakers in the same age range.

Interspeaker variations in coarticulation are explicitly investigated. Indeed, speaker-dependent effect of instructed rate changes has been reported in the literature addressing the effects of rate manipulations on V-to-V coarticulation and on kinematics. In this study, patterns of changes in coarticulation under the different tempo conditions are compared across speakers.

The slow and fast tempo conditions were paced with a moving bar on the screen. The rate was set at 1.4 syll/sec for the slow speech, and at 6 syll/sec for the fast speech, in order to create conditions that could be comparable across speakers.

7.2 Methods

7.2.1 Participants

Five female speakers in their twenties (22<>28) participated in the experiment. Four of them where students in speech sciences at the Sorbonne Nouvelle University in Paris, one in the PhD program (Speaker3) and three in the master program (Speaker2, Speaker4 and Speaker5). Speaker1 was recruited outside of the university among friends. All speakers lived in the region of Ile de France at the time of the recordings, but they originally came from different regions in France: Speaker1 and Speaker3 from Ile de France, Speaker2 and Speaker5 from Alsace, Speaker4 from Auvergne-Rhône-Alpes. However, no speaker presented a regional accent.

7.2.2 Speech material and conditions

The target words *papi* and *papa* were each embedded in a sentence:

- *Le* ***papi*** *de Tashri loue des skis de fond.* "Tashri's grandpa rents some cross-country skis"
- *Le ch'ti c'était comme le chinois pour* ***papa****.* "Ch'ti was like Chinese for dad"

Sentences were read by the speakers 20 times in a row, for a total of 20 tokens per target word.

Tempo modifications were obtained by asking the speakers to follow a visual prompt on the screen, i.e. a moving bar:

- Tempo was set at **1.4 syllables** per second for the **slow** tempo condition
- Tempo was set at **6 syllables** per second for the **fast** tempo condition

For the two tempo conditions, speakers were instructed to read the sentence continuously without inserting pauses.

For the baseline condition, no particular reading instructions were given to the speakers.

7.2.3 Recording procedure

The **baseline** and the **fast** condition were recorded by the speakers in the same session. The **slow** condition was recorded in a later session.

Indeed, the experiment was part of a larger protocol that targeted different modifications of speech (Didirkova *et al.,* 2020). Five sentences were recorded by the speakers in different conditions. Besides the instructed tempo changes, other conditions included auditory feedback modifications and modified sensory feedback (biteblock).

The entire protocol was recorded by each speaker in three sessions.

- In Session1, speakers recorded the **baseline**, the modified auditory feedback (cocktail party noise) and the **fast** tempo conditions;
- In Session2, speakers recorded the f0 modifications and delayed auditory feedback conditions;
- in Session3, speakers recorded the bite-block and the **slow** tempo condition.

The baseline and the two tempo modifications were thus not recorded in the same session.

Since the protocol was long and tiring for the speakers, the three sessions were made in different days or different moments of the same day (morning and afternoon). The participants were either workers or students, and the schedule of the recording sessions depended on their availability. The time lag between the sessions was:

- Speaker1:
    - Session1-Session2 (1 day)
    - Session2-Session3 (morning-afternoon)
- Speaker2:
    - Session1-Session2 (1 day)
    - Session2-Session3 (4 days)

- Speaker3:
  - Session1-Session2 (1 day)
  - Session2-Session3 (4 days)
- Speaker4:
  - Session1-Session2 (1 day)
  - Session2-Session3 (2 days)
- Speaker5:
  - Session1-Session2 (2 days)
  - Session2-Session3 (1 day)

The recordings were made by Ivana Didirkova, who was a postdoc at the LPP at the time. For a thorough description of the other tasks and first general results see Didirkova, Lancia, and Fougeron (2020).

7.2.4 Acoustic measures

7.2.4.1 V-to-V coarticulation

Coarticulation is measured as the contextual difference in *F2-F1 compacity* between [a]i of *papi* and [a]a of *papa,* and as the acoustic assimilation of [a]i to /i/ in *papi* as captured by the *acoustic assimilation index.* The two measures are described in section 4.5.

7.2.4.2 Articulation rate

For each sentence, articulation rate is manually calculated as the number of expected phonemes over sentence duration, minus pause duration.

Articulation rate is measured in order to assess whether speakers have successfully complied to the task, decreasing their rate in the **slow** tempo condition and increasing their rate in the **fast** tempo condition.

7.2.5 Statistical analyses

Two analyses are run. To test whether there is coarticulation, measured as the contextual difference between [a]a and [a]i in all conditions and for all speakers, and whether the magnitude of this difference varies across instructed tempo for each speaker, a first model with *F2-F1 compacity* of V1 /a/ as dependent variable and **context** (V2 /a/ vs /i/), **condition** and **speaker** as fixed factors, in interaction, was built. Random intercepts for repetition (1-20) and random slopes for all the three fixed factors were modeled. Random slopes and intercept were de-correlated to avoid convergence warnings. The fitted linear regression was:

***F2-F1 compacity*~ Context*Condition*Speaker + (Context + Condition + Speaker||Repetition)**

To test whether coarticulation, measured with the *acoustic assimilation index*, varies in degree depending on speaker and instructed tempo, a second model with *acoustic assimilation index* as dependent variable and **condition** and **speaker** as fixed factors, in interaction, was built. Random intercepts were modeled for repetition (1-20) and random slopes for **condition** and **speaker**. Random slopes and intercept were de-correlated due to convergence issues. The fitted linear regression was:

Assimilation Index ~ Condition*Speaker + (Condition+Speaker||Repetition)

In the results section, the two analyses will be presented in the same order as this section.

Before presenting the results on the coarticulation measures, articulation rate will be presented for each speaker and condition in order to assess how speakers have responded to the task.

7.3 Results

7.3.1 Response to the task: articulation rate

Articulation rate is presented to assess whether speakers have successfully complied to the task, decreasing and increasing their speech rate, and to assess their rate in the baseline condition. Since the tempo imposed in the **slow** and the **fast** condition was the same for all speakers, depending on their habitual rate speakers could have decreased or increased rate to different extents.

In Figure 23, articulation rate is shown for each speaker and condition, for the two target sentences.

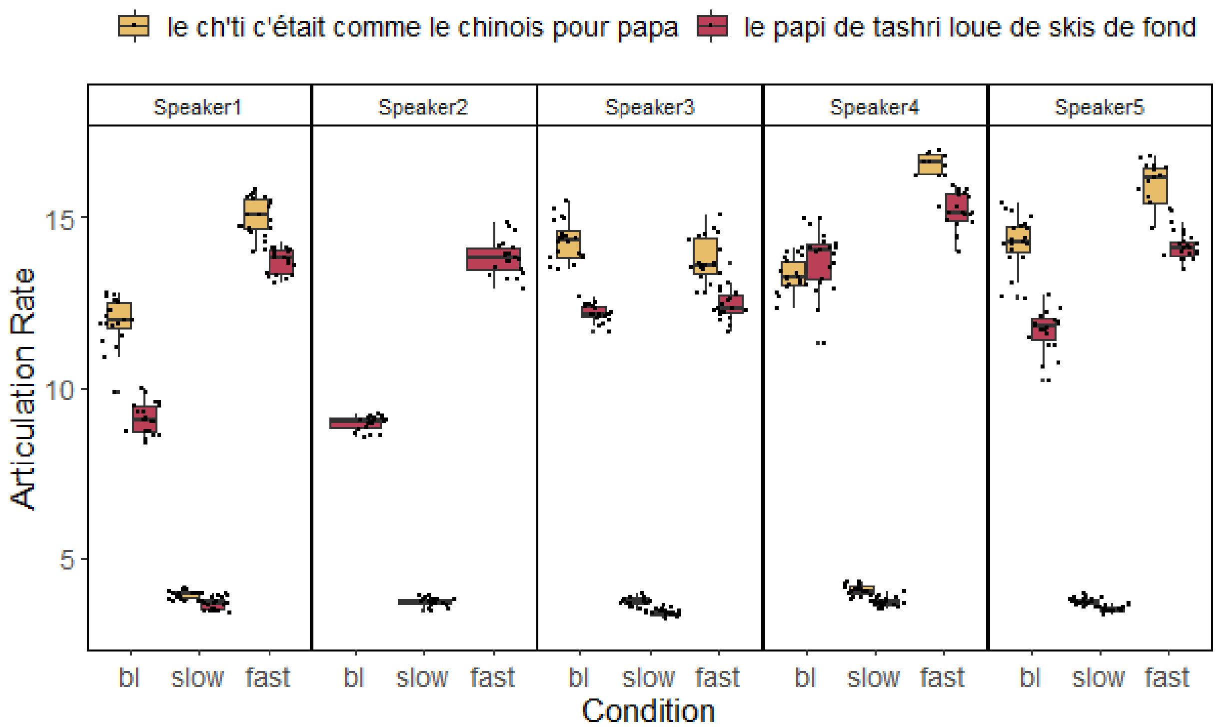

Figure 23. Tempo study. Articulation rate for each speaker and condition for the target sentences for papi (in red) and papa (in yellow).

It can be noticed that articulation rate could not be calculated for Speaker2 in the *papa* sentence, due to missing data. Since she decreases her rate at slow tempo and increases her rate at fast tempo for the *papi* sentence (and the other sentences of the corpus), it can be considered that she also complies to the task for the *papa* sentence.

As expected, articulation rate is speaker dependent in the baseline condition, with faster and slower speakers, e.g. Speaker1 vs Speaker5. It can be noticed that, for all speakers but Speaker4 in the baseline condition, the *papa* sentence is faster than the *papi* sentence in each condition

All speakers considerably slow down their rate in the **slow** condition, adopting all similar rates approximately spanning from 3.40 ph/s to 4 ph/.

In the **fast** condition, four speakers out of five increase their rate. Speaker3 does not increase her articulation rate in this condition. However, for the speakers who increase their rate, the **fast** condition differs less from the baseline than the **slow** condition, and there are some differences in the rates adopted by the speakers.

7.3.2 V-to-V coarticulation

7.3.2.1 Contextual difference index

Figure 24 displays, for each speaker, *F2-F1 compacity* values of [a]i compared to [a]a in the baseline, slow and fast condition for each speaker. The higher the *compacity* values of [a]i with respect to [a]a, the more there is coarticulation.

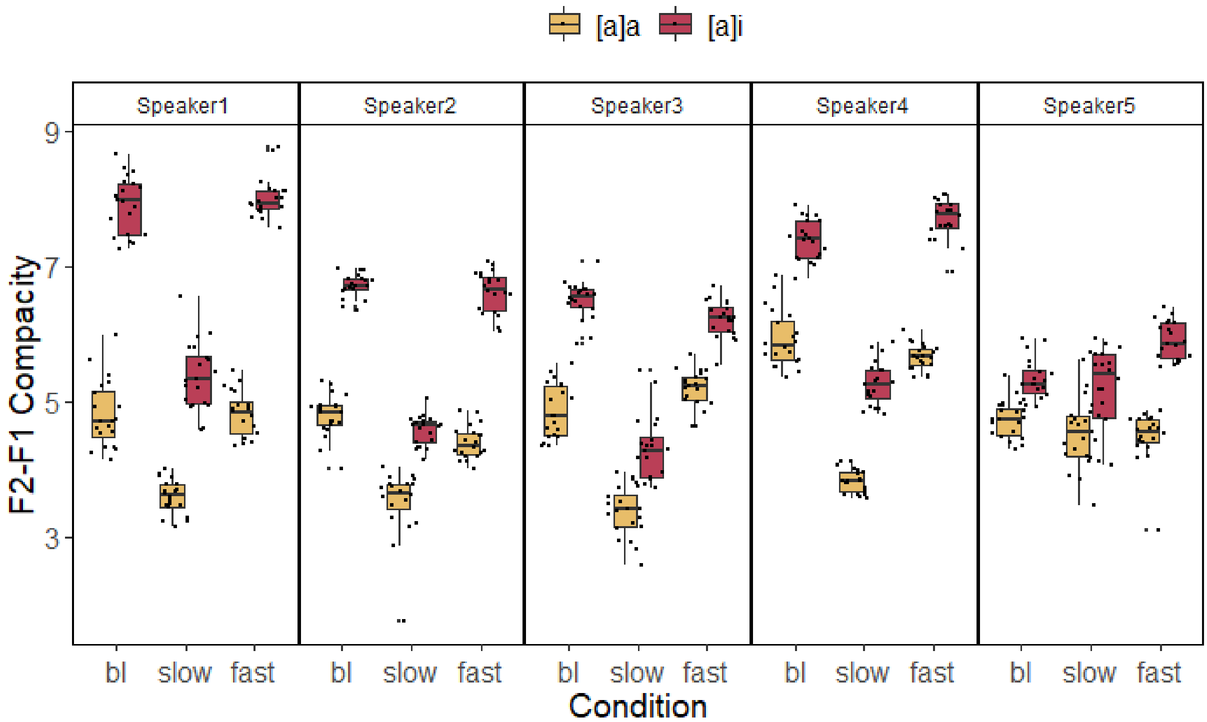

Figure 24. Tempo study. F2-F1 compacity values of [a]i (in red) and [a]a (in yellow) in the baseline ((bl), slow and fast tempo conditions for each speaker. The higher the compacity values of [a]i with respect to [a]a, the greater the coarticulation.

The R^2 of the fitted model, a summary of the fixed effects and interactions, and the results of the pairwise comparisons are given in Table 8.

F2-F1 compacity of V1 /a/ is affected by the nature of V_2, but this effect interacts with **speaker** and **condition**. The *F2-F1 compacity* values are higher for [a]i than [a]a. However, the extent of this contextual difference varies according on condition (**baseline**, **slow**, **fast**), with a speaker-dependent pattern.

The pairwise comparisons show that every speaker presents an idiosyncratic degree of contextual difference in the baseline.

In the **slow** condition, this "baseline" degree of contextual difference is reduced for three speakers out of five (Speaker1, 2 and 3). Speaker4 and Speaker5 present the same contextual difference in the two conditions. Moreover, in this condition both [a]i and [a]a present a more "compact" realization, meaning that they are more open with respect to the baseline, for four out of five speakers (Speaker 5 present the same compacities in the two conditions for the two contexts).

In the **fast** condition, only two speakers present an increase in the contextual difference between [a]i and [a]a with respect to the baseline, Speaker5 and Speaker4. Of them, Speaker5 is the one who shows the greatest increase in coarticulation. Speaker1 and Speaker2 present the same degree of contextual difference in this condition as for the baseline. Speaker3 present less difference in this condition, but note that she does not increase articulation rate.

***F2-F1 compacity~* Context*Condition*Speaker + (Context + Condition + Speaker||Repetition)** (R2m=0.92, R^2c=0.93)	
Summary of the fixed effects	
Context	$\chi^2(1)$= 98.7, *p*=<.0001
Condition	$\chi^2(2)$= 91.2 ,*p*=<.0001
Speaker	$\chi^2(4)$= 177, *p*=<.0001
Context*Condition	$\chi^2(2)$= 47.4, *p*=<.0001
Context*Speaker	$\chi^2(4)$= 174.3, *p*=<.0001
Condition*Speaker	$\chi^2(8)$= 335.1, *p*=<.0001
Context*Condition*Speaker	$\chi^2(8)$= 85.4, *p*=<.0001
Pairwise comparisons [a]i vs [a]a	

Speaker1	**baseline**	β=3.1, SE=0.11, *p*=<.0001
	slow	β= 1.8, SE=0.11, *p*=<.0001
	fast	β = 3.2, SE=0.11, *p*=<.0001
Speaker2	**baseline**	β=1.9, SE=0.11, *p*=<.0001
	slow	β= 1.1, SE=0.11, *p*=<.0001
	fast	β=2.2, SE=0.11, *p*=<.0001
Speaker3	**baseline**	β=1.6, SE=0.11, *p*=<.0001
	slow	β= 0.9, SE=0.11, *p*=<.0001
	fast	β=1.0, SE=0.11, *p*=<.0001
Speaker4	**baseline**	β= 1.5, SE=0.11, *p*=<.0001
	slow	β= 1.4, SE=0.11, *p*=<.0001
	fast	β= 2.0, SE=0.11, *p*=<.0001
Speaker5	**baseline**	β= 0.6, SE=0.11, *p*=<.0001
	slow	β= 0.7 , SE=0.11, *p*=<.0001
	fast	β= 1.4, SE=0.11, *p*=<.0001

Table 8. Tempo study. Effect of context, condition and speaker on F2-F1 compacity of V1 /a/. For the fixed effects, χ2, degrees of freedom, and p values are reported. Coefficient estimates, standard errors, and p values are reported for the pairwise comparisons.

7.3.2.2 Acoustic assimilation index

Figure 25 displays, for each speaker, the *acoustic assimilation index* in the three conditions.

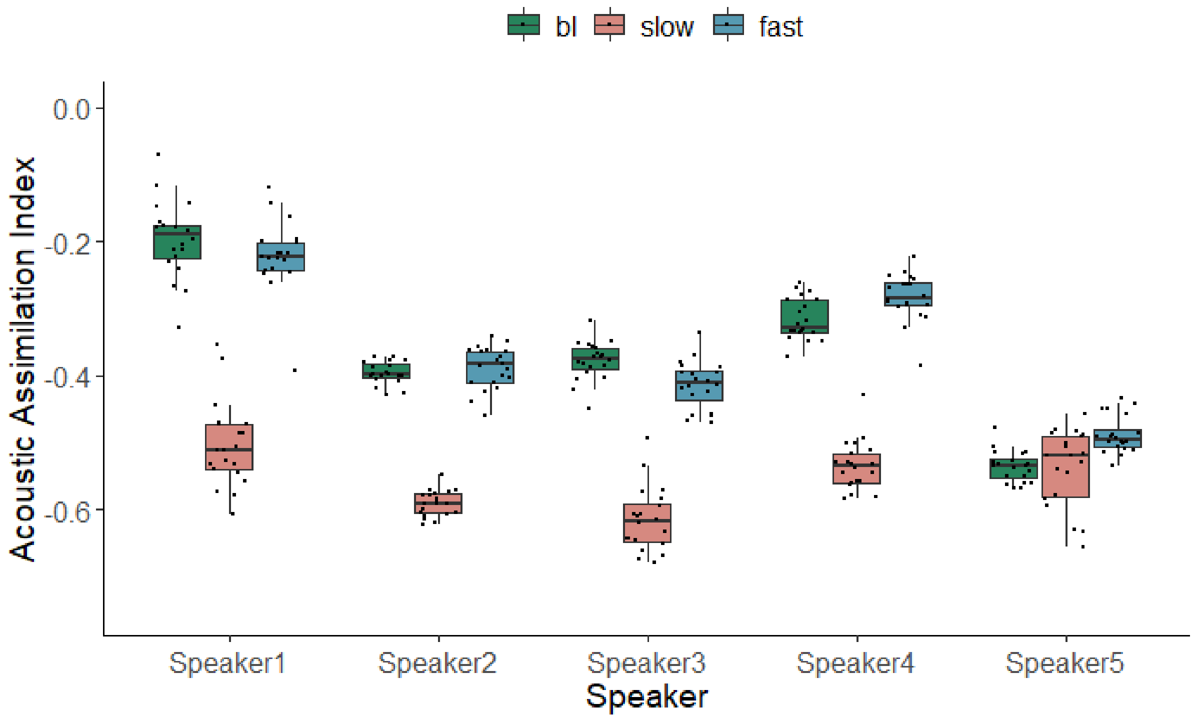

Figure 25. Tempo study. Acoustic assimilation index per speaker and condition, baseline (bl), slow, fast. The higher the index is, the more there is acoustic assimilation of [a]i to /i/.

The R^2 of the fitted model, a summary of the fixed effects and interactions, and the results of the pairwise comparisons are given in Table 9.

The effect of **condition** on the *acoustic assimilation index* is modulated by **speaker**.

Pairwise comparisons show that, in the **slow** condition, the degree of the *acoustic assimilation* of [a]i to /i/ is significantly reduced for four speakers out of five with respect to the **baseline**. Speaker5 does not reduce the *assimilation index*, but present an increase in the variability of the index in this condition.

In the **fast** condition, two speakers significantly increase *acoustic assimilation* degree with respect to the baseline, Speaker4 and Speaker5. However, the effects of **fast** tempo on coarticulation, when significant, are somewhat weaker in comparison to the effects of **slow** tempo.

Speaker3 shows lower *acoustic assimilation index* in the **fast** than in the **baseline** condition. Note that this speaker does not increase rate in this condition.

Assimilation Index ~ Condition*Speaker + (Condition+Speaker||Repetition) (R^2m=0.91, R^2c=0.91)		
Summary of the fixed effects		
Condition	$\chi^2(2)$= 85.5, *p*=<.0001	
Speaker	$\chi^2(4)$= 227.9, *p*=<.0001	
Condition*Speaker	$\chi^2(8)$= 254.3, *p*=<.0001	
Pairwise comparisons **Slow tempo** *vs* **baseline** and **fast tempo** *vs* **baseline**		
Speaker1	**Slow** vs baseline	β= -0.31, SE=0.01, *p=<.0001*
	Fast vs baseline	ns
Speaker2	**Slow** vs baseline	β= -0.19, SE=0.01, *p=<.0001*
	Fast vs baseline	ns
Speaker3	**Slow** vs baseline	β= -0.24, SE=0.01, *p=<.0001*
	Fast vs baseline	β= -0.036, SE=0.01, *p*= 0.02
Speaker4	**Slow** vs baseline	β= -0.22, SE=0.013, *p=<.0001*
	Fast vs baseline	β= 0.03, SE=0.01, *p*= 0.025
Speaker5	**Slow** vs baseline	ns
	Fast vs baseline	β= 0.05, *p*=0.0005

Table 9. Tempo study. Effect of condition and speaker on the acoustic assimilation index. For the fixed effects, χ2, degrees of freedom, and p values are reported. Coefficient estimates, standard errors, and p values are reported for the pairwise comparisons.

7.4 Summary of results

In this study, individual patterns of V-to-V coarticulation in five speakers are investigated across an instructed slow and fast speech tempo, compared to a comfortable self-paced tempo. Coarticulation is measured as the contextual difference between the *F2-F1 compacity* of [a]i and [a]a, and as the acoustic assimilation of V_1 [a]i to V_2 /i/ in *papi* as measured by the *acoustic assimilation index*.

The two analyses on coarticulation differ in the individual results of the slow condition. Indeed, the analysis on *F2-F1 compacity* reveals a reduction of coarticulation at slow speech for three speakers, Speaker4 and Speaker5 not presenting a reduction in contextual difference. On the other hand, a reduction of the *acoustic assimilation index* at slow tempo is found for four speakers, only Speaker5 presenting the same degree of *acoustic assimilation* in the baseline and at slow tempo. Data observation show that for Speaker4, the results on *acoustic assimilation* reflect the production of a more open /a/ vowel, which stays more distinct from /i/. However, this speaker still presents a clear contextual difference between the two vowel targets [a]i and [a]a in this condition on F2, as revealed by the formant observations.

At fast tempo, two speakers out of the four who increase rate in this condition show a greater coarticulation degree. Indeed, only Speaker4 and Speaker5 present an increase of the contextual difference between the ***compacity*** of [a]i and [a]a and a higher ***acoustic assimilation index*** in the fast condition. Speaker1 and Speaker2, albeit they increase rate in the fast tempo condition, present the same coarticulation degree as the baseline. Speaker3 does not increase her articulation rate in this condition, not successfully complying to the task.

8 Individual variations in V-to-V coarticulation according to boundary type

8.1 Goals of the study

In this study, inter and intraspeaker variations in V-to-V coarticulation are investigated across three different boundary conditions. The influence of V_2 on V_1 is examined across syllable boundary within a word, across word boundary and across clause boundary in 45 repetitions per condition per speaker.

The first goal of this study is to investigate the nature of units over which the movements for the achievement of a speech target can be planned, that is, the domain of planning of coarticulation. To this end, the degree of coarticulation, and the variability of coarticulation, are examined within a word, across a word boundary and across a clause boundary. A higher degree of variability of coarticulation and a lower degree of coarticulation between two syllables belonging to different words or clauses (or rather, a degree of coarticulation depending on boundary strength) could indicate that the two syllables are not planned together. Indeed, if two gestures are not coordinated with each other, more variability can be expected in the degree to which the two gestures overlap with each other, and less coarticulation the further the gestures are from each other.

The second goal of this study is to investigate whether inter and intraspeaker variations in the degree of coarticulation depend on specificities in prosody. Previous studies have shown speaker-dependent effects in coarticulation across boundary types. The possibility that individual variations in coarticulation depending on boundary conditions are to be accounted by interspeaker and intertoken differences in the prosodic phrasing is tested. Two analyses are carried out for this purpose. Interspeaker differences in the prosodic implementation of the word boundary and clause boundary conditions are measured on three acoustic parameters of prosody, fundamental frequency, vowel duration and the distance between the two vowels (used to detect a pause). Intraspeaker (intertoken) variations in boundary strength in the word and clause boundary conditions are measured perceptually as the degree of separation between the two words judged by six listeners.

8.2 Methods

8.2.1 Participants

The same five female participants who participated in the **tempo** study participated in the study. Participants' pool can be thus found in section 7.2.1.

8.2.2 Speech material

The target sequences /papi/ and /papa/ were embedded in sentences where the boundary between the two syllables is manipulated, from syllable boundary within a word, to word boundary, to clause boundary. These three conditions are defined as **within word**, **word boundary** and **clause boundary** conditions.

Forty-five tokens have been recorded by the speakers for each of the target sequences in each condition. The sentences were read by the speakers in five sessions recorded over a variable timespan (see *infra*). In each session, each of the target sequences occurred 9 times. This procedure was followed to obtain a large number of repetitions while avoiding the effect of fatigue and testing the stability of the productions over time.

For each boundary condition, to ensure variation in the corpus, three carrier sentences with the same syntactic structure and number of syllables were created. The part of the sentence that contains the target sequence stays the same, while lexical changes are introduced after the target sequence.

In the next sections, the boundary conditions and the recording procedure will be described.

8.2.2.1 Target sequences and conditions

In this section, target sequences for the boundary conditions are presented and an example of sentence type is given for each condition. The entire list of sentences per condition is given in Appendix B.

- **Within word condition (/papV_2/).** Target sequences correspond to the bisyllabic words ***papi*** and ***papa***, which are the subject of the sentence.
  - Ex. sentence **/papi/**: *Quand **papi** l'aura vu, il nous croira* "when grandpa sees it he will believe us".
  - Ex. sentence **/papa/**: *Quand **papa** s'en va au sport, il baille* "when dad goes to the gym, he yawns".
- **Word boundary condition (/pa#pV_2/).** The two syllables of the target sequences belong to two words, respectively subject and verb of the carrier sentence.
  - Ex. sentence **/pa#pi/**: *Pa**pa pi**lote un hélicoptère* "dad pilots an helicopter".
  - Ex. sentence **/pa#pa/**: *Pa**pa pa**sse par chez toi en velo* "dad passes by your house on his bike".
- **Clause boundary condition (/pa##pV2/).** The two syllables of the target sequences belong to two words, respectively subject and apposition for the [a]i context and subject and participle verb for the [a]a context.
  - Ex. sentence **/pa##pi/**: *Pa**pa, pi**lote à Toulouse, rentre tard* "dad, pilot in Toulouse, comes home late";
  - Ex. sentence **/pa##pa/**: *Pa**pa, pa**ssant par Toulon, l'a vu* "dad, passing by Toulon, saw it".

To elicit inter and intraspeaker variability in the prosodic phrasing of the sentences in the word and clause boundary condition, no reading instructions were given to the speakers.

8.2.2.2 Prosodic phrasing

The sentences, and thus the target sequences, were susceptible to be realized with different prosodic phrasings. We will describe the possible realizations of each sentence by referring to the Autosegmental-Metrical model of French intonation by Jun and Fougeron (2000, 2002).

8.2.2.2.1 Reference for French intonation: Jun and Fougeron's model (2000, 2002)

In this model the basic tonal unit of French is the Accentual Phrase (AP). The AP is characterized by a final rising pitch accent H*, which is associated with the final full syllable of the AP. The syllable associated with H* accent is characterized by final lengthening and by a rise in F0. The AP can present an optional initial accent Hi, which is usually associated with one of the first syllables of the first content word of the AP. The syllable associated with the Hi initial accent can present a weaker F0 rise than the AP final accent, and is not always significantly longer than the other domain medial syllables. Therefore, the basic tonal pattern of the AP is /LHiLH*/, with each tone realized on one syllable. However, depending on speech rate or phrase length, either the initial L or the initial Hi could not surface. In AP shorter than four syllables, the final LH* can be realized on the last syllable of the AP.

The higher level of the prosodic hierarchy is represented by the Intonational Phrase (IP). The IP is marked by a major continuation rise or a final fall, represented as H% or a L% boundary tones, which are realized on the last syllable of the IP. The syllable associated with the boundary tone exhibit final lengthening and is thus longer than the last syllable of an AP. The end of an IP can be marked by a pause.

8.2.2.2.2 Expected phrasing for the target sequences

The possible realizations of the target sequences in the word boundary and clause boundary conditions are listed below (Table 10). The syllables on which the accents can be realized are evidenced in bold.

Since the prosodic analysis was aimed at investigating how differences in phrasing could affect coarticulation, the statistical analyses on the acoustic and perceptual measures of prosodic boundary were carried out only on the [a]i context. However, the analysis of coarticulation as contextual difference between the *compacity* values of V_1 considers also [a]a. For this reason, possible differences in the prosodic realization of the two contexts in each condition were controlled for. Therefore, in Table 10 the two contexts are presented.

Condition	Target sequence	Possible realizations
Word boundary	/pa#pi/	[L Hi L H*] pa**pa** pi**lot**e
		[L H*] [L H*] pa**pa** pi**lot**e
	/pa#pa/	[L Hi LH*] Pa**pa** **pass**e
		[L L H*] Pa**pa** **pass**e
Clause boundary	/pa##pi/	{L H%}{[L H*]…} Pa**pa,** pi**lot**e
		{L H%}{[Hi L…]} Pa**pa**, **pi**lote
	/pa##pa/	{L H%]}{[L H*]…} Pa**pa**, pas**san**t
		{L H%}{[Hi L..]} Pa**pa**, **pas**sant

Table 10. Boundary study. Expected prosodic phrasing for the target phrases of the word boundary and clause boundary conditions.

In particular, variation in phrasing was expected in the word boundary condition, where the target phrases were susceptible to be realized as one or two APs.

In the clause boundary condition, the realization of an IP boundary was expected, but speakers could produce or not produce a pause after it.

8.2.3 Recording procedure

8.2.3.1 Recording sessions

The five sessions were recorded by the speakers over several days. The time lag between the sessions was not controlled for and went from some hours to a month. For Speakers 1, 3, 4 and 5 the recordings were carried out over approximately one of two working weeks. Speakers were either workers or students, and the scheduling of the sessions depended on their schedule (as described for the **tempo** study in section 7.2.3). The lag between sessions was much longer for Speaker2. Indeed, these recordings were taken during the confinement, and difficulties related to the circumstances delayed the recordings.

The effect of session was tested on coarticulation to control (and articulation rate) for learning effects and for the differences in the lags between recording among the speakers. No effect of session was found.

The time lag between sessions for each speaker was:

- Speaker1:
    - Session1-Session2: 1 day
    - Session2-Session3: 2 days
    - Session3-Session4: morning-afternoon
    - Session4-Session5: 1 day
- Speaker2:
    - Session1-Session2: 2 weeks
    - Session2-Session3: 1 month
    - Session3-Session4: 1 day
    - Session4-Session5: eight days
- Speaker3:

    - Session1-Session2: 2 days
    - Session2-Session3: 4 days
    - Session3-Session4: 1 day
    - Session4-Session5: 2 days
- Speaker4:
    - Session1-Session2: 1 day
    - Session2-Session3: 2 days
    - Session3-Session4: 3 days
    - Session4-Session5: 1 day

8.2.3.2 Reading task

The sentences were presented on the screen through a slideshow in which each slide contained a sentence, with other sentences of the corpus which are not presented here. After the participant had read each sentence, the experimenter manually passed to the following slide. Each recording session was divided in two blocks, and participants were offered to take a break between blocks. Each session took from 30 to 40 minutes.

The sentences were pseudo-randomized to have a different order in each session, but the order was the same for all speakers.

8.2.4 Perception experiment for the determination of the prosodic boundary strength

A perception experiment was conducted in order to measure the degree of prosodic boundary strength of speakers' productions of *papa pilote* in the word boundary and clause boundary conditions.

8.2.4.1 Stimuli

All the occurrences of *papa pilote* were extracted from the productions of the sentences of the word and clause boundary conditions by all speakers with a Praat script (Garcia, 2015, modified by the author). A silence of 20 milliseconds were added thanks to a second Praat script (Hirst, 2011, modified by the author).

8.2.4.2 Listeners

Three “internal” and 15 "external" listeners rated the degree of separation between words. The three listeners defined “internal” are the author of this dissertation, her supervisor and another PhD student working on the same project, A. Yildiz. The listeners defined “external” were recruited inside and outside of the university and were both phoneticians of naïve speakers.

Of these 18 listeners (14 F, 4 M), 15 were native French speakers. Two were bilingual, a male kurdish-french speaker and female polish-french speaker. Both speakers had lived in France for more than 40 years and used French as primary language. One was a native Italian speaker, the author, who had been living in France for 5 years. Listener’s age was comprised between 20 and 66 years old (*M=34.2, SD=15.4).*

Details of the listeners’ pool are given in Appendix C.

8.2.4.3 Online perception text

The perception test was built on OpenSesame by A. Yildiz.

Since the goal was to judge intraspeaker variation, a perception test was built for the productions of each speaker.

Productions of *papa pilote* by each speaker were auditory and visually presented in the test in a random order, all conditions (word and clause boundary) confounded, with other corpus items that are not analyzed in this study. Listeners could listen to the sentence up to two times. Then, they had to indicate the degree of separation between the two words on a scale from 0 to 3:

0 (no separation)	1 (weak separation)	2 (strong separation)	3 (very strong separation)

The test lasted on average 20 minutes.

Before the actual experiment, listeners had a test practice where 11 stimuli were presented. Test sentences were recorded by a female French speaker who did not participate in the production experiment.

8.2.4.4 Listeners distribution and score averaging

In order to keep the test from being too long, each "external" listeners evaluated the productions of a single speaker (225 items). The three “internal” listened evaluated the productions of all five speakers (225*5 items). Each speaker was therefore evaluated by three "external" listeners and by the three “internal” listeners. Since the responses of the "external" judges did not differ significantly from those of the “internal” ones ($t = -1.89$, $p=0.06$), the judgment scores were averaged to obtain a score expressing degree of *perceived boundary strength* for each item.

8.2.5 Measures

8.2.5.1 Prosody

8.2.5.1.1 Acoustic measures

Individual differences in the prosodic phrasing of the **/pa#pi/** and **/pa##pi/** target sequences in the word boundary and clause boundary conditions are measured acoustically on *F0* of V_1, *V_1 duration* in ms and *V-to-V lags*, from the offset of V_1 to the onset of V_2, in ms.

After extraction in Hertz in Praat, F0, V_1 duration and V-to-V lags were log-transformed in order to normalize the distribution of the residuals in linear mixed models.

These three cues allow determining whether, in word and clause boundary conditions speakers produce a Hi, a H* or a H% accent. To this end, F0, V1 duration and V-to-V lags of productions in the word boundary and clause boundary conditions are compared to the within word condition, used as baseline, since the vowel is unaccented.

8.2.5.1.2 Score of perceived boundary strength

Intraspeaker (intertoken) differences in the prosodic phrasing of the target phrases of the word boundary and clause boundary conditions are analyzed perceptually with a score of *perceived boundary strength*, obtained with the procedure specified in 8.2.4. Scores spanned from 0 to 3 and were treated as a continuous variable.

8.2.5.2 V-to-V coarticulation

Coarticulation degree and variability are measured for this experiment.

Coarticulation is measured as the contextual difference in *F2-F1 compacity* between [a]i of *papi* and [a]a of *papa,* and as the acoustic assimilation of [a]i to /i/ in *papi* as captured by the *acoustic assimilation index.* The two measures are described in Chapter section 4.5.

Variability in coarticulation degree is measured as the variability of the *acoustic assimilation index* in the productions of one condition by one speaker. The absolute distance of the index value for each token to the mean index value for each condition and speaker is calculated to obtain an *assimilation variability index.*

8.2.6 Statistical analysis

First, it is tested whether speakers differ in the prosodic phrasing of the target sequences of the [a]i context in the three conditions: within word (/papi/) word boundary (pa#pi), and clause boundary (pa##pi). In particular, it is tested whether the word boundary and clause boundary condition differ from the within word condition, used as baseline. Three models are run with *F0* (log), *V1 duration* (log) and *V-to-V lag* (log) as dependent variables (VD), and **condition** and **speaker** as fixed effects in interaction. Random intercepts were modeled for **repetition** (1-45) and random slopes for **condition and speaker**. Random intercept and slopes were decorrelated to avoid convergence issues (as in the other models of this study). The fitted regressions were:

VD ~ Condition*Speaker+ (Condition+Speaker||Repetition)

To test whether there is coarticulation, measured as the *contextual difference* between [a]a and [a]i, in all positions and for all speakers, and whether the magnitude of this difference varies depending on boundary type for each speaker, a model with *F2-F1 compacity* of V_1 /a/ as dependent variable and **context** (V_2 /i/ vs /a/), **condition** and **speaker** as fixed factors, in interaction, was built. Random intercepts for repetition and random slopes for all fixed factors were modeled.

The fitted linear regression was:

***F2-F1 compacity*~ Context*Condition*Speaker + (Context + Condition+ Speaker || Repetition)**

Two models were built to test whether the degree of *acoustic assimilation* of [a]i to /i/, measured by the *acoustic assimilation* index, and the variability in the degree of acoustic assimilation, measured by the *assimilation variability index* vary depending on boundary type and whether the pattern of this variation depends on the speaker. The *acoustic assimilation index* and *assimilation variability index* were modeled as dependent variables (VD) and **condition** and **speaker** as fixed factors, in interaction. Random intercepts for repetition and random slopes for **condition** and **speaker** were modeled. The fitted linear regression were:

VD ~ Condition*Speaker + (Condition+Speaker||Repetition)

Finally, to analyze whether intraspeaker variability in the acoustic assimilation of [a]i to /i/ can be accounted for by the degree of prosodic boundary strength, Pearson's correlations are run between the *perceived boundary strength* and the *acoustic assimilation index* in the word boundary and clause boundary condition separately, for each speaker.

In the next section, the results of the analyses will be presented in the same order as they were presented in this section.

8.3 Results

8.3.1 Interspeaker differences in prosody

8.3.1.1 F0

Figure 26 displays the F0 of V1 [a]i for each speaker in the **within word**, **word boundary** and **clause boundary** conditions.

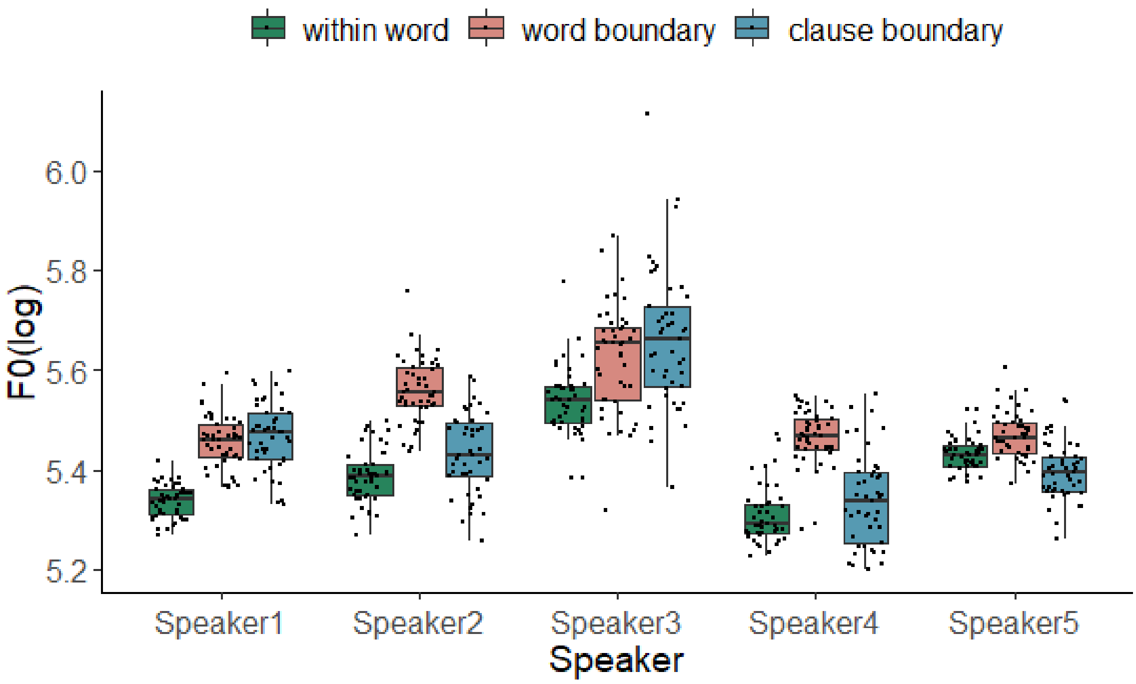

Figure 26. Boundary study. F0 of [a]i for each condition (within word, word boundary and clause boundary) and speaker. F0 in the within word condition is considered the baseline cause the vowel is unaccented.

The **within word** condition can be considered as the baseline, since the vowel is unaccented. A higher F0 in the **word boundary** condition (in *papa pilote*) or in the **clause boundary** condition (in *papa, pilote à…*) can be associated to an accented vowel carrying a high pitch accent H* or boundary tone H%. No difference in F0 or a lower F0 in the **clause boundary** condition with respect to the **within word** condition can be associated to a vowel carrying a L% boundary tone.

The R^2 of the fitted model, a summary of the fixed effects and interactions, and the results of the pairwise comparisons are given in Table 11.

The **condition** has a significant effect on the F0 of [a]i, and this effect interacts with **speaker**.

Pairwise comparisons show that in the **word boundary** condition, all speakers present a higher F0 than the **within word** condition. However, the difference between these conditions is less striking for Speaker 5.

In the **clause boundary** condition, three speakers out of five present a higher F0 than the **within word** condition. Speaker4 does not present a difference in F0 between the **within word** and **clause boundary** conditions. Speaker5 presents a lower F0 in the **clause boundary** than the **within word** condition. Three speakers (Speaker2, Speaker4 and Speaker5) present a lower F0 in **clause boundary** than **word boundary** condition.

The results on F0 can be schematized as follows:

- Speaker1: within word < word boundary = clause boundary
- Speaker2: within word < clause boundary < word boundary
- Speaker3: within word< word boundary < clause boundary
- Speaker4: within word = clause boundary < word boundary
- Speaker5: clause boundary < within word < word boundary

These results will be commented in the summary of results section together with the other results on prosody (section 8.4.1).

F0 (log) ~ Condition*Speaker+ (Condition+Speaker||Repetition) ($R^2m=0.65$, $R^2c=0.74$)	
Summary of fixed effects	
Condition	$\chi^2(2)=152$, p=<.0001
Speaker	$\chi^2(4)=146$, p=<.0001
Condition* Speaker	$\chi^2(8)=173$, p=<.0001

Pairwise comparisons		
Speaker1	**Word boundary** vs **within word**	β= 0.12, SE=0.01, *p*=<.0001
	Clause boundary vs **within word**	β= 0.13, SE=0.01, *p*=<.0001
	Clause boundary vs **word boundary**	ns
Speaker2	**Word boundary** vs **within word**	β= 0.17, SE=0.01, *p*=<.0001
	Clause boundary vs **within word**	β=0.05, SE=0.01, *p*=<.0001
	Clause boundary vs **word boundary**	β= -0.13, SE=0.01, *p*=<.0001
Speaker3	**Word boundary** vs **within word**	β=0.09, SE=0.01, *p*=<.0001
	Clause boundary vs **within word**	β=0.12, SE=0.01, *p*=<.0001
	Clause boundary vs **word boundary**	β= 0.04, SE=0.01, *p*=0.03
Speaker4	**Word boundary** vs **within word**	β=0.16, SE=0.01, *p*=<.0001
	Clause boundary vs **within word**	ns
	Clause boundary vs **word boundary**	β= -0.13, SE=0.01, *p*=<.0001
Speaker5	**Word boundary** vs **within word**	β=0.04, SE=0.01, *p*=<.01
	Clause boundary vs **within word**	β=-0.03 , SE=0.01, *p*=0.04
	Clause boundary vs **word boundary**	β= -0.07, SE=0.01, *p*=<.0001

Table 11. Boundary study. Effect of condition and speaker on F0 of [a]i. For the fixed effects, χ2, degrees of freedom, and p value are reported. Coefficient estimates, standard errors, and p values are reported for the pairwise comparisons.

8.3.1.2 V_1 duration

Figure 27 displays V_1 [a]i duration for each speaker in the **within word**, **word boundary** and **clause boundary** conditions. The **within word** condition can be considered as the baseline because unaccented. A longer vowel in **word boundary** and **clause boundary** conditions compared to **within word** condition can be respectively associated to an AP final accented vowel or an IP final vowel. In the **clause boundary** condition, vowel lengthening is expected, while in **word boundary** condition the duration of the vowel could either be similar to the **within word** condition or be in-between the **within word** and **clause boundary** condition.

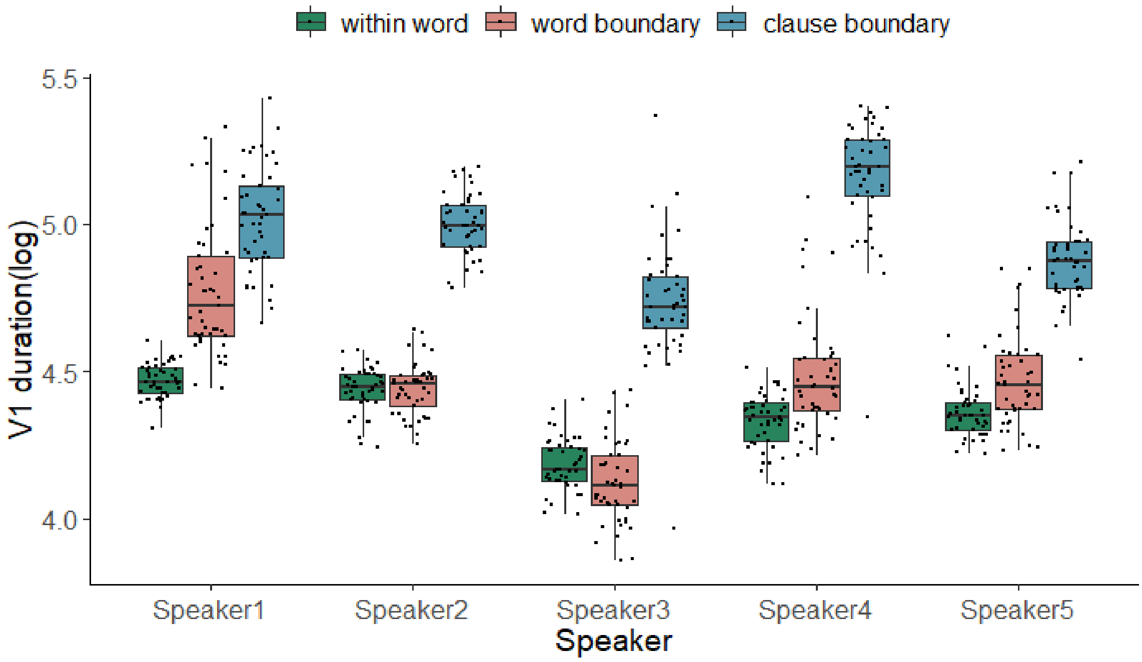

Figure 27. Boundary study. Duration of [a]i for each condition (within word, word boundary, clause boundary) and speaker. V_1 duration in the within word condition is considered as the baseline cause the vowel is unaccented.

The R^2 of the fitted model, a summary of the fixed effects and interactions, and the results of the pairwise comparisons are given in Table 12.

Vowel duration is affected by the **condition**, but this effect is **speaker-**dependent.

Pairwise comparisons show that in the **word boundary** condition, three speakers produce longer vowels than in the **within word** condition (Speaker1, Speaker4 and Speaker5). Speaker2 and Speaker3 present no differences between the **word boundary** and the **within word** condition (Speaker3 exhibits slightly shorter vowels in the **word boundary** condition, but the effect is not significant).

In the **clause boundary** condition, all speakers produce longer vowels compared to the **within word** condition and to the **word boundary** condition.

The results on V_1 duration can be schematized as follows:

- Speaker1: within word < word boundary < clause boundary
- Speaker2: within word = word boundary < clause boundary
- Speaker3: within word = word boundary < clause boundary
- Speaker4: within word < word boundary < clause boundary
- Speaker 5: within word < word boundary < clause boundary

These results will be commented in the summary of results section together with the other results on prosody (section 8.4.1).

V1 duration (log) ~ Condition*Speaker+ (Condition+Speaker||Repetition) (R^2m=0.82, R^2c=0.84)		
Summary of fixed effects		
Condition	$\chi^2(2)$=230, p=<.0001	
Speaker	$\chi^2(4)$=189, p=<.0001	
Condition* Speaker	$\chi^2(8)$=184, p=<.0001	
Pairwise comparisons **Word boundary** vs **within word** and **clause boundary** vs **within word**		
Speaker1	**Word boundary** vs **within word**	β=0.31, SE=0.03, p=<.0001
	Clause boundary vs **within word**	β=0.55, SE=0.03, p=<.0001
	Clause boundary vs **word boundary**	β=0.25, SE=0.03, p=<.0001

Speaker2	**Word boundary** vs **within word**	ns
	Clause boundary vs **within word**	β=0.57, SE=0.03, p=<.0001
	Clause boundary vs **word boundary**	β=0.56, SE=0.03, p=<.0001
Speaker3	**Word boundary** vs **within word**	ns
	Clause boundary vs **within word**	β=0.55, SE=0.03, p=<.0001
	Clause boundary vs **word boundary**	β=0.61, SE=0.03, p=<.0001
Speaker4	**Word boundary** vs **within word**	β=0.16, SE=0.03, p=<.0001
	Clause boundary vs **within word**	β=0.83, SE=0.03, p=<.0001
	Clause boundary vs **word boundary**	β=0.68, SE=0.03, p=<.0001
Speaker5	**Word boundary** vs **within word**	β=0.11, SE=0.03, p=<.0001
	Clause boundary vs **within word**	β=0.52, SE=0.03, p=<.0001
	Clause boundary vs **word boundary**	β=0.41, SE=0.03, p=<.0001

Table 12. Boundary study. Effect of condition and speaker on the duration of [a]i. For the fixed effects, $\chi 2$, degrees of freedom and p value are reported. Coefficient estimates, standard errors, and p values are reported for the pairwise comparisons.

8.3.1.3 V-to-V lag

Figure 28 displays the duration of the V-to-V lag for each speaker in the **within word, word boundary** and **clause boundary** conditions.

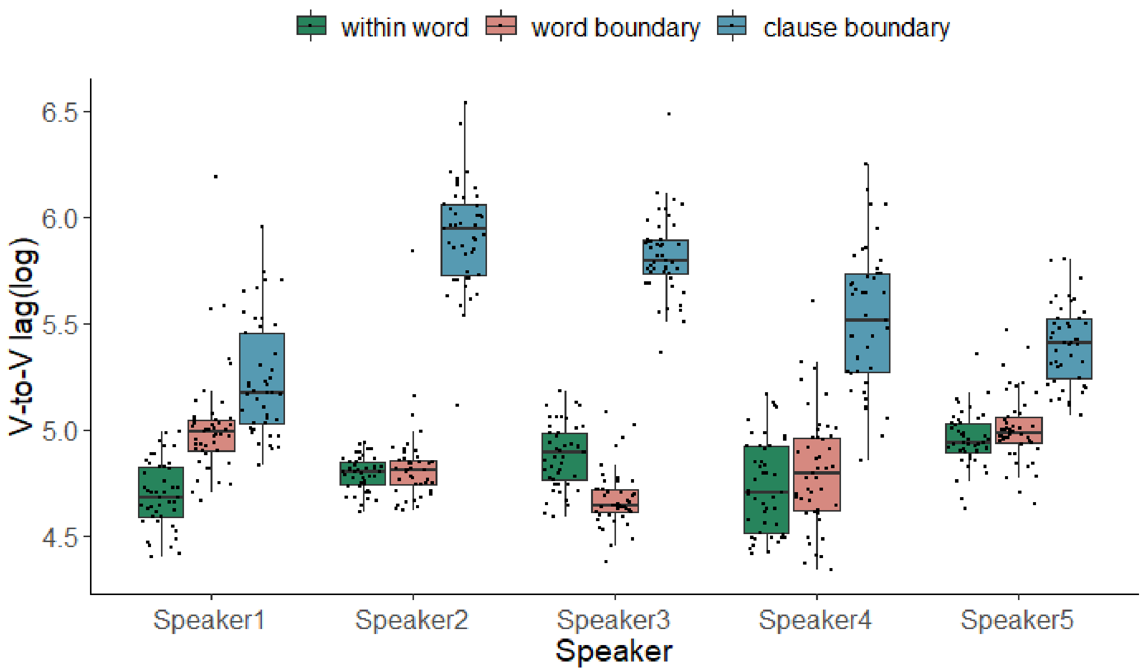

Figure 28. Boundary study. V-to-V lag duration for each condition (within word, word boundary, and clause boundary) and speaker. V-to-V lags in the within word condition correspond to the duration of the intervocalic /p/ and are considered as the baseline.

In the **within word** condition, which can be considered the baseline, the V-to-V lags correspond to the duration of C_2 /p/ in *papa* and *papi.* In **word boundary** and **clause boundary** condition, they correspond to the duration of C_2 /p/ plus a possible pause between *papa* and *pilote.* In the **clause boundary** condition, a pause after the boundary, and thus longer lags, is expected. In the **word boundary** condition, longer lags are not necessarily expected and can be associated to the insertion of a short pause between the noun and the verb.

The R^2 of the fitted model, a summary of the fixed effects and interactions, and the results of the pairwise comparisons are given in Table 13.

The duration of the V-to-V lag is affected by the **condition**, but the effect is **speaker**-dependent.

Pairwise comparisons show that, in the **word boundary** condition, only one speaker produce longer V-to-V lags than in the **within word** condition, Speaker1. Speaker2, Speaker4 and Speaker5 do not present a difference between these conditions. Speaker3 produces shorter V-to-V lags in the **word boundary** condition than the **within word** condition.

In the **clause boundary** condition, all speakers present significantly longer lags than the **within word** and the **word boundary** conditions.

The results on V-to-V lags can be schematized as follows:

- Speaker1: within word < word boundary < clause boundary
- Speaker2: within word = word boundary < clause boundary
- Speaker3: word boundary < within word < clause boundary
- Speaker4: within word = word boundary < clause boundary
- Speaker5: within word = word boundary < clause boundary

These results will be commented in the summary of results section together with the other results on prosody (section 8.4.1).

V-to-V lag (log) ~ Condition*Speaker+ (Condition+Speaker||Repetition) (R^2m=78, R^2c=82)		
Summary of fixed effects		
Condition	$\chi^2(2)$=159, *p*=<.0001	
Speaker	$\chi^2(4)$=50.5, *p*=<.0001	
Condition* Speaker	$\chi^2(8)$=344, *p*=<.0001	
Pairwise comparisons		
Speaker1	**Word boundary** vs **within word**	β=0.33, SE=0.04, *p*=<.0001

	Clause boundary vs **within word**	β=0.54, SE=0.04, *p*=<.0001
	Clause boundary vs **word boundary**	
Speaker2	**Word boundary** vs **within word**	ns
	Clause boundary vs **within word**	β=1.12, SE=0.04, *p*=<.0001
	Clause boundary vs **word boundary**	
Speaker3	**Word boundary** vs **within word**	β=-0.20, SE=0.04, *p*=<.0001
	Clause boundary vs **within word**	β=0.94, SE=0.04, *p*=<.0001
	Clause boundary vs **word boundary**	
Speaker4	**Word boundary** vs **within word**	ns
	Clause boundary vs **within word**	β=0.79, SE=0.04, *p*=<.0001
	Clause boundary vs **word boundary**	
Speaker5	**Word boundary** vs **within word**	ns
	Clause boundary vs **within word**	β=0.45, SE=0.04, *p*=<.0001
	Clause boundary vs **word boundary**	

Table 13. Boundary study. Effect of condition and speaker on V-to-V lags of [a]i. For the fixed effects, χ2, degrees of freedom, and p values are reported. Coefficient estimates, standard errors, and p values are reported for the pairwise comparisons.

8.3.2 V-to-V coarticulation

8.3.2.1 Contextual difference index

Figure 29 displays, for each speaker, *F2-F1 compacity* values of [a]i compared to [a]a in the **within word**, **word boundary**, and **clause boundary** conditions. The higher the *compacity* values of [a]i with respect to [a]a, the more there is coarticulation.

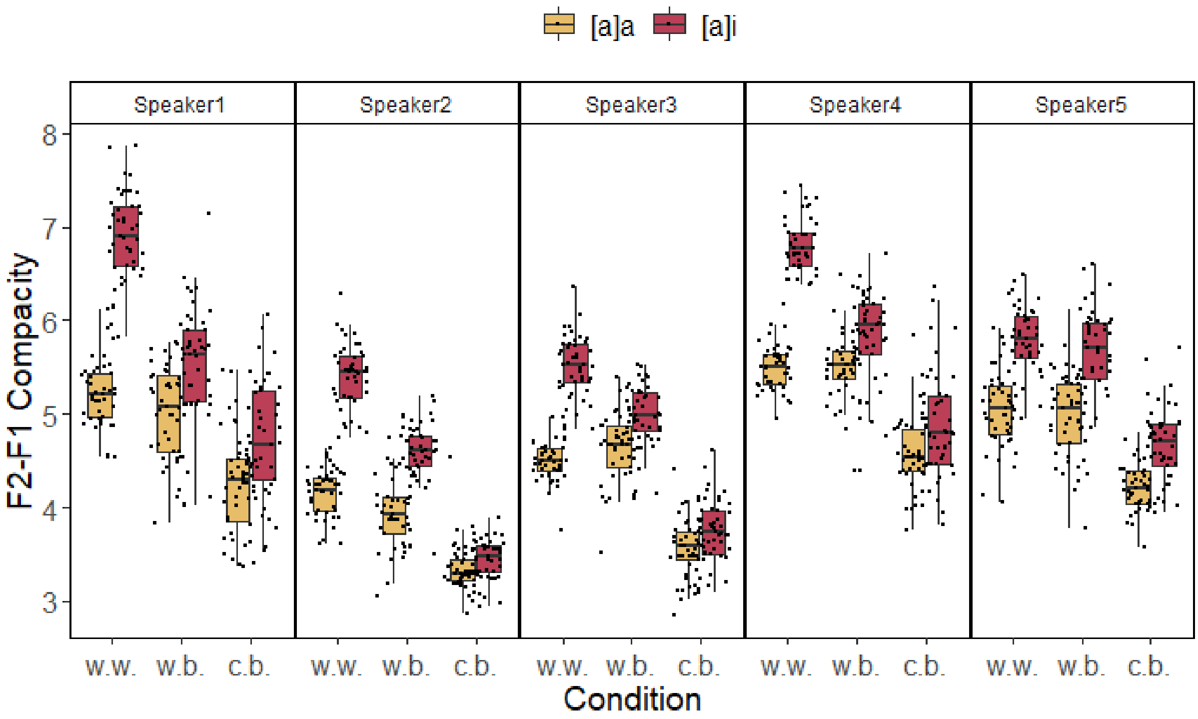

Figure 29. Boundary study. F2-F1 compacity values for [a]i (in red) and [a]a (in yellow), for each condition (w.w.= within word, w.b.=word boundary, c.b.=clause boundary) and speaker. The higher the compacity values of [a]i with respect to [a]a, the more there is coarticulation.

The R^2 of the fitted model, a summary of the fixed effects and interactions, and the results of the pairwise comparisons are given in Table 14.

F2-F1 compacity of V_1 /a/ is significantly influenced by the nature of V_2. That is, *F2-F1 compacity* values are higher for [a]i than [a]a. However, the extent of the difference between the two contexts depends on the **condition**, in interaction with **speaker**.

Pairwise comparisons show that in the **within word** condition speakers coarticulate to different degrees, that is, each speaker present her extent of contextual difference between [a]i and [a]a within a word.

In the **word boundary** condition, all speakers present a significant contextual difference between the *compacity* of [a]i and [a]a. This contextual difference is reduced with respect to the **within word** condition (as showed by the coefficient estimates) for four speakers out of five. Speaker5 presents approximately the same difference between the two contexts in the **word boundary** as for the **within word** condition.

In the **clause boundary** condition, the contextual difference between the *F2-F1 compacity* of [a]i and [a]a is significant for four speakers. For these speakers, the difference between [a]i and [a]a is reduced in this condition with respect to **the within word** and **word boundary** condition. Speaker2 does not present a significant difference between the two contexts in this condition, meaning that she does not present coarticulation.

***F2-F1 compacity*~ Context*Condition*Speaker + (Context + Condition + Speaker||Repetition)** (R2m=0.83, R^2c=0.85)	
Summary of the fixed effects	
Context	$\chi^2(1)$= 143, *p*=<.0001
Condition	$\chi^2(2)$= 258, *p*=<.0001
Speaker	$\chi^2(4)$= 383, *p*=<.0001
Context*Condition	$\chi^2(2)$= 270, *p*=<.0001
Context*Speaker	$\chi^2(4)$= 33.8, *p*=<.0001
Condition*Speaker	$\chi^2(8)$= 120, *p*=<.0001
Context*Condition*Speaker	$\chi^2(8)$= 77.4, *p*=<.0001
Pairwise comparisons [a]i vs [a]a	

Speaker1	**within word**	β= 1.70, SE=0.08, *p*=<.0001
	word boundary	β=0.54, SE=0.08, *p*=<.0001
	clause boundary	β=0.47, SE=0.08, *p*=<.0001
Speaker2	**within word**	β= 1.23, SE=0.08, *p*=<.0001
	word boundary	β=0.67, SE=0.08, *p*=<.0001
	clause boundary	ns
Speaker3	**within word**	β=1.03, SE=0.08, *p*=<.0001
	word boundary	β=0.34, SE=0.08, *p*=<.0001
	clause boundary	β=0.19 SE=0.08, *p*=0.015
Speaker4	**within word**	β=1.31, SE=0.08, *p*=<.0001
	word boundary	β=0.35, SE=0.08, *p*=<.0001
	clause boundary	β=0.30, SE=0.08, *p*=<.0001
Speaker5	**within word**	β= 0.73, SE=0.08, *p*=<.0001
	word boundary	β=0.68, SE=0.08, *p*=<.0001
	clause boundary	β=0.44, SE=0.08, *p*=<.0001

Table 14. Boundary study. Effect of context, condition and speaker on the F2-F1 compacity of V1 /a/. For the fixed effects, χ2, degrees of freedom, and p value are reported. Coefficient estimates, standard errors, and p values are reported for the pairwise comparisons.

8.3.2.2 Acoustic assimilation index

Figure 30 displays the *acoustic assimilation index* for each speaker in the **within word**, **word boundary** and **clause boundary** conditions. The higher the index value, the greater the acoustic assimilation of [a]i to /i/ in *papi*.

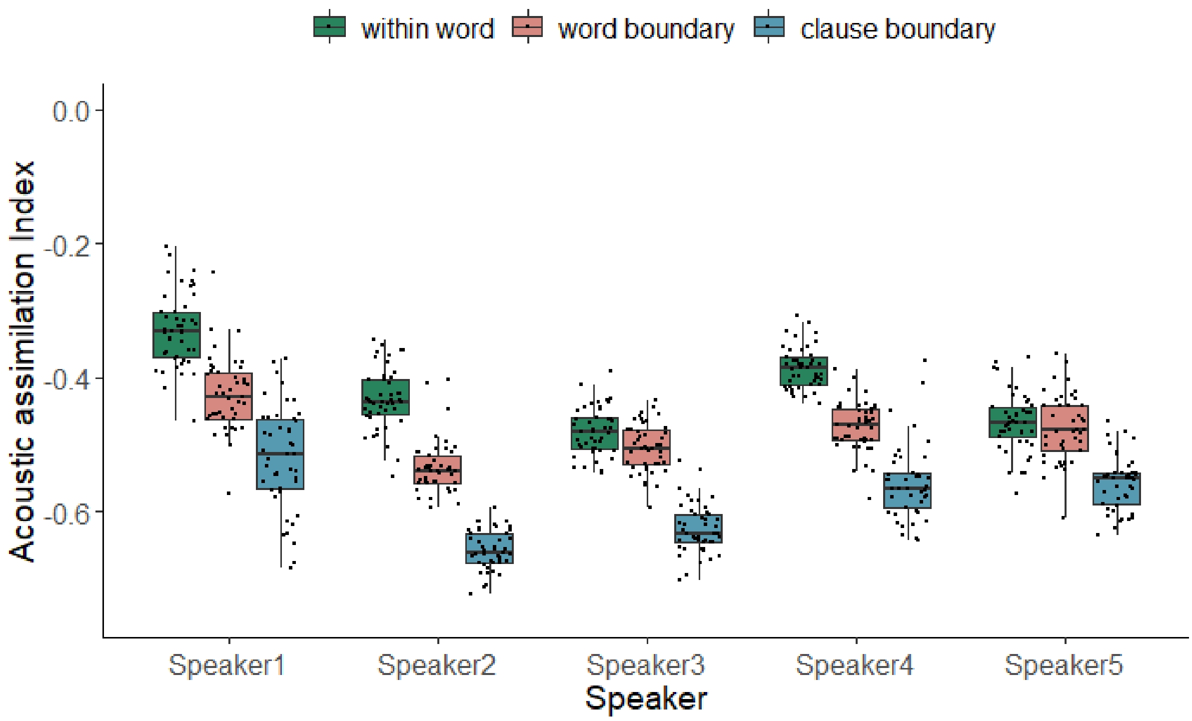

Figure 30. Boundary study. Acoustic assimilation index for each condition (within word, word boundary, and clause boundary) and speaker. The higher the index, the more there is acoustic assimilation of [a]i to /i/.

The R^2 of the fitted model, a summary of the fixed effects and interactions, and the results of the pairwise comparisons are given in Table 15.

Boundary **condition** has a significant effect on the *acoustic assimilation index*, and this effect is modulated by **speaker**.

The pairwise comparisons show that, in the **word boundary** condition, four speakers out of five reduce the acoustic assimilation of [a]i to /i/ with respect to the **within word** condition. There is no significant difference between the *acoustic assimilation indices* of the **within word** and **word boundary** conditions for Speaker5.

In the **clause boundary** condition, all speakers reduce the degree of *acoustic assimilation* of [a]i to /i/ with respect to both **within word** and **word boundary** condition.

It can be noticed that Speaker1, and to a lesser extent Speaker4, present, for several of the repetitions, the same values of the *acoustic assimilation index* in the **clause boundary** and **word boundary** conditions: in other words, in some repetition these speakers present the same coarticulation degree in these two conditions. A similar pattern, but for less repetition and thus less overlap between the conditions, is showed by Speaker5.

Acoustic assimilation index ~ **Condition*Speaker+ (Condition+Speaker||Repetition)** (R^2m=0.76 , R^2c=0.78)		
Summary of fixed effects		
Condition	$\chi^2(2)$=265, *p*=<.0001	
Speaker	$\chi^2(4)$=311, *p*=<.0001	
Condition* Speaker	$\chi^2(8)$=131, *p*=<.001	
Pairwise comparisons		
Speaker1	**Word boundary** vs **within word**	β= -0.09, SE=0.01, *p*=<.0001
	Clause boundary vs **within word**	β= -0.19, SE=0.01, *p*=<.0001
	Clause boundary vs **word boundary**	β= -0.09, SE=0.01, *p*=<.0001
Speaker2	**Word boundary** vs **within word**	β= -0.10, SE=0.01, *p*=<.0001
	Clause boundary vs **within word**	β= -0.23, SE=0.01, *p*=<.0001
	Clause boundary vs **word boundary**	β= -0.12, SE=0.01, *p*=<.0001
Speaker3	**Word boundary** vs **within word**	β= -0.03, SE=0.01, *p*=0.03
	Clause boundary vs **within word**	β= -0.15, SE=0.01, *p*=<.0001
	Clause boundary vs **word boundary**	β= -0.12, SE=0.01, *p*=<.0001
Speaker4	**Word boundary** vs **within word**	β= -0.09, SE=0.01, *p*=<.0001

	Clause boundary vs **within word**	β= -0.17, SE=0.01, *p*=<.0001
	Clause boundary vs **word boundary**	β= -0.08, SE=0.01, *p*=<.0001
Speaker5	**Word boundary** vs **within word**	ns
	Clause boundary vs **within word**	β= -0.09, SE=0.01, *p*=<.0001
	Clause boundary vs **word boundary**	β= -0.08, SE=0.01, *p*=<.0001

Table 15. Boundary study. Effect of condition and speaker on the acoustic assimilation index. For the fixed effects, χ2, degrees of freedom, and p values are reported. Coefficient estimates, standard errors, and p values are reported for the pairwise comparisons.

8.3.2.3 Assimilation variability index

Figure 31 displays the *assimilation variability index* in the **within word**, **word boundary** and **clause boundary** conditions for all speakers. The higher the *assimilation variability index,* the more the *acoustic assimilation index* vary across tokens for each speaker in each condition, thus the more variable coarticulation is in that condition.

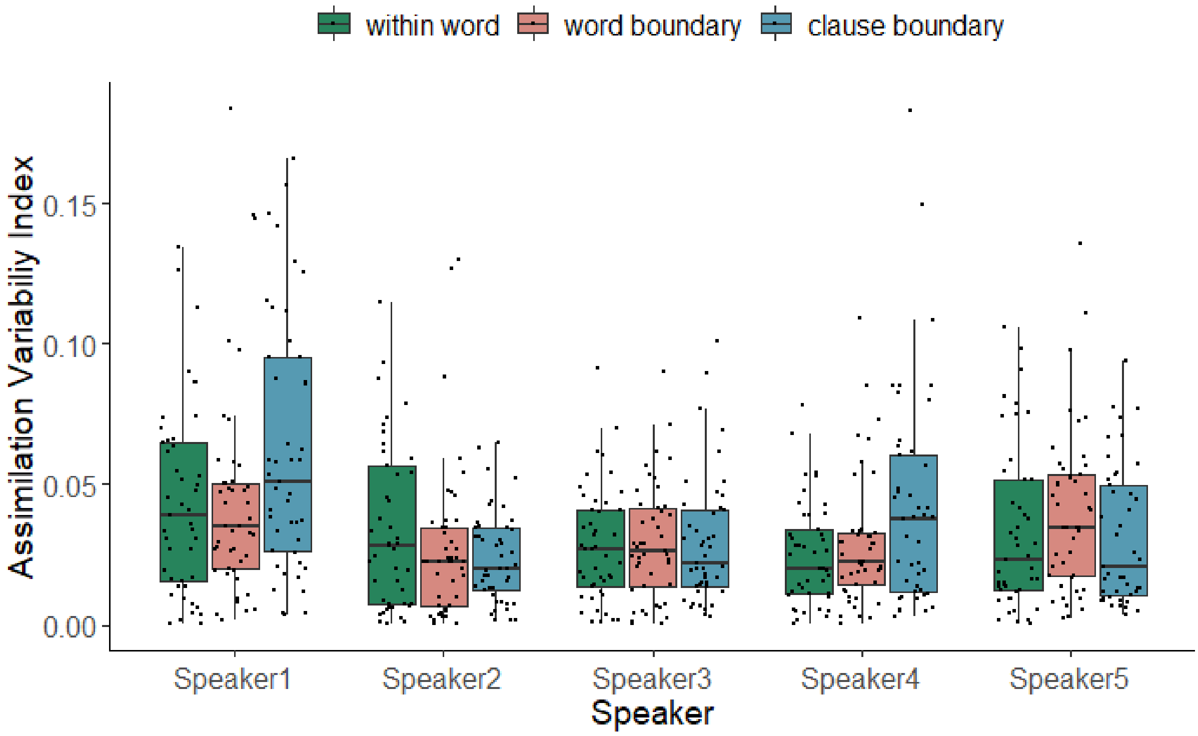

Figure 31. Boundary study. Assimilation variability index for each condition (within word, word boundary, and clause boundary) and speaker.

The R^2 of the model and a summary of the fixed effects are given in Table 16. Since few pairwise comparisons were significant, they are presented in the text.

The *assimilation variability index* significantly varies as a function of **speaker**, in interaction with **condition**. No effect of **condition** is found on the index.

Pairwise comparisons show that only two speakers present a variation in *acoustic assimilation variability* depending on condition, i.e. Speaker1 and Speaker4, for whom the *acoustic assimilation* index is more variable in **clause boundary** condition than the other two conditions.

None of the speakers present a difference in the *assimilation variability index* between the **word boundary** and the **within word** conditions.

Assimilation Variability Index ~ Condition*Speaker+ (Condition+Speaker||Repetition) (R^2m=0.1 , R^2c=0.12)	
Summary of fixed effects	
Condition	ns
Speaker	$\chi^2(4)=48.6$, p=<.0001
Condition*Speaker	$\chi^2(8)=23.1$, p=0.003

Table 16. Boundary study. Effect of condition and speaker on the assimilation variability index: χ2, degrees of freedom, and p values are reported for the fixed effects.

8.3.3 Relationship between the acoustic assimilation index and perceptual boundary strength

Figure 32 displays the relationship between the *perceived boundary strength*, on the x axis, and the *acoustic assimilation index*, on the y axis, for each speaker in **word boundary** (in pink) and **clause boundary** (in blue) condition.

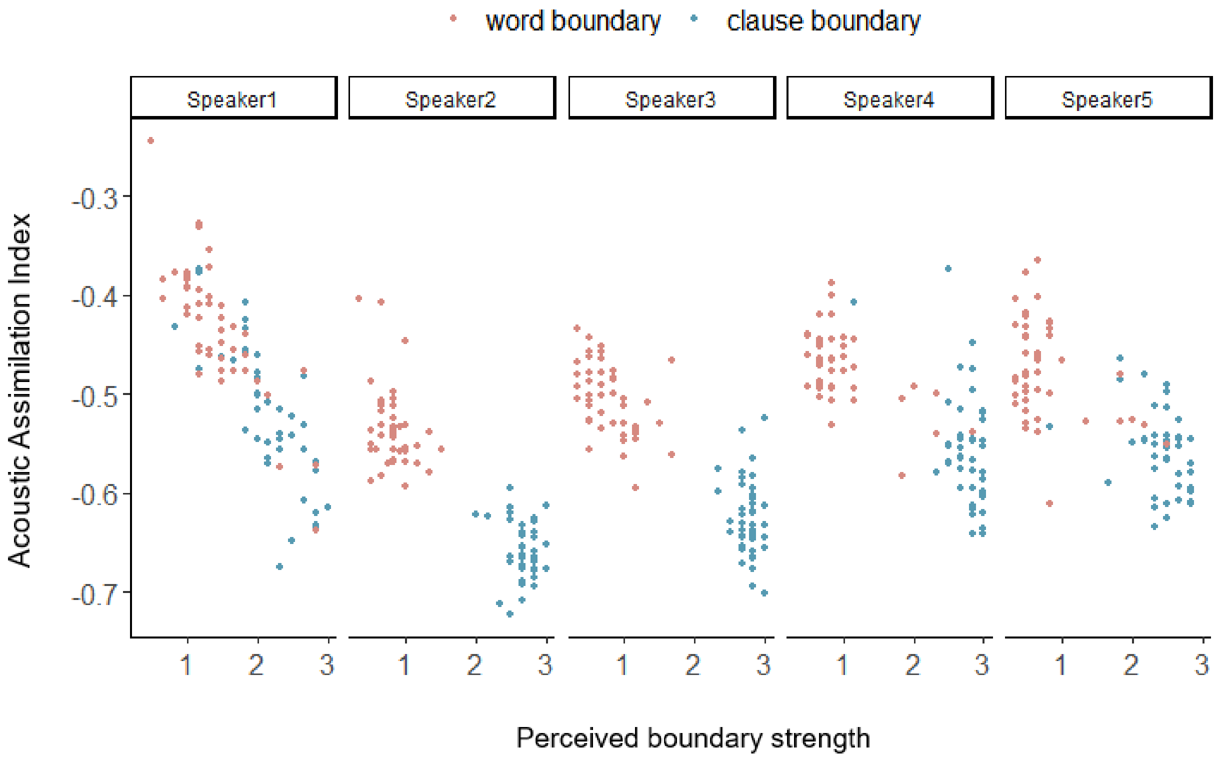

Figure 32. Boundary study. Relationship between the perceived boundary strength and the acoustic assimilation index in the word and clause boundary conditions for all speakers.

For **Speaker1**, in terms of *perceived boundary strength*, no clear distinction can be observed between *pa#pi* sequences in **word boundary** and *pa##pi* sequences **in clause boundary** position. Regardless of boundary type, there is a decrease of the *acoustic assimilation index* at the increase in the *perceived boundary strength*, confirmed by the correlation coefficients (word boundary, r.=.81, clause boundary, r=.8).

Speaker2 presents a weak correlation between the *acoustic assimilation index* and *perceived boundary strength* (word boundary: r = -.30; clause boundary: r = -.15), and a categorical distinction between the two conditions. This pattern is accompanied by an overall low degree of *acoustic assimilation* of [a]i to /i/ for this speaker in these two conditions.

Speaker3 shows a categorical distinction in terms of *perceived boundary strength* between the two positions. In **word boundary** condition, there is a tendency to have a reduction in *acoustic assimilation* with increasing *perceived boundary strength* (r = -.48). In **clause boundary** condition the *acoustic assimilation index* is low and weakly, if not, correlated to *boundary strength* (r=-.17).

Speaker4 shows a moderate correlation between the *acoustic assimilation index* and *perceived boundary strength* in the **word boundary** condition (r=-.54) though this is due mainly to some repetitions with low assimilation and high boundary strength. In the **clause boundary** condition, correlation is also moderate (r=-.49), but this is solely due to one repetition of the condition that presents a high degree of *acoustic assimilation* and low *boundary strength*.

Speaker5 shows a weak to moderate correlation between *acoustic assimilation* and *perceived boundary strength* in the **word** and **clause boundary** condition (respectively, r=-.38, r=-.34) also caused by some repetitions with low assimilation and high boundary strength in the **word boundary** condition and some repetitions with somewhat high assimilation and lower boundary strength.

Speaker4 and **Speaker5** show a similar pattern. On one hand, they produce in the **clause boundary** position (in the upper part of the blue scatterplots) items with a high degree of *acoustic assimilation.* However, these items are still perceived with greater **boundary strength** than the items in the **word boundary** condition that share a similar degree of assimilation. On the other hand, the items in the **word boundary** condition that present the same *boundary strength* as items of the **clause boundary** condition show also a low *acoustic assimilation index*.

8.4 Summary of results

In this study, individual patterns of V-to-V anticipatory coarticulation, and its variability, are investigated in three different conditions, **within word**, across **word boundary** and across **clause boundary**, over forty-five repetition. Coarticulation is measured as the contextual difference between the *F2-F1 compacity* of V_1 /a/ in the two contexts and as the acoustic assimilation of [a]i to /i/. The variability of coarticulation is measured as the *variability of acoustic assimilation* for each speakers and condition. In word and clause boundary position, the relationship between the degree of acoustic assimilation of [a]i to /i/ and the strength of the prosodic boundary is also investigated, the goal being determining whether the intraspeaker variation in coarticulation in these positions is to be attributed to an intertoken variation in the prosodic phrasing. Moreover, an analysis of three acoustic cues of prosody taken in the [a]i context, F0, V1 duration and V-to-V lag, is carried out in order to determine whether interspeaker variation in coarticulation has to be associated to interspeaker specificities in the prosodic phrasing of the word boundary and clause boundary conditions. The results of this last analysis will be summarized first, in order to highlight individual differences in the reading that could eventually influence coarticulation patterns.

8.4.1 Summary of the results on prosody

In particular, interspeaker variations in the prosodic phrasing were expected in the word boundary condition, where the target sequence *papa pilote* could be realized either as a single AP, with a possible Hi initial accent on the second syllable of *papa,* or as two APs, with a final H* pitch accent on the second syllable of *papa*. A H* accent on [a]i could affect coarticulation degree.

In the clause boundary condition, less variations in phrasing where expected among the speakers: a IP boundary was likely to be produced, either with or without a pause. Pauses could be more or less long.

As expected, some individual differences do emerge in the word boundary condition. H* accents marking an AP boundary seem to be produced in this condition by Speaker1 and Speaker4, as shown by the presence of longer vowels and higher F0 than in the within word condition. The production of domain final high accents by these speakers is confirmed by a manual prosodic annotation carried out by the author and by a second expert. However, Speaker1 seems to produce stronger boundaries: this speaker also produces longer V-to-V lags, which could be interpreted as indicating the realization of a pause after the accented vowel. AP boundaries are not usually marked by a pause in French (Jun & Fougeron 2000), so this leaves a doubt on the categorization of this boundaries, which still seem weaker than the IP boundaries produced in the clause condition by the same speaker.

Speaker2, Speaker3 and Speaker5 show another pattern for the word boundary condition. Indeed, these speakers seem to produce *papa pilote* as a single AP marked by a high initial accent (Hi) on the second syllable of *papa*. Speaker2 and Speaker3 present in this context only a higher F0 than in the within word condition. On the other hand, Speaker5 presents both higher F0 and longer vowels, but she increases F0 to a lesser extent compared to the other speakers. Since Hi initial accents in French are not always significantly longer than other domain medial syllables, and can be weaker in pitch than AP final syllables, the characteristics exhibited by these three speakers can be indicators of a Hi accent (Jun & Fougeron, 2000). The prosodic annotation supports this indication given by the acoustic data.

In clause boundary condition, all speakers produce stronger boundaries as expected. Indeed, all present longer vowels and V-to-V lags in this condition than the within word and word boundary condition. However, there are differences as for F0: if Speaker1, Speaker2, and Speaker3 present higher F0 in this condition than the within word condition, suggesting the realization of a H% boundary tone, Speaker5 shows lower F0 in this condition than within word, suggesting the realization of a L%

boundary tone. Speaker4 show more variable F0 values in this condition, suggesting that she produced both H% and L% tones depending on the production.

8.4.2 Summary of the results on V-to-V coarticulation

The main outcome of this study relates to interspeaker differences in the coarticulation patterns in the word boundary condition. If four speakers reduce coarticulation across word boundary compared to within word, Speaker5 present the same degree of coarticulation in the two conditions.

Can these interspeaker differences be explained by the results on F0? All speakers increased F0 in the word boundary condition with respect to within word, producing either a H* or a Hi accent associated with the second syllable of *papa* in *papa pilote*. However, Speaker5, who presents no reduction in coarticulation in this condition, rises her F0 to a lesser extent, compared to the patterns exhibited by the other speakers. Since the two measures of coarticulation used in this analysis are partly based on F1, F0 height could have an indirect influence on the coarticulation measure. Indeed, a higher F0 would have the tendency to increase the value of F1, resulting in less F2-F1 distance, and thus in a less compact vowel. This effect of F0 on F1 would counteract the influence of V_2 /i/ on V_1 /a/. To test this hypothesis, a Pearson's correlation is run between F0 and the *acoustic assimilation index*, all speakers and conditions confounded. The two factors being weakly correlated (r=-0.33), it is reasonable to exclude that changes in coarticulation degree, as measured in this study, are dependent on F0 height. Prosodic differences can only partially explain the differences in coarticulation patterns between speakers in the word boundary condition.

A second outcome relates to the results on the clause boundary condition. If, as expected, all speakers show less coarticulation between clauses than between or within words, a certain degree of coarticulation is found for four speakers. Indeed, the analysis on the *F2-F1 compacity* of [a]i and [a]a shows that Speaker2 is the only speaker who does not present a significant difference between the vowel /a/ in the two contexts. The other speakers coarticulate to different extents.

Speaker-dependent is also the relationship between the degree of *acoustic assimilation* and the *perceived boundary strength.* Indeed, only Speaker1 shows a reduction of the *acoustic assimilation index* at the increase in *perceived boundary strength*. For the other speakers, either there is no relationship between these two factors or it holds only for some repetitions.

Finally, the last outcome of this study concerns the variability of coarticulation. Contrary to the expectations, no more variability in the degree of the acoustic assimilation of [a]i to /i/ measured by the *assimilation variability index* is found across word boundary than within word. Two speakers, Speaker1 and Speaker4, present a higher *assimilation variability index* across clause boundary than in the other conditions.

9 Discussion and conclusions

"Nature, however, is not very economical with respect to patterns of coordination"

Turvey 1990

9.1 Summary of the main results

In four studies, variations in the degree of anticipatory V-to-V coarticulation have been investigated according to population, in different Motor Speech Disorders (AoS and Dysarthria associated to ALS, Wilson Disease, Parkinson's Disease) and in adults spanning 20 to 93 years old; and according to condition in five speakers: in a slow and fast tempo compared to speakers' habitual tempo, and in three different boundary conditions, syllable boundary within a word, across word boundary and across clause boundary.

Overall, population-dependent and speaker-dependent variations in V-to-V coarticulation degree are found. V-to-V coarticulation is reduced in speakers affected by both AoS and dysarthria, suggesting that an impairment in extrasyllabic V-to-V coarticulation can be a characteristic of disordered speech, regardless of pathology type.

A reduction of V-to-V coarticulation is also found with age in neurotypical adults and is showed to follow a non-linear pattern: coarticulation degree gradually decreases starting from 20/30 years old up to approximately 50 years old, with a steeper decrease after approximately 70 years of age. More varied coarticulation patterns are showed for middle aged speakers. The results on the relationship between coarticulation and segmental durations in MSD and between coarticulation and articulation rate in neurotypical adults shows that, if there is a tendency for these factors to covary to a certain extent, a direct relationship cannot be drawn.

Analysis on individual responses to tempo manipulations show speaker-specific variations in V-to-V coarticulation depending on speech tempo. Indeed, out of five speakers, three to four speakers decreased coarticulation while speaking slower. While speaking faster, only two speakers increased their coarticulation degree.

Inter-speaker variations also emerge in the analysis of V-to-V coarticulation patterns depending on boundary type. Indeed, across word boundary, four speakers out of five reduce coarticulation degree with respect to coarticulation within word, while one speaker present the same coarticulation degree in the two conditions. Across clause boundary, four out of five speakers coarticulate to different extent, while only one speaker does not present coarticulation between clauses. No more variability is found in the word boundary condition than within word, while more variability is found for two speakers in clause boundary condition with respect to the other two boundary conditions. The inter and intraspeaker variation in coarticulation degree is not entirely accountable for by prosody, at least for four out five speakers.

9.2 Preface to the discussion

Two were the main research questions of this dissertations, as presented in the first chapter (section 2.3): 1) to what extent V-to-V coarticulation covaries with articulation rate? 2) What are the units of speech over which the movements for a string of speech targets are planned?

Point on question 1: to what extent V-to-V coarticulation covaries with articulation rate?

The relationship between coarticulation and rate has been investigated depending on population in MSD and with aging and depending on condition by eliciting changes in speech tempo. The results on MSD suggest that an impairment in anticipatory V-to-V coarticulation, albeit not specific to a type of disorder, it cannot be uniquely accounted for by a slowing of speech, which characterizes AoS, ALS and Wilson Disease. Indeed, a reduction of coarticulation was found also for Parkinson's Disease speakers, who do not show longer segmental durations. Overall, the results of the age study indicate that there is a tendency for coarticulation to covary with rate for speakers aged 20 to 50, and for speakers aged 50 to 70. However, this relationship is age- and speaker- dependent: overall, middle-aged speakers tend to coarticulate less than younger speakers at the same rate, and the relationship between coarticulation and rate is not linear for all speakers, with some showing faster rates and a lower coarticulation degree than slower speakers. The results on tempo changes also indicate that coarticulation does not linearly decreases with a decrease in rate and increases with an increase in rate. Therefore, it can be concluded that the relationship between rate and coarticulation is not linear. The different results listed here will be further addressed in the discussion in relation to the units of motor planning, therefore in relation to research question 2 (*what are the units of speech over which the movements for a string of speech targets are planned?*), and to population-specific and speaker-specific variations in coarticulation.

Point on coarticulatory direction

In the MSD study, anticipatory V-to-V coarticulation has been compared to carryover C-to-V coarticulation. While one of the reasons the two were compared is that they differ in the domain over which coarticulation is observed (extra- vs intrasyllabic), another reason was that they differ in direction of the effect. One hypothesis was that impaired V-to-V coarticulation and unimpaired C-to-V coarticulation for AoS (who present impairment in motor planning), and unimpaired V-to-V coarticulation and impaired C-to-V coarticulation for dysarthric speakers (who present impairment in the execution of motor programs) could indicate different mechanisms underlying these coarticulation types. However, this prediction was not supported by the results, that showed impaired V-to-V and unimpaired C-to-V coarticulation for all MSD groups.

These results will be therefore discussed in relation to the coarticulation domain, that is, intrasyllabic vs extrasyllabic.

9.3 What are the units of speech over which the movements for a string of speech targets are planned?

The question of the nature of the units of planning in speech production has been mainly addressed in regards to the phonological units of encoding. In Levelt, Roelofs, and Meyer (1999) the unit of phonological encoding is considered the phonological word, defined as a stressed word + the unstressed words attached to it. Thus, phonological planning at the post-lexical level is considered only in the case of forms like word + clitic, in order to account for phenomena of resyllabification in English (e.g. in *escort us*). However, psycholinguistic studies on phonological priming, a paradigm in which it is tested whether the presentation of a phonologically overlapping word speeds up the production of the target word, have suggested that the scope of phonological planning can extend from a syllable, to a word, to more than one word (Schriefers & Teruel, 1999; Meyer *et al.*, 2003; Alario & Caramazza, 2002; Schnur, Costa, & Caramazza, 2006; Schnur, 2011). In French, phenomena like the liaison, that is, the phenomenon in which a latent final consonant is obligatory pronounced in adjective noun phrases when the adjective is followed by a noun starting with a vowel, suggest that phonological planning can extend over at least two words. This has been experimentally tested by Michel Lange and Laganaro (2014), who investigated the effect of priming of the noun or of the adjective in adjective + noun phrases that contained an obligatory liaison in French (e.g. *les trois aimants* "the three magnets"), showing that either the first word either the whole sequence was encoded before starting the production. But *over which units the planning of the phonetic detail is made?* The interplay between phonological planning and the planning of phonetic details has been addressed in studies on prosody (among many others, Fougeron & Keating, 1997;

Fougeron, 2001; Keating & Shattuck-Hufnagel, 2002; Cho, 2016). Indeed, speakers are likely to plan the prosodic structure of an utterance prior to the motor execution. Therefore, some phonetic details of an utterance such as the relative timing of the segments, stress and F0 contour are planned over a prosodic constituent or over an entire utterance (Cho, 2016). A certain degree of correspondence between units of phonological planning, the units defined by the prosodic organization of the utterance, and units where the movements for the achievement of a target in a particular context are planned, is expected (e.g. Ma *et al.,* 2015). However, it is not clear whether the units over which coarticulation is specified correspond to a syllable, a word, or more words within a prosodic unit.

Is there a special cohesion between syllables belonging to the same word?

Anticipatory V-to-V coarticulation, as seen in the introduction (section 1), has been mainly investigated within a word, and evidence is provided that this coarticulation is planned, therefore that this coordination is specified. The arguments adduced span from Whalen's (1990) experiment, that showed that extensive anticipatory V-to-V coarticulation has to be planned prior to execution to appear, to language-specific patterns of coarticulation, which indicate that specific patterns of coarticulation are acquired by native speakers (e.g. Beddor *et al.,* 2002; Ma *et al.,* 2015). But, *is this coordination between two vowels specified because they are in the same word?* V-to-V coarticulation has also been observed across words (e.g. Abry & Lallouache, 1995; Grosvald, 2009; Mok, 2010), and it has not usually compared to coarticulation within a word. Hardcastle (1985) who compared the overlap between two consonants in a CC cluster across syllable boundary within a word, word boundary, clause and sentence boundary, reported inconsistent results as for the comparison between syllable and word boundary (see section 3.2.4). The question of whether the coordination between two vowels belonging to different syllables within a word differs from the coordination between two vowels belonging to different words is still open.

The difference between intraword coordination and interword coordination has been previously addressed in the literature by comparing coordination within a monosyllabic word vs coordination between two monosyllabic words. In an electropalatographic study, Byrd (1996a) investigated the degree of overlap between two consonants in CC clusters of different compositions (/dg, gd, ks, sg, sk/) in onset position within a monosyllabic word (#CC, where # is a word boundary), in coda position within a monosyllabic word (CC#), and in heterosyllabic position between two monosyllabic words (C#C). Note that here heterosyllabic = across word boundary. Her results showed:

- The first consonant is less overlapped by the following consonant (less anticipation) in an onset cluster than a coda or heterosyllabic cluster. A tendency is found toward less overlap in heterosyllabic than coda cluster, but it is not systematic: indeed, it depends on the speaker. Moreover, the stop-fricative sequence has a specific behavior, with more overlap in heterosyllabic than coda clusters and overall little overlap in coda clusters. The general pattern is:
    - Overlap ⇨ onset < heterosyllabic < coda
- Sequence duration is longer in onset clusters than in coda or heterosyllabic clusters:
    - Sequence duration ⇨ onset > heterosyllabic > coda
- Less variability in relative timing is found in the onset clusters than in the coda or heterosyllabic clusters. More variability is found in heterosyllabic than coda sequences only for the /sk/ cluster. The other clusters do not show differences in variability in heterosyllabic vs coda sequences.

If we consider the results relative to coda vs heterosyllabic/between words clusters (leaving aside onset clusters, which can exhibit a different behavior), Byrd's results indicate overall no more variability across a word boundary than within a word, and a tendency toward less overlap in heterosyllabic clusters, thus across word boundary, than in coda clusters, thus within a word, with speaker-dependent variations. Therefore, her findings are similar to the findings of the **boundary** study (chapter 8): more V-to-V coarticulation is generally found within (*papi)* than across two words (*papa pilote*) that are or are not separated by a prosodic boundary, with the exception of one speaker, and no more intraspeaker variability in coarticulation degree is found across than within word. These results suggest that elements belonging to the same word do present a pattern of coordination that is different from elements belonging to different words. In our specific case, V-to-V coarticulation within a word would be specified in the gestural score, therefore at a lexical level.

A parenthesis can be opened here about the *measure of variability of coarticulation* used in the **boundary** study. Indeed, it was postulated that the stability of coarticulation could be an indicator of a specified coordination, while more intraspeaker variability could be associated to an uncontrolled (random) overlap between gestures. However, no difference in variability is found across words than within words (and only for two speakers there is a difference between clause boundary and within word), raising the question: *is intraspeaker (intertoken) variability over 45 repetition a good index of the stability of intergestural coordination?* The answer is no. There are three possibilities: 1) variability of coarticulation is not a good assessment of a specified coordination; 2) 45 repetitions are not enough to test this. Whalen & Chen (2019) who observed the stability of coarticulation across multiple repetitions, had collected 300 tokens of each word; 3) rather than intraspeaker variability, interspeaker variability has to be tested: I have not considered it in this analysis, but it would be interesting to see whether there is more interspeaker variation in coarticulation across words than within a word. More speakers would have to be examined to test this.

The next question I would like to raise is:

Is the word <u>always</u> a unit where coarticulation is specified?

Some of our results suggest that, if the word is a unit in which V-to-V coarticulation is specified, there are some cases where this unit is "broken up", and speakers plan on smaller units. In particular there are two results to discuss in this respect: the results on V-to-V coarticulation in AoS speakers and the results on V-to-V coarticulation at slow tempo.

"Breaking" of word-sized units in AoS:

Less extrasyllabic V-to-V coarticulation and unimpaired intrasyllabic C-to-V coarticulation is found in AoS speakers with respect to control speakers.

How to explain a reduction of V-to-V coarticulation in AoS speakers? A reduction of extrasyllabic anticipatory coarticulation in AoS has been previously reported in the literature (e.g. Ziegler & Von Cramon, 1986a; Dogil & Meyer, 1998) and could be attributed to an impairment in motor planning (e.g. Ziegler & von Cramon, 1986a). Indeed, the impairment for AoS is considered to be located at a stage of speech encoding where speech units are assembled and coordinated with each other (e.g. Kelso & Tuller, 1981; see section 3.3.2.2). The reduction found in anticipatory V-to-V coarticulation in AoS indicate that these speakers could have difficulties in planning the movements over a word-sized speech plan, and therefore would not plan coarticulation across syllables. In other words, in AoS coarticulation is not specified between the two syllables *pa* and *pi* within a word - at least not in our results: a characteristic of AoS is the inconsistency of errors (e.g. Ziegler, 2008) so it would be interesting to look at the variability of their extrasyllabic coarticulation to see if sometimes they are planning on word-sized unit -. A certain degree of anticipatory coarticulation found in the productions of these speakers could be attributed to a non-controlled overlap between gestures.

How to explain unimpaired intrasyllabic C-to-V coarticulation for AoS speakers? The finding that C-to-V coarticulation is unimpaired in AoS suggests that these speakers restructure the organization of articulation over syllable-sized units. The hypothesis that the syllable, as a unit of speech encoding, is resistant to the apraxic degeneration, has been supported by previous studies on speech errors. In their main idea, these studies refer to Levelt's proposition that motor plans for high frequency syllables are stored in a mental syllabary, and that therefore the retrieval of a certain syllable should be influenced by its frequency of use (Levelt and Wheeldon, 1994; Levelt *et al.,* 1999). Studies on the error rates in syllable production reported a tendency toward less production errors for high frequency than low frequency syllables in AoS (Aichert & Ziegler, 2004; Laganaro, 2008; Staiger & Ziegler, 2008). Syllable-level motor plans would be intact in AoS speakers, which would reflect in unimpaired C-to-V coarticulation. This planning over smaller units would reflect in another characteristic of AoS speech, that is, syllable segregation (Rosenbeck *et al.,* 1984; Duffy, 2012).

"Breaking" of word-sized units in slow speech:

How to explain the patterns of coarticulation found at slow tempo? Studies on kinematic responses to rate changes have raised the question whether speaking at a slow rate entails a different mechanism than speaking at normal or fast rate. Adams, Weismer, and Kent (1993) and Perkell, Zandipour, Matthies, and Lane (2002) reported at slow rate velocity profiles characterized by multiple velocity peaks, which suggested less smoothness at slow speech, and thus an "unnaturalness" of slow speech. An "unnatural" articulatory behavior at slow speech could reflect a change in the organization of intra and intergestural coordination. In this direction could be interpreted the patterns of V-to-V coarticulation showed by the five speakers while speaking slower, which can indicate a reorganization of intergestural coordination over smaller units in this condition (a "syllabified" speech). In four speakers, this reflects in a reduction of the coarticulation between the two vowels, with some residual degree of coarticulation due to uncontrolled overlap between the vocalic gestures. In one speaker, this leads to more variability in coarticulation: since the stability of the coordination is lost, the two vowels randomly overlap, thus there can be more or less coarticulation than at normal rate depending on the repetition. This view is supported by previous results of the literature on coarticulation at slow rate (reviewed in 3.1.1). As for V-to-V anticipatory coarticulation, Hertrich and Ackermann (1995) reported inconsistent patterns of coarticulation at slow speech, with either unchanged or increased coarticulation with respect to a normal rate, depending on speaker. If the overlap between two vowels is uncontrolled, inter and intraspeaker variations, with same, more (or less) coarticulation can arise. In the direction of a restructuration of planning over syllables rather than larger units in slow speech can be interpreted also the results of Tjaden and Wilding (2005) who reported a decrease of extrasyllabic VC, but not intrasyllabic CV coarticulation at slow rate.

How to explain coarticulation beyond word unit?

The phrase *papa pilote* has been produced by the speakers in the word boundary and clause boundary condition with the following prosodic phrasing.

Word boundary:

[papa pilote] AP (3 speakers)

[papa] AP [pilote…] AP (2 speakers)

Clause boundary:

{papa} IP {pilote…} (all speakers, with pauses of different durations at the IP boundary)

I will consider first the cases where either an AP boundary or an IP boundary is produced between the two words. Across AP and IP boundary less coarticulation is found with respect to the within word condition. This finding supports previous results by Cho (2004) showing a decrease in V-to-V coarticulation at the increasing of the strength of the prosodic boundary from word to ip to IP, and is in line with other studies showing a decrease in coarticulation at the increasing of boundary strength (also in V-to-V coarticulation, Cho, 2004; in CV coarticulation, e.g. Meynadier 2004).

Is coarticulation across an AP or IP boundary planned?

Across IP and AP boundary, I do not expect the two vowels spanning the boundary to be part of a single planning unit over which coarticulation is specified. Indeed, an utterance can be considered a succession of units that are tuned by the prosodic organization, and thus a unit where coarticulation is specified would be comprised in one prosodic domain, but is not expected to straddle two domains (Fougeron, 2022, personal communication). Being the overlap between the two vowel gestures not specified at the level of the gestural score, it is "free" to vary under the influence of prosody, for instance under the action of a π-gesture (see section 3.2.1). The π-gesture acts transgesturally by slowing the temporal unfolding of the constriction gestures, that become longer and less overlapped. The action of the π-gesture occurs when the gestural score is already specified (Byrd & Saltzman, 2003) thus at a post-lexical level. Since the π-gesture activation strength depends on the strength of the prosodic boundary, coarticulation is expected to decrease at the increase of the boundary. This prediction is reflected in the results of Speaker1: this speaker produces in the word and clause boundary conditions a continuum of boundaries of different strength, and shows a strong linear correlation between coarticulation degree and boundary strength. The weak relation between prosodic boundary and coarticulation for the other speakers across IP boundary can be explained by the fact that the other speakers vary less in their degrees of prosodic boundary strength in the clause condition.

Is some overlap between vowels in VCV sequences inevitable in connected speech?

Before addressing the next point on planning, I would like to raise a point about the results on coarticulation across IP boundary. All speakers presented coarticulation to a certain extent across IP boundaries that were often marked by a pause, except for the speaker who made the longest pauses. This result suggests that a certain degree of overlap between the two vowel is inevitable, if the vowel gestures are not too far apart to overlap. This result is compatible with accounts that consider vowels and consonants as inherently different gestures. Öhman (1966) originally proposed that the articulation of vowels constitutes the substrate of articulation on which consonants are superimposed, and this view has been adopted by other models of coproduction (see section 2.1.2.1). In Articulatory Phonology consonant and vowels are on different tiers of the gestural score (Browman & Goldstein 1990). This means that vowels are always adjacent in connected speech. If the articulation between vowels is continuous regardless of intervening consonants, some coarticulation would inevitably occur, so there would be always a certain degree of V-to-V coarticulation if the two gestures are not too far apart (e.g. in the case of a long pause).

Is coarticulation across words inside an AP planned?

Three speakers produce *papa pilote* as one AP, with an accent (a high initial accent, see section 8.2.2.2.1) on the second syllable of *papa.* Of these three speakers, two speakers present less coarticulation than within word, and one (Speaker5) present the same coarticulation than within word. *How to explain this difference? Is coarticulation specified or not specified in this case?* I will propose two interpretations of these results:

1. Speaker5 encodes the subject + verb phrase *papa pilote* as a single planning unit over which coarticulation is specified, thus it is not affected by the presence of a prominence, characterized by higher F0 and vowel lengthening; on the other hand, coarticulation is reduced across two words than within word for two speakers because the two words are encoded in two different planning units. Coarticulation is not specified, and it is affected by the presence of a prominence (characterized by a rise of F0), who affects the spatial magnitude of the gesture of [a]i leading to less coarticulation. In the direction of individual differences in

the size of planning units can be interpreted the results of previous studies reporting interspeaker differences in the scope of coarticulation across words (e.g. Grosvald, 2009). Moreover, the idea that the units of planning can be speaker-dependent is supported by studies on phonological planning, who have reported interspeaker differences in the effects of phonological priming (Meyer *et al.*, 2003; Michel Lange & Laganaro, 2014). For instance, Michel Lange and Laganaro (2014), in the study on the liaison mentioned earlier, showed that speakers could encode either only the first word (so not producing the liaison between the two words) or the phrase composed by two words (so producing the liaison) before starting production.

2. Coarticulation is specified within an AP, and the degree of coarticulation is regulated by the action of a modulation gesture or μ-gesture (see section 3.2.1). The μ-gesture (Saltzman *et al.*, 2008) shapes the effects of prominence on the constriction gestures in a similar way as the π-gesture, that is, gestures under the action of a μ-gesture can be longer and less overlapped. However, the μ-gesture is implemented in the gestural score at its creation, thus is specified at a lexical level. The action of a μ-gesture could account for a reduction of coarticulation between two words in which coarticulation is specified, by changing the specification of this coordination in the gestural score. In the case of the two speakers who reduce coarticulation and present an accent characterized by a F0 rise but no vowel lengthening, the reduction of coarticulation could be accounted for by a spatial μ-gesture (μ_S) that acts on the spatial properties of the constriction gestures ("resizing" the gesture) and has been proposed to account for effects of prominence which are different from the effects of a prosodic boundary. But *how to explain the pattern of coarticulation of Speaker5, who also presents an accent?* The absence of a reduction of coarticulation under prominence for this speaker can be related to specific characteristics of her speech. Indeed, this speaker shows a low degree of coarticulation, as it can

be observed both within word and in the baseline of the **tempo** study, and tends to keep “stable” the degree of coarticulation across the tempo and boundary conditions. In particular, she tends to not decrease overlap further. It is possible that this speaker presents already the least overlap possible (for her) between the two vowels, therefore she cannot decrease coarticulation further in a fluent speech, when there is not a pause between the two words (see also the fact that at slow rate she does not decrease systematically coarticulation, but the effects of a slowing of speech reflects in the “loss of consistency” of the overlap, thus in an increased variability).

9.4 How to explain variations in coarticulation within word?

The results of the different studies, and in particular of the **MSD**, **age** and **tempo** studies (chapters 5, 6, and 7), have highlighted variations in coarticulation depending on population or on speaker’s identity that cannot be accounted for by articulation rate. Indeed, the relationship between coarticulation and rate depends on age, but is also speaker-dependent, which in turns reflects on the different responses showed by speakers to an increase in speech tempo. Speaker-specific patterns of coarticulation have been reported in several studies. For instance, interspeaker variations of coarticulation have been shown in the degree of anticipatory vowel nasalization in American English (Zellou, 2017), in the scope of anticipatory labial coarticulation in French (Abry & Lallouache, 1995; Robert *et al.*, 2005) and in the degree of C-to-V coarticulation in Dutch (van den Heuvel *et al.*, 1996). In modeling coarticulation, speaker-specific, but also population-specific variations in coarticulation can be taken into account by allowing a constrained variability in speech targets. In other terms, by considering speech targets not as points, but as ranges of acceptable values, that is, as windows, as originally proposed by Keating (1990). Windows can be seen from a spatial point of view, allowing a modification of the speech targets as a function of the surrounding context, or also from a temporal point of view, allowing variation in the overlap or temporal coordination between gestures. In a spatial perspective, the view of speech targets as windows has been adopted by models of optimal planning such as GEPPETO (e.g. Perrier, 2014) and DIVA (e.g. Guenther, 1994), as reviewed in section 2.1. In these models, speech goals are target ranges in the acoustic or orosensory spaces, which include all the acceptable realizations for a given speech sound. Anticipatory coarticulation is done by choosing, within the acceptable range for a speech goal, a target that is closer to a following speech target in a compatible dimension in order to minimize the displacement between the two targets (Figure 33).

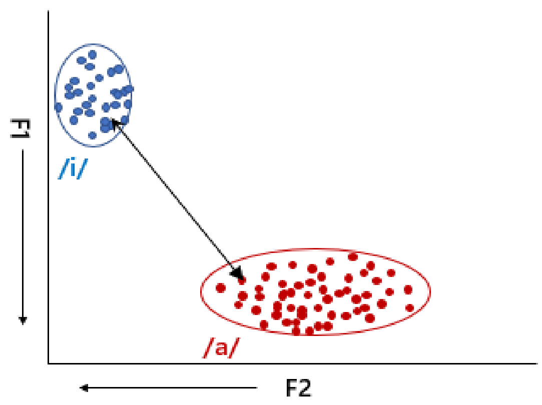

Figure 33. Coarticulation between /a/ and /i/ in the F1 and F2 dimensions in a model where windows are considered from a spatial perspective. The arrow represents the coarticulated targets chosen to limit the displacement in the articulators. This illustration of coarticulation is inspired by the GEPPETO model, in which speech goals are ellipsoids in the acoustic space (Perrier, 2014; see also Barbier, 2006, for a similar illustration).

Here, I will focus on windows in a temporal perspective. The possibility for a variability in the temporal specification of the overlap between speech gestures has been considered by Byrd (1996b; Byrd & Saltzman, 2000) by implementing in the Task Dynamic framework phase windows, inspired by the window model (Keating 1990). In the next paragraphs I will address the population-dependent and speaker-dependent variability in our data by adopting this perspective of variability in intergestural coordination.

Variability in intergestural phasing

In the original Articulatory Phonology framework, the relative phasing between gestures in word production was specified "by hand" in terms of invariant point-to-point phase relations. The need for the implementation of a mechanism that allows to a certain extent variability in the relative timing between gestures has been addressed in the Task Dynamics framework by Byrd (1996b) and Saltzman and Byrd (2000). Byrd (1996b), in order to account for variability in intergestural coordination according to linguistic factors, such as phrasal structure, and extralinguistic factors such as speech rate, postulated the implementation in Task Dynamics of phase windows. The phase windows are temporal ranges in which a specific phasing between two gestures can occur. The specific range of the window depends on the type of gesture and on the relationship that gestures have with each other. For instance, it is hypothesized that the phase window is wider for intergestural coordination between words than within a word (Byrd 1996b).

An example of phase window given by Byrd for the phasing of two consonants is illustrated in Figure 34.

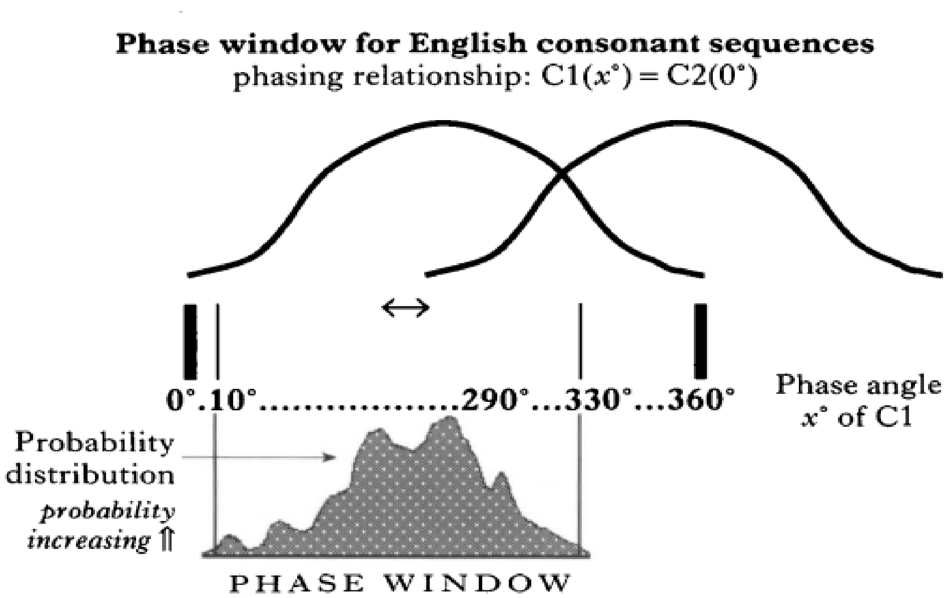

Figure 34. Phase window for english consonant sequence. Extracted by Byrd (1996b)

The lower and upper limits of a particular phase window are constrained by physical and linguistic factors, and inside this window there are some points in time where the phasing between the gestures is more probable to occur. The probability that a particular phasing within the window could happen is represented in terms of probability density (displayed in the figure as a density curve). Virtually the phasing between the two gestures could occur at any timepoint of the phase window, but the probability that this happens is small, because several factors determine the final probability density for a window of a specific segment in a particular utterance. The temporal region of the window where the actual phasing between the gestures is implemented will determine the degree of temporal overlap between the two gestures.

Different linguistic and extralinguistic variables weight on the probability density of the window and therefore influence the probability that a particular phasing will be implemented. Considered variables are for instance context, prosody and speaking rate (Byrd 1996b). I will focus in particular on speaking rate, to address how it interacts with factors related to speaker's identity. These variables, called influencers, are considered to have an active influence on the window. For instance, an increase in speech rate would skew the probability density toward a more overlapped phasing within the window. Different influencers may allow different amounts of variability, i.e. if an influencer favors a narrow phase window, there will be less variability. Finally, different influencers can be active at the same time on a window, and they could have different weights on a window, that is, one could be stronger than another in weighting the window.

How to explain the results of the age and tempo studies (chapters 7 and 6) with phase windows?

The degree of overlap between V_1 /a/ and V_2 /i/, and thus the degree of observed spectral assimilation between the two vowels, would depend on where (at what timepoint) within the phase window for one particular production the gesture for the /i/ starts, and thus where the phasing with /a/ occurs. The phase window for the overlap between the two vowels could look something like this (Figure 35):

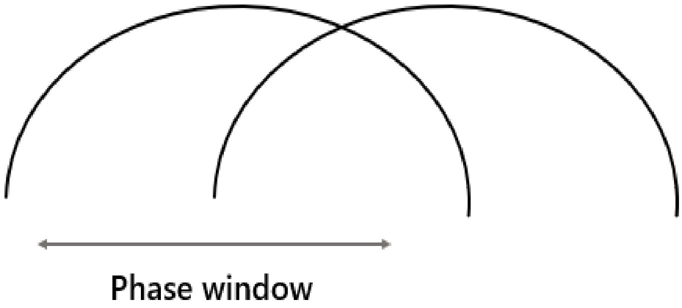

Figure 35. Phase window for the overlap between /a/ and /i/. The two vowels are considered as contiguous (as they would be if they are on the same tier) thus the intervening consonant is not represented here. The width of the arrow represents the width of the phase window, containing all the timepoints in which the actual overlap can be implemented.

Inside this window, the probability that a particular phasing occurs is influenced by three main factors, rate, speaker's age and speaker's identity. Indeed, speaker's identity, with all that it entails, that is, physical specificities but also style of speaking/reading, age, would be an influencer that weights on the probability density of a window and interacts with rate. If the influencer *speaker's identity* is coactive with the influencer *rate*, and if *speaker's identity* weights more on the window than the influencer *rate*, we will observe interspeaker variations. Let us observe this closer.

Expected influence of an increase in rate on coarticulation: example based on the results for the young speakers of the age study and for two speakers in the tempo study.

An increase in rate could skew the probability density of the window toward a more overlapped phasing. The size of the window is the same, but inside this window more overlap is made more probable by the action of fast rate (Figure 36).

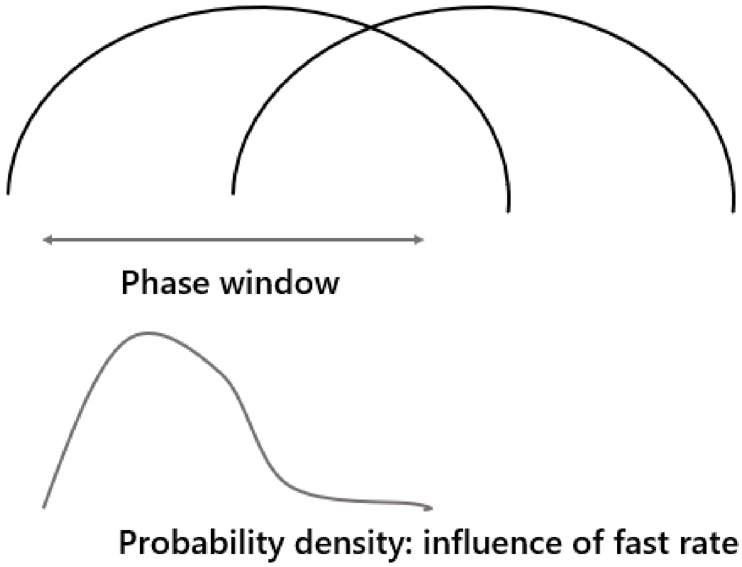

Figure 36. Phase window for the overlap between /a/ and /i/, with the effect, on the probability density curve, of the influencer "fast rate". A fast rate makes more probable a more overlapped phasing within the window. The representation of the probability is only illustrative and it is as postulated by the author.

If the only influencer on this window is the rate, or if the coactive influencers *age* and *speakers' identity* do not oppose this effect, a more overlapped phasing can be expected to be implemented, and more coarticulation would be observed (Figure 37). This is what it is overall observed for the speakers aged 20 to approximately 50 of the **age** study (chapter 6): speakers with higher rates generally present also more coarticulation. The same pattern is exhibited in the tempo study by the two speakers out of five who increase their coarticulation degree at fast tempo.

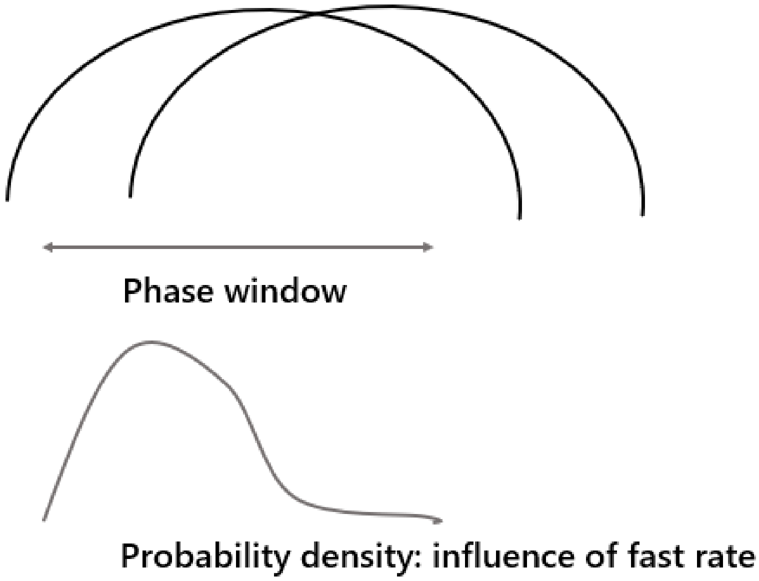

Figure 37.Phase window for the overlap between /a/ and /i/, with the action of the influencer "fast rate", which causes more overlap.

However, in the case of older speakers, or depending on speakers' identity, other factors are at play in weighting the window.

Influence of speakers' age, considered together with rate, on coarticulation: results for the ***middle-aged*** *speakers of the age study, compared to younger speakers.*

Middle-aged speakers exhibit less coarticulation than younger speakers. If coarticulation covaries with age for these speakers, that is, faster speakers have more coarticulation than slower speakers, they present overall less coarticulation than younger speakers at equal articulation rates. These speakers could favor a more "hyperarticulated" speech style, which can be assimilated to clear speech and is characterized by less coarticulation, as found in other studies looking at anticipatory coarticulation in different speech styles (Duez, 1992; Scarborough & Zellou, 2013). A tendency for hyperarticulated speech targets in speakers in their 50ies is also suggested by the longitudinal study of Gahl and Baayen (2019), who showed that middle-aged speakers produce more peripheral targets compared to when they were younger, and this hyperarticulation is found even for short vowels, when there was less time to reach the target. *Speakers' age* can be considered an influencer that acts on the probability density of the phase window. In middle aged speakers the probability density of the window would be skewed toward a less overlapped phasing between V_1 and V_2. On the contrary, in younger speakers, the probability density of the window would be skewed toward a more overlapped phasing between V_1 and V_2. - This hypothesis, that is, that for younger speakers there is an "active" tendency toward more overlap, is based also on the results that we found in a previous study, where younger speakers (<40) overall increased coarticulation degree in a sentence repetition at a comfortable rate vs a reading task, and in a fast repetition vs a repetition at a comfortable rate (D'Alessandro, Bourbon, & Fougeron 2020) - . If the influencer *age* weights more on the probability density of the window than *rate*, this would lead to different degrees of overlap for two speakers of different ages, at the same rate. In Figure 38 the overlap between /a/ and /i/ at a moderate rate (12.5 ph/sec) is modeled for a 54 years old speaker (on the left) and a 25 years old speaker (on the right), based on the data of the age study.

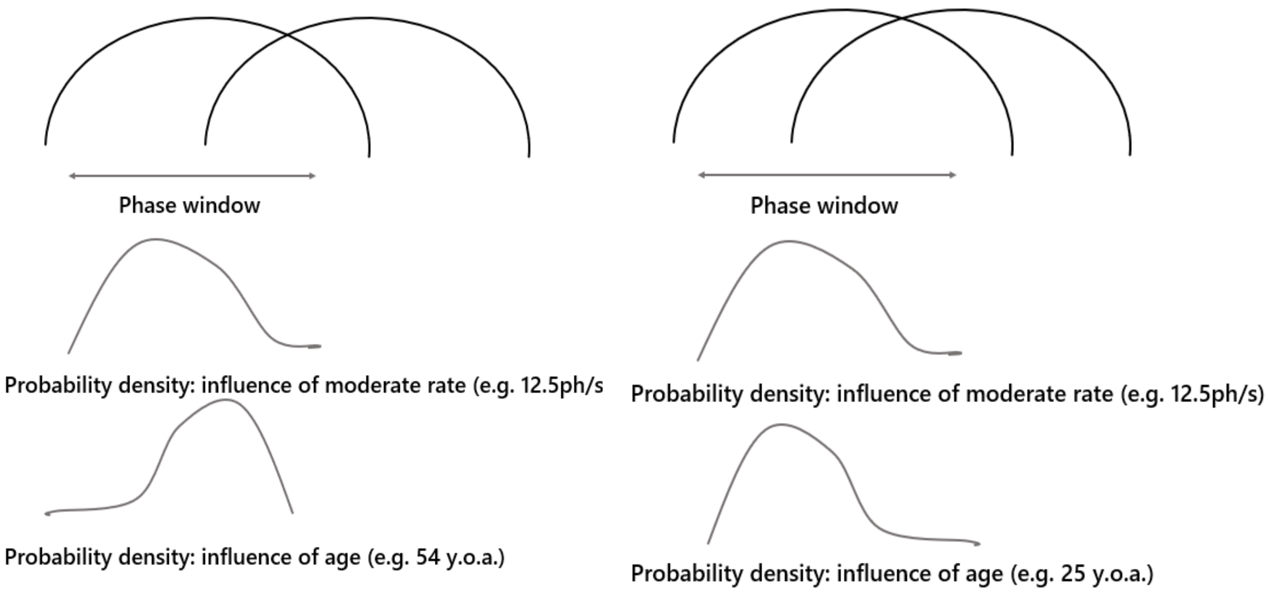

Figure 38. Phase windows for the overlap between /a/ and /i/. Postulated effect of rate 12.5 ph/s and ages 54 vs 25 on intergestural phasing on the basis of the observed data. The influencer age weights more than rate, leading middle-age speaker to exhibit less overlap than younger speakers at the same rate.

At equal rate, the 54 years old speaker has less coarticulation, because their tendency to favor a less overlapped phasing counteracts the influence of a moderate rate.

Influence of speakers' age, considered together with rate, on coarticulation: results for the ***older (> 70)*** *speakers of the age study.*

Speakers above 70 y.o.a. show low coarticulation, overall slower rates, but a weak relationship between these two factors. Indeed, there are some speakers presenting low degree of coarticulation and slower rate, but a majority of speakers show lower coarticulation degree than younger speakers at similar articulation rates, and some speakers show really low coarticulation degree at moderate speaking rates. The weak covariation of rate and coarticulation for these speakers too could be explained by a more hyperarticulated reading and speaking style. In the study already mentioned on the effects of an increase in speaking rate on coarticulation and on articulatory precision in younger and older speakers (D'Alessandro *et al.,* 2020), we found that older speakers (>68 years old) applied a different strategy than younger speakers to increase rate. Indeed, with an increase in articulation rate, younger speakers decreased articulatory precision, resulting in vowel undershoot, and increased coarticulation, whereas older speakers increased articulation rate (with the same acceleration as younger speakers) without changing their degree of articulatory precision and without increasing coarticulation. This suggests a tendency to hyperarticulate in older speakers which would reflect also in a reduced coarticulation degree, regardless of speech rate, maybe as a strategy to increase intelligibility. A tendency toward a more careful and hyperarticulated speech for older speakers is supported by the results of Mücke, Hermes, and Tilsen (2020), on the production of CCV syllables in German by speakers aged 70 to 80 and younger speakers. Older speakers showed a more symmetrical organization of the consonants of the cluster and the vowel, which suggests a tendency to hyperarticulate. Considering the phase windows for these older speakers of the **age** study, similarly to the scenario postulated for middle-aged speakers, the influencer *age* would favor the implementation of a less overlapped phasing. With respect to a middle-aged speaker, an older speaker could show a stronger skew toward a less overlapped phasing, leading the /i/ gesture to start closer to the right edge of the window (Figure 39).

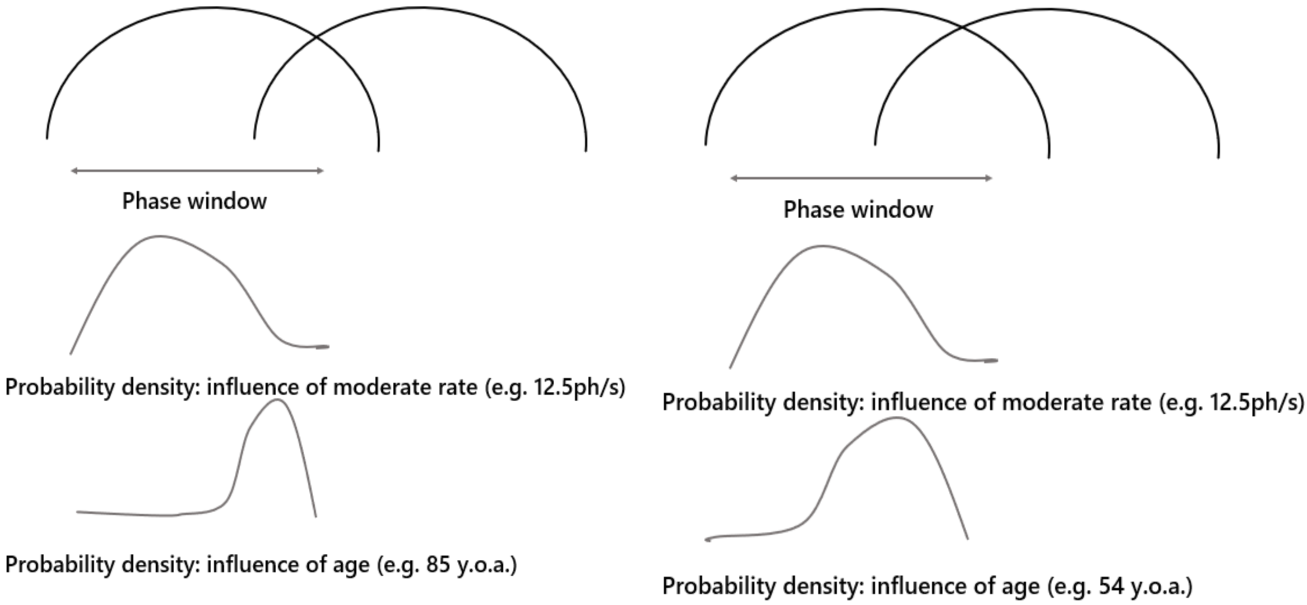

Figure 39. Phase windows for the overlap between /a/ and /i/. Postulated effect of rate 12.5 ph/s and ages 85 vs 54 on intergestural phasing on the basis of the observed data. The influencer age weights more than rate, leading the older speaker to exhibit less overlap than the middle-aged speaker.

Influence of speaker's identity: interspeaker variations in the effect of fast tempo on coarticulation in the tempo study.

Of the four speakers who increase their articulation rate in the tempo condition, only two speakers increased their coarticulation degree at fast tempo. Speaker-specific effects of rate changes on anticipatory V-to-V coarticulation have been previously reported (e.g. Matthies *et al.*, 2001), and can reflect an influence of speakers' identity or speakers' speaking style on the relationship between rate changes and intergestural coordination, similarly to what has been reported in studies on kinematics (Berry, 2011; review in section 3.1.2). An increase in rate would favor an increase in overlap, as already seen, skewing the probability density of the window toward a more overlapped phasing between /a/ and /i/. However, some speakers could favor a less overlapped speech style, and would adopt another strategy to increase rate than increasing overlap, for example by increasing the velocity of the articulators to reach the same /a/ target in less time (e.g. Van Son & Pols, 1992; Ostry & Munhall, 1985; Goozée *et al.*, 2003). This preference toward a less overlapped phasing in a speaker can be represented by a probability density favoring a narrow region of the phase window. If *speaker's identity* weights more than *rate* (as seen with age), we will observe interspeaker variation in the effect of an increase in rate in coarticulation. In Figure 40 the overlap between /a/ and /i/ in the baseline, that is, at a normal rate for the speaker, and with an increase in rate in the fast tempo condition, is modeled for a speaker of the **tempo** study who does not increase coarticulation. The influencer *speaker's identity* favors a narrow window for the overlap, therefore the increase in rate, who skews the probability density of the window toward more overlap, has no effect on the implemented degree of overlap between the two vowels.

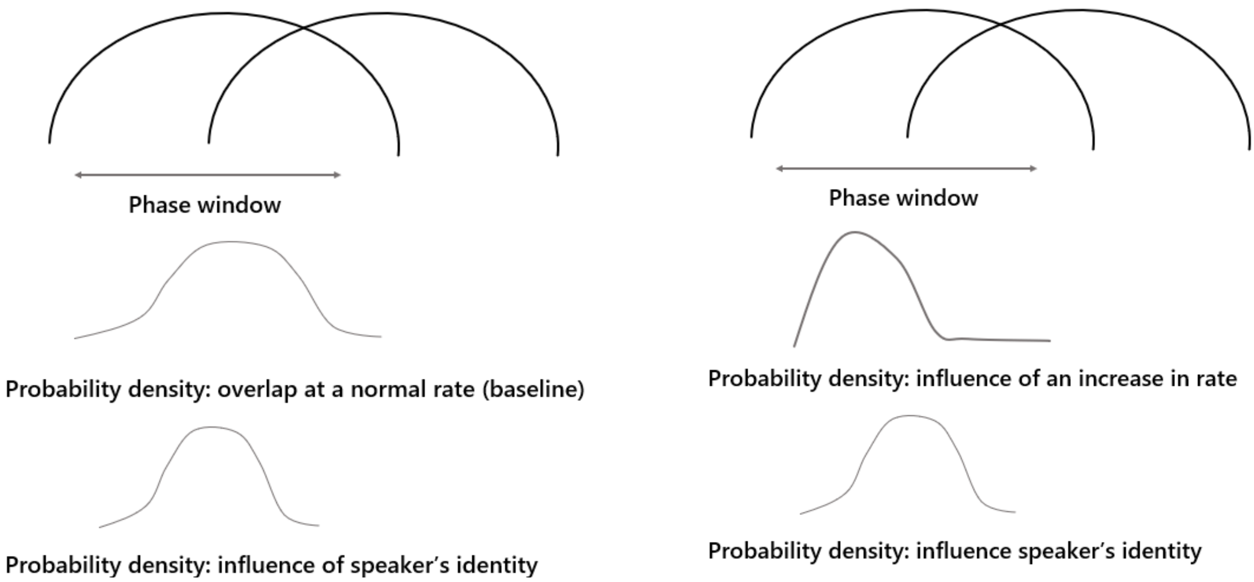

Figure 40. Phase windows for the overlap between /a/ and /i/ at a normal (baseline, on the left) and a fast tempo (increase in rate, on the right) for a speaker who does not increase coarticulation with the increase in rate. The influencer speaker's identity « overrides » the effect of rate on the overlap.

In Figure 41 the overlap between /a/ and /i/ in the baseline, that is, at a normal rate for the speaker, and with an increase in rate in the fast tempo condition, is modeled for a speaker of the tempo study who increases coarticulation. The influencer *speaker's identity* favors a large window for the overlap, therefore the increase in rate, that skews the probability density toward a more overlapped phasing, has the effect of increasing overlap (the more overlapped phasing is implemented).

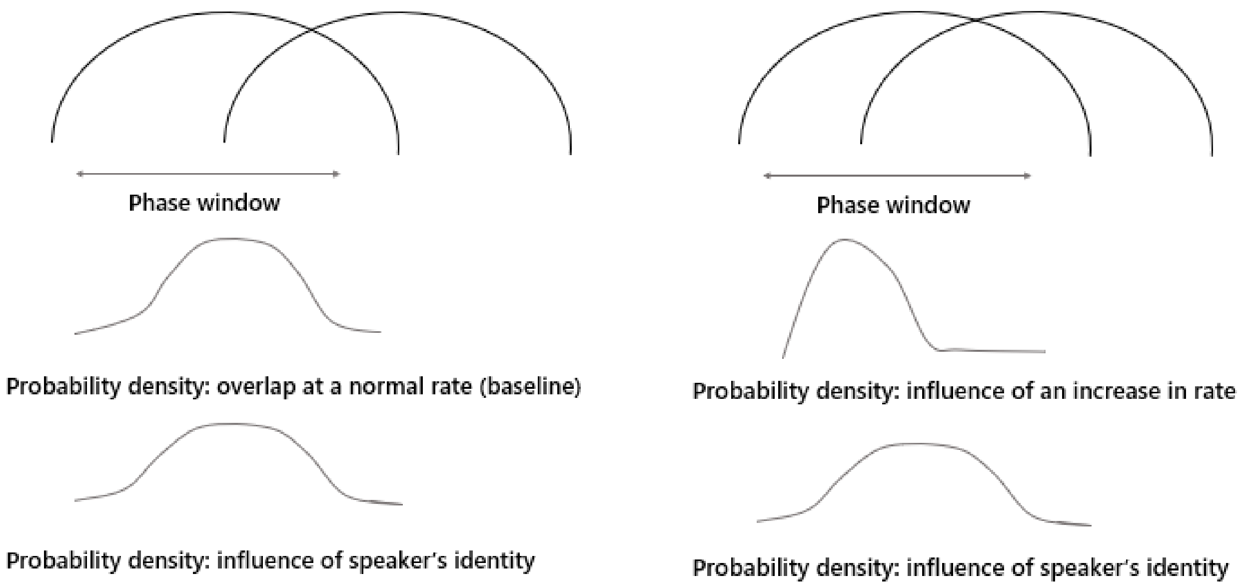

Figure 41. Phase windows for the overlap between /a/ and /i/ at a normal (baseline, on the left) and a fast tempo (increase in rate, on the right) for a speaker who increases coarticulation with the increase in rate. The influencer rate skews the probability density toward an increase in overlap, that is implemented.

In the last section of this discussion, I will address the results of the **MSD** study (chapter 5) on V-to-V and C-to-V coarticulation.

How to account for the coarticulation patterns observed in dysarthria?

Since the impairment in dysarthria is situated at the level of the execution of motor programs, and not at the level of planning, C-to-V coarticulation was expected to be more impaired than anticipatory V-to-V coarticulation in dysarthria. Our results do not support these predictions. Indeed, a reduction of V-to-V coarticulation is found for all groups of dysarthric speakers, ALS, Wilson Disease and Parkinson's Disease. Conversely, they present unimpaired C-to-V coarticulation. As for the temporal organization of speech, these three groups present a different behavior. Longer segmental durations are found in ALS and Wilson Disease for both the items on which V-to-V and C-to-V coarticulation are observed, while Parkinson's Disease speakers, on the other hand, show no difference in segmental durations with respect to the control group. This difference in the temporal organization of speech across groups can be attributed to the different dysarthria types associated to these pathologies. Indeed, a slowing of speech characterizes most dysarthria types, and has been reported for ALS and Wilson Disease, while hypokinetic dysarthria associated with Parkinson's Disease can be characterized by "rushes" in rate, i.e. abnormally rapid articulation rates (Darley *et al.,* 1975; Weismer, 1997; Berry, Darley, Aronson, & Goldstein, 1974; Ackermann, Konczak, & Hertrich, 1997). A reduction of V-to-V coarticulation and apparently unimpaired C-to-V coarticulation for these speakers can be accounted for by the temporal organization of their speech, together with impairment in the articulation of movements that characterizes dysathria. Indeed, a reduction of movement velocity and in movement displacement have been found in different dysarthria types (Weismer, 1997). Smaller movement amplitude has been documented especially for ALS (Watanabe, Arasaki, Nagata, & Shouji, 1994; Turner, Tjaden, & Weismer, 1995; Lee, Bell & Simmons, 2018; Lee & Bell, 2018), and Parkinson's Disease (Forrest, Weismer, & Turner 1989; Caligiuri, 1989; Audibert & Fougeron, 2012; Roland *et al.,* 2016; Mefferd & Dietrich, 2019). On the basis of the temporal and spatial characteristics of the speech in these dysarthria types, I will propose an interpretation for the results of ALS and Wilson Disease speakers on one side, and for Parkinson's Speakers on the other.

ALS and Wilson Disease

ALS and Wilson Disease are associated with mixed dysarthria, spastic/flaccid for ALS and ataxic, spastic, and/or hypo and hyperkinetic for Wilson Disease (Duffy, 2012; Pernon *et al.,* 2013; a recap of the dysarthria types in our population is in section 3.3.2.1). Speakers affected by ALS have been reported to present longer durations, slower tongue body movements, and a reduction of tongue displacement (e.g. Lee & Bell, 2018). Speakers affected by Wilson Disease have been reported to present longer durations, long silent pauses, imprecise articulation (Pernon *et al.,* 2013; Berry *et al.* 1974). A reduction in V-to-V coarticulation in these speakers could be accounted for by an effort to compensate for a difficulty in articulating smoothly the two vocalic gestures. This would lead them to produce longer and less overlapped movements, resulting in longer vowels and longer V-to-V lags, and less coarticulation. Similar patterns of C-to-V coarticulation in these speakers as for the controls can be accounted again for by longer movements and reduction in displacement. Indeed, these speakers could find difficult to produce the open vowel /a/ after the /ʃ/. In other terms, a difficulty in the phasing of the movements of the vowel and the consonant would cause the long consonant gesture to overlap more on the (long) vowel gesture, resulting in a degree of coarticulation similar to that of control speakers, even in presence of long vowel durations (see also the results of Hertrich & Ackermann, 1999, who reported unimpaired carryover coarticulation in ataxia, in presence of longer segmental durations, cited in section 3.3.2.3.1). Being the pattern of coarticulation and segmental durations exhibited by these speakers similar to the one showed for AoS, a question could be opened whether these speakers also restructure their speech over syllable-sized unit, as a response to movement perturbation.

Parkinson's Disease

A reduction in V-to-V coarticulation in Parkinson's Disease and similar patterns of C-to-V coarticulation as for control speakers can be accounted for by a reduction in tongue displacement (Mefferd & Dietrich, 2019) and in lip/jaw displacement (Forrest *et al.,* 1989). As pointed out in the results (section 5.5) these speakers show a lowering of F1, resulting in an undershoot of the /a/ vowel target, in both V_2 context, thus for both [a]i and [a]a. This undershoot, which is accompanied by short segmental durations for these speakers, can be seen as the result of overall smaller movements, but also as an abnormal overlap of the /p/ gesture on the vowel in /pa/. Similarly, the pattern of C-to-V coarticulation exhibited by these speakers can be related to this decrease in lip/jaw movement amplitude. Indeed, speakers produce more closed /a/ vowels, which would actually reflect an increase in coarticulation between the /ʃ/ consonant and the /a/ vowel. A tendency toward increased CV coarticulation attributed to a reduction in movement displacement has been previously reported in Parkinson's Disease (Tjaden, 2000; Roland *et al.*, 2016). In our data, a tendency toward an increase in the overlap between /ʃ/ and /a/ for these speakers is supported by the observation of the *F2-F1 compacity* values of p[a]

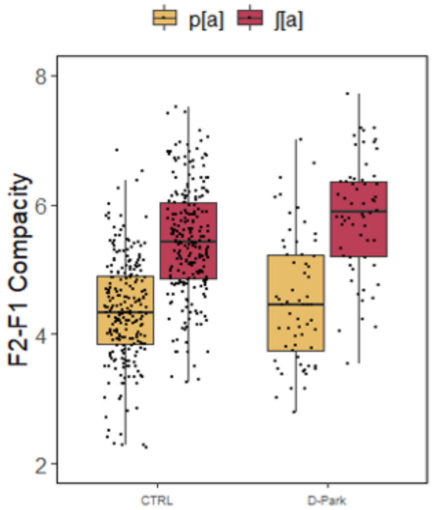

Figure 42. F2-F1 compacity of p[a], in yellow, and ʃ[a], in red, for CTRL and D-Park speakers. Compacity tends to be higher for D-Park speakers, especially in ʃ[a].

and ʃ[a] (Figure 42): it can be observed that the vowel targets /a/ in both contexts (but especially for ʃ[a]) tend to be less compact in Parkinson's Disease with respect to the vowels of control speakers. Higher *compacity* in ʃ[a] can reflect an increased effect of /ʃ/ on /a/ which translates in a lowering of F1 and a rise in F2. A tendency toward less *compacity* in p[a] could be due to a lowering of F1, reflecting a less open /a/ vowel due to the overlap of /p/ (Figure 42).

9.5 Final remarks and future perspectives

In this dissertation, the variations that V-to-V coarticulation undergoes under the influence of several factors such as pathology, speaker's identity and age, speech tempo and prosody have been investigated. The necessity to account for interspeaker variations in models of coarticulation has been emphasized, and one of the propositions of the literature that can be adapted in this respect has been discussed.

However, this work has some limitations, and many things can still be observed relatively to coarticulation. In listing these limitations and future perspectives, I will start from the methodology, to move to the single studies.

Methodological remarks:

- I have observed only one case of V-to-V coarticulation, that is, coarticulation of /a/ with /i/, and with only one intervening consonant. In this dissertation this allowed to focus on changes in coarticulation related to the different factors examined. However, in considering the possible interspeaker variability of coarticulation, future works will have to take into account how this interspeaker variability interact with different degrees of constraints of the intervening consonants.
- I have observed (almost) only anticipatory V-to-V coarticulation. There are still questions to address related to coarticulatory direction, and one open question is whether we would observe the same changes depending on speaker or condition in anticipatory and carryover coarticulation.

Remarks on the single studies:

- In the MSD study, I have compared anticipatory V-to-V and carryover C-to-V coarticulation, and I have interpreted our results as evidence that AoS speakers plan over syllable-sized units. Future work will have to compare intrasyllabic anticipatory and extrasyllabic anticipatory coarticulation in order to support or disprove this claim, and in order to explore possible patterns of differences in coarticulation depending on coarticulatory direction in patients.
- In the tempo and boundary studies, interspeaker variations in coarticulation have been highlighted depending on condition. One speaker in particular stood out as for coarticulatory behavior. Increasing the number of participants is necessary in order to test whether other participants share this pattern, or if other patterns of coarticulation depending on condition emerge.
- In the rate study, we relied on fixed slow and fast tempos to elicit the rate changes. However, if in the slow tempo all speakers decreased rate to a similar extent, differences in the articulation rates reached by the speakers at fast

tempo are reported. Moreover, one speaker did not increase her articulation rate. Future investigations would have to either increase the rate of the fast tempo condition, either rely on other tasks such as self-paced slowing and accelerating rates, previously used in the literature.

References

Abry, C., & Lallouache, T. (1995). Le MEM: un modèle d'anticipation paramétrable par locuteur: Données sur l'arrondissement en français. *Les Cahiers de l'ICP. Bulletin de la communication parlée*, (3), 85-99.

Abry, C., Lallouache, M. T., & Cathiard, M. A. (1996). How can coarticulation models account for speech sensitivity to audio-visual desynchronization? In *Speechreading by humans and machines* (pp. 247-255). Springer, Berlin, Heidelberg.

Ackermann, H., Konczak, J., & Hertrich, I. (1997). The temporal control of repetitive articulatory movements in Parkinson's disease. *Brain and language*, 56(2), 312-319.

Adams, S. G., Weismer, G., & Kent, R. D. (1993). Speaking rate and speech movement velocity profiles. *Journal of Speech, Language, and Hearing Research*, *36*(1), 41-54.

Agwuele, A., Sussman, H. M., & Lindblom, B. (2008). The effect of speaking rate onconsonant vowel coarticulation. *Phonetica*, *65*(4), 194-209.

Aichert, I., & Ziegler, W. (2004). Syllable frequency and syllable structure in apraxia of speech. *Brain and Language*, 88, 148–159.

Alario, F. X., and Caramazza, A. (2002). The production of determiners: evidence from French. *Cognition* 82, 179–223.

Albuquerque, L., Oliveira, C., Teixeira, A. J., Sa-Couto, P., & Figueiredo, D. (2019). Age-Related Changes in European Portuguese Vowel Acoustics. In *INTERSPEECH* (pp. 3965-3969).

Albuquerque, L., Oliveira, C., Teixeira, A., Sa-Couto, P., & Figueiredo, D. (2020). A comprehensive analysis of age and gender effects in European Portuguese oral vowels. *Journal of Voice*.

Albuquerque, L., Valente, A. R., Barros, F., Teixeira, A. J., Silva, S. S., Martins, P., & Oliveira, C. (2021). The age effects on EP vowel production: an ultrasound pilot study. In *IberSPEECH* (pp. 245-49).

Audibert, N. (2014, personal communication). Extract acoustic parameters from selected labels v14 [Praat script]

Audibert, N., & Fougeron, C. (2012). Distortions de l'espace vocalique : quelles mesures? Application à la dysarthrie. In *Actes des 29èmes Journées d'Etudes sur la Parole/JEP-TALN-RECITAL 2012* (pp. 217-224).

Auzou, P., & Rolland-Monnoury, V. (2019). BECD : Batterie D'évaluation Clinique de la Dysarthrie [Clinical Dysarthria Assessment Battery].

Barbier, G. (2016). *Contrôle de la production de la parole chez l'enfant de 4 ans : l'anticipation comme indice de maturité motrice* (Doctoral dissertation, Université Grenoble Alpes (ComUE)).

Barbier, G., Perrier, P., Payan, Y., Tiede, M. K., Gerber, S., Perkell, J. S., & Ménard, L. (2020). What anticipatory coarticulation in children tells us about speech motor control maturity. *Plos one*, *15*(4), e0231484.

Bartle-Meyer, C. J., & Murdoch, B. E. (2010). A kinematic investigation of anticipatory lingual movement in acquired apraxia of speech. *Aphasiology*, *24*(5), 623-642.

Barton, K. (2009) Mu-MIn: Multi-model inference. R Package Version 0.12.2/r18.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). LME4: Linear Mixed-Effects Models Using Eigen and S4. R Package version 1.1-4.

Beddor, P. S., Harnsberger, J. D., & Lindemann, S. (2002). Language-specific patterns of vowel-to-vowel coarticulation: Acoustic structures and their perceptual correlates. *Journal of Phonetics*, *30*(4), 591-627.

Bell-Berti, F., & Harris, K. S. (1981). A temporal model of speech production. *Phonetica*, *38*(1-3), 9-20.

Benguerel, A. P., & Cowan, H. A. (1974). Coarticulation of upper lip protrusion in French. *Phonetica*, *30*(1), 41-55.

Berry, W. R., Darley, F. L., Aronson, A. E., & Goldstein, N. P. (1974). Dysarthria in Wilson's disease. *Journal of Speech and Hearing Research*, 17(2), 169-183.

Berry, J. (2011). Speaking rate effects on normal aspects of articulation: Outcomes and issues. *Perspectives on Speech Science and Orofacial Disorders*, *21*(1), 15-26.

Bilodeau-Mercure, M., Kirouac, V., Langlois, N., Ouellet, C., Gasse, I., & Tremblay, P. (2015). Movement sequencing in normal aging: Speech, oro-facial, and finger movements. *Age*, *37*(4), 1-13.

Boehmke, B., & Greenwell, B. (2019). *Hands-on machine learning with R*. Chapman and Hall/CRC.

Boersma, P., & Weenink, D. (2021). Praat: doing phonetics by computer [computer program] (2011). *Version*, 5(3), 74.

Bombien, L., Mooshammer, C., Hoole, P., Kühnert, B., & Schneeberg, J. (2006, December). An EPG study of initial clusters in varying prosodic conditions. In *7th International Seminar on Speech Production (ISSP)* (pp. pp-35).

Boucher, K. R. (2007). *Patterns of anticipatory coarticulation in adults and typically developing children*. Master Dissertation, Brigham Young University.

Bourbon, A., & Hermes, A. (2020). Have a break: Aging effects on sentence production and structuring in French. In *12th International Seminar on Speech Production* (*ISSP*) (pp. 102-105).

Browman, C. P., & Goldstein, L. M. (1986). Towards an articulatory phonology. *Phonology*, *3*, 219-252.

Browman, C. P., & Goldstein, L. (1988). Some notes on syllable structure in articulatory phonology. *Phonetica*, *45*(2-4), 140-155.

Browman, C. P., & Goldstein, L. (1990). Tiers in articulatory phonology, with some implications for casual speech. *Papers in laboratory phonology I: Between the grammar and physics of speech*, 341-376.

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, *49*(3-4), 155-180.

Byrd, D. (1996a). Influences on articulatory timing in consonant sequences. *Journal of phonetics*, *24*(2), 209-244

Byrd, D. (1996b). A phase window framework for articulatory timing. *Phonology*, *13*(2), 139-169.

Byrd, D., & Tan, C. C. (1996). Saying consonant clusters quickly. *Journal of Phonetics*, *24*(2), 263-282.

Byrd, D. (2000). Articulatory vowel lengthening and coordination at phrasal junctures. *Phonetica*, *57*(1), 3-16.

Byrd D, & Saltzman, E. (2003). The elastic phrase: modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics*, 31,149–80

Byrd, D., & Choi, S. (2010). At the juncture of prosody, phonology, and phonetics—The interaction of phrasal and syllable structure in shaping the timing of consonant gestures. *Laboratory phonology*, *10*, 31-59.

Caçola, P., Roberson, J., & Gabbard, C. (2013). Aging in movement representations for sequential finger movements: a comparison between young-, middle-aged, and older adults. *Brain and cognition*, *82*(1), 1-5.

Caligiuri, M. P. (1989). The influence of speaking rate on articulatory hypokinesia in Parkinsonian dysarthria. *Brain and Language*, 36(3), 493-502.

Chicherio, C., Genoud-Prachex, T., Assal, F., & Laganaro, M. (2019). E-GeBAS: Electronic Geneva bedside aphasia scale. Computer Program.

Cho, T. (2004). Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics*, *32*(2), 141-176.

Cho, T. (2011). Laboratory phonology. *The continuum companion to phonology*, 343-368.

Cho, T. (2016). Prosodic boundary strengthening in the phonetics–prosody interface. *Language and Linguistics Compass*, *10*(3), 120-141.

Cho, T., Kim, D., & Kim, S. (2017). Prosodically-conditioned fine-tuning of coarticulatory vowel nasalization in English. *Journal of Phonetics*, *64*, 71-89.

Conklin, J. T., & Dmitrieva, O. (2020). Vowel-to-Vowel Coarticulation in Spanish Nonwords. *Phonetica*, *77*(4), 294-319.

Croot, K. (2002). Diagnosis of AOS: definition and criteria. *Seminars in speech and language* (Vol. 23, No. 04, 267-280).

D’alessandro, D., & Fougeron, C. (2021). Changes in Anticipatory VtoV Coarticulation in French during Adulthood. *Languages*, *6*(4), 181.

D’Alessandro, D., Bourbon, A., Fougeron, C. (2020). Effect of age on rate and coarticulation across different speech-tasks. In *12th International Seminar on Speech Production* (*ISSP*), 14-18.

D'Alessandro, D., Fougeron, C. (2018). Réduction de la coarticulation et vieillissement. *In Actes des 32ème Journées d’Études 491 sur la Parole,* 410-418.

Daniloff, R. G., & Hammarberg, R. E. (1973). On defining coarticulation. *Journal of Phonetics*, *1*(3), 239-248.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969a). Differential diagnostic patterns of dysarthria. *Journal of speech and hearing research*, *12*(2), 246-269.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969b). Clusters of deviant speech dimensions in the dysarthrias. *Journal of speech and hearing research*, 12(3), 462-496.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1975). *Motor speech disorders*. Saunders.

Didirková, I., Lancia, L., & Fougeron, C. (2020). Adaptations sur le F1 et le débit en réponse à diverses perturbations. In *6e conférence conjointe Journées d'Études sur la Parole (JEP, 33e édition), Traitement Automatique des Langues Naturelles (TALN, 27e édition), Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RÉCITAL, 22e édition). Volume 1: Journées d'Études sur la Parole* (163-171). ATALA; AFCP.

Dogil, G., & Mayer, J. (1998). Selective phonological impairment: a case of apraxia of speech. *Phonology*, *15*(2), 143-188.

Duez, D. (1992). Second formant locus-nucleus patterns: an investigation of spontaneous French speech. *Speech communication*, 495 11(4-5), 417-427.

Duffy, J. R. (2012). *Motor speech disorders: Substrates, differential diagnosis, and management*. Mosby; 3rd edition.

Eichhorn, J. T., Kent, R. D., Austin, D., & Vorperian, H. K. (2018). Effects of aging on vocal fundamental frequency and vowel formants in men and women. *Journal of Voice*, *32*(5), 644-e1.

Elvira-Garcia W. (2015). Extract_and_save_intervals, v.1. [Praat script]

Enderby, P. (2013). Disorders of communication: dysarthria. *Handbook of clinical neurology*, *110*, 273-281.

Engstrand, O. (1988). Articulatory correlates of stress and speaking rate in Swedish VCV utterances. *The journal of the Acoustical society of America*, *83*(5), 1863-1875.

Farnetani, E. (1990). VCV lingual coarticulation and its spatiotemporal domain. In *Speech production and speech modeling* (93-130). Springer, Dordrecht.

Farnetani, E., & Recasens, D. (2010). Coarticulation and connected speech processes. In *The handbook of phonetic sciences*, 316-352.

Flege, J. E. (1988). Effects of speaking rate on tongue position and velocity of movement in vowel production. *The Journal of the Acoustical Society of America*, *84*(3), 901-916.

Fletcher, A. R., McAuliffe, M. J., Lansford, K. L., & Liss, J. M. (2015). The relationship between speech segment duration and vowel centralization in a group of older speakers. *The Journal of the Acoustical Society of America*, *138*(4), 2132-2139.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research*, *12*(3), 189-198.

Forrest, K., Weismer, G., & Turner, G. S. (1989). Kinematic, acoustic, and perceptual analyses of connected speech produced by Parkinsonian and normal geriatric adults. *The Journal of the Acoustical Society of America*, 85(6), 2608-2622.

Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. The journal of the acoustical society of America, 101(6), 3728-3740.

Fougeron, C. (1999). Prosodically conditioned articulatory variations: a review. *UCLA Working Papers in Phonetics* 97. 1–74.

Fougeron, C. (2001). Articulatory properties of initial segments in several prosodic constituents in French. Journal of phonetics, 29(2), 109-135.

Fougeron, C., Delvaux, V., Menard, L., & Laganaro, M. (2018). The MonPaGe_HA database for the documentation of spoken French throughout adulthood. In

Proceedings of the Eleventh International Conference on Language Resources and Évaluation (LREC 2018).

Fougeron, C., Guitard-Ivent, F., & Delvaux, V. (2021). Multi-dimensional variation in adult speech as a function of age. *Languages*, *6*(4), 176.

Fourakis, M. (1991). Tempo, stress, and vowel reduction in American English. *The Journal of the Acoustical society of America*, *90*(4), 1816-1827.

Fowler, C. A. (1980). Coarticulation and theories of extrinsic timing. *Journal of phonetics*, *8*(1), 113-133.

Fowler, C. A. (1981). Production and perception of coarticulation among stressed and unstressed vowels. *Journal of Speech, Language, and Hearing Research*, *24*(1), 127-139.

Fowler, C. A., & Saltzman, E. (1993). Coordination and coarticulation in speech production. *Language and speech*, *36*(2-3), 171-195.

Fox, J., & Weisberg, S. (2018). *An R companion to applied regression*. Sage publications.

Friedman, J. H. (1991). Multivariate adaptive regression splines. *The annals of statistics*, 19(1), 1-67.

Gahl, S., & Baayen, R. H. (2019). Twenty-eight years of vowels: Tracking phonetic variation through young to middle age adulthood. *Journal of Phonetics*, *74*, 42-54.

Gay, T. (1978). Effect of speaking rate on vowel formant movements. *The journal of the Acoustical society of America*, *63*(1), 223-230.

Gay, T. (1981). Mechanisms in the control of speech rate. *Phonetica*, *38*(1-3), 148-158.

Goldman, J. P. (2011). EasyAlign: an automatic phonetic alignment tool under Praat. In *Interspeech'11, 12th Annual Conference of the International Speech Communication Association*.

Goozée, J. V., Lapointe, L. L., & Murdoch, B. E. (2003). Effects of speaking rate on EMA-derived lingual kinematics: a preliminary 508 investigation. *Clinical linguistics & phonetics*, 17(4-5), 375-381.

Grimme, B., Fuchs, S., Perrier, P., & Schöner, G. (2011). Limb versus speech motor control: A conceptual review. *Motor control*, *15*(1), 5-33.

Grosvald, M. (2009). Interspeaker variation in the extent and perception of long-distance vowel-to-vowel coarticulation. *Journal of Phonetics*, *37*(2), 173-188.

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. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological review*, *102*(3), 594.

Guenther, F. H., Hampson, M., & Johnson, D. (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychological review*, *105*(4), 611.

Guenther, F. H., Ghosh, S. S., Nieto-Castanon, A., & Tourville, J. (2006). A neural model of speech production. *Speech production: Models, phonetic processes and techniques*, 27-40.

Guitard-Ivent, F. (2018a). Effets de la durée vocalique et du locuteur sur le degré de coarticulation C-à-V en français : étude sur grands corpus. In *Actes des 32ème Journées d'Étude sur la Parole (JEP)*.

Guitard-Ivent, F. (2018b). *Coarticulation C-à-V en français : interaction avec le type de voyelle, la position prosodique et le style de parole* (Doctoral dissertation, Université Sorbonne Nouvelle-Paris 3).

Guitard-Ivent, F., Turco, G., & Fougeron, C. (2021). Domain-initial effects on C-to-V and V-to-V coarticulation in French: A corpus-based study. *Journal of Phonetics*, *87*, 101057.

Hardcastle, W. J. (1985). Some phonetic and syntactic constraints on lingual coarticulation during /kl/ sequences. *Speech Communication*, *4*(1-3), 247-263.

Henke, W. L. (1966). *Dynamic articulatory model of speech production using computer simulation.* Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Hertrich, I., & Ackermann, H. (1995). Coarticulation in slow speech: Durational and spectral analysis. *Language and Speech*, *38*(2), 159-187.

Hertrich, I., & Ackermann, H. (1999). Temporal and spectral aspects of coarticulation in ataxic dysarthria: An acoustic analysis. *Journal of Speech, Language, and Hearing Research*, *42*(2), 367-381.

Hertrich, I., & Ackermann, H. (2000). Lip–jaw and tongue–jaw coordination during rate-controlled syllable repetitions. *The Journal of the Acoustical Society of America*, *107*(4), 2236-2247.

Hirai, T., Tanaka, O., Koshino, H., & Yajima, T. (1991). Ultrasound observations of tongue motor behavior. *The Journal of Prosthetic Dentistry*, *65*(6), 840-844.

Hirst, D. (2011). Insert silence. [Praat script]

Hodge, M. M. (1989). *A comparison of spectral-temporal measures across speaker age: Implications for an acoustic characterization of speech maturation*. The University of Wisconsin-Madison.

Horton, W. S., Spieler, D. H., & Shriberg, E. (2010). A corpus analysis of patterns of age-related change in conversational speech. *Psychology and aging*, *25*(3), 708.

Iraci, M. (2017). *Vowels, consonants and co-articulation in Parkinson's Disease.* Unpublished PhD Dissertation. University of Salento, Lecce, Italy.

Jacewicz, E., Fox, R. A., O'Neill, C., & Salmons, J. (2009). Articulation rate across dialect, age, and gender. *Language variation and change*, *21*(2), 233-256.

Jacewicz, E., Fox, R. A., & Wei, L. (2010). Between-speaker and within-speaker variation in speech tempo of American English. *The Journal of the Acoustical Society of America*, *128*(2), 839-850.

Jang, J., Kim, S., & Cho, T. (2018). Focus and boundary effects on coarticulatory vowel nasalization in Korean with implications for cross-linguistic similarities and differences. *The Journal of the Acoustical Society of America*, *144*(1), EL33-EL39.

Joo, H., Jang, J., Kim, S., Cho, T., & Cutler, A. (2019). Prosodic structural effects on coarticulatory vowel nasalization in Australian English in comparison to American English. In *19th International Congress of Phonetic Sciences (ICPhS 2019)* (835-839). Australasian Speech Science and Technology Association Inc.

Jun, S. A., & Fougeron, C. (2000). A phonological model of French intonation. In *Intonation* (209-242). Springer, Dordrecht.

Jun, S., & Fougeron, C. (2002). Realizations of accentual phrase in French intonation. *Probus - International Journal of Latin and Romance Linguistics,* 14(1), 147-172.

Katz, W., Orth, U., Machetanz, J., & Schönle, P. (1990). Anticipatory labial coarticulation in the speech of German-speaking anterior aphasic subjects: Acoustic analyses. *Journal of Neurolinguistics*, *5*(2-3), 295-320.

Keating, P. A. (1990). The window model of coarticulation: articulatory evidence. *Papers in laboratory phonology I*, *26*, 451-470.

Keating, P., & Shattuck-Hufnagel, S. (2002). A prosodic view of word form encoding for speech production. *UCLA working papers in phonetics*, 112-156.

Kelso, J. S., & Tuller, B. (1981). Toward a theory of apractic syndromes. *Brain and Language*, *12*(2), 224-245.

Kent, R. D., & Moll, K. L. (1972). Cinefluorographic analyses of selected lingual consonants. *Journal of speech and hearing research*, *15*(3), 453-473.

Kent, R. D., & Rosenbek, J. C. (1983). Acoustic patterns of apraxia of speech. *Journal of Speech, Language, and Hearing Research*, *26*(2), 231-249.

Kent, R. D., Kent, J. F., Duffy, J., & Weismer, G. (1998). The dysarthrias: speech-voice profiles, related dysfunctions, and neuropathology. *Journal of Medical Speech-language pathology*.

Krivokapić, J. (2020). Prosody in articulatory phonology. *Prosodic theory and practice*.

Kuehn, D. P., & Moll, K. L. (1976). A cineradiographic study of VC and CV articulatory velocities. *Journal of phonetics*, *4*(4), 303-320.

Laganaro, M. (2008). Is there a syllable frequency effect in aphasia or in apraxia of speech or both? *Aphasiology*, *22*(11), 1191-1200.

Laganaro, M., Fougeron, C., Pernon, M., Levêque, N., Borel, S., Fournet, M.,... & Delvaux, V. (2021). Sensitivity and specificity of an acoustic-and perceptual-based tool for assessing motor speech disorders in French: The MonPaGe-screening protocol. *Clinical Linguistics & Phonetics*, *35*(11), 1060-1075.

Lee, J., & Bell, M. (2018). Articulatory range of movement in individuals with dysarthria secondary to amyotrophic lateral sclerosis. *American Journal of Speech-Language Pathology*, 27(3), 996-1009.

Lee, J., Bell, M., & Simmons, Z. (2018). Articulatory kinematic characteristics across the dysarthria severity spectrum in individuals with amyotrophic lateral sclerosis. *American Journal of Speech-Language Pathology*, *27*(1), 258-269.

Levelt, W. J., & Wheeldon, L. (1994). Do speakers have access to a mental syllabary? *Cognition*, *50*(1-3), 239-269.

Levelt, W. J. M., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*. 22, 1–38.

Li, H., Kim, S., & Cho, T. (2020). Prosodic structurally conditioned variation of coarticulatory vowel nasalization in Mandarin Chinese: Its language specificity and cross-linguistic generalizability. *The Journal of the Acoustical Society of America*, *148*(3).

Lindblom, B. (1963). Spectrographic study of vowel reduction. *The journal of the Acoustical society of America*, *35*(11), 1773-1781.

Liss, J. M., Weismer, G., & Rosenbek, J. C. (1990). Selected acoustic characteristics of speech production in very old males. *Journal of gerontology*, *45*(2), P35-P45.

Ma, L., Perrier, P., & Dang, J. (2015). Strength of syllabic influences on articulation in Mandarin Chinese and French: Insights from a motor control approach. *Journal of Phonetics*, *53*, 101-124.

Magen, H. (1984). Vowel-to-vowel coarticulation in English and Japanese. *The Journal of the Acoustical Society of America*, *75*(S1), S41-S41.

Manuel, S. (1999). Cross-language studies: Relating language-particular coarticulation patterns to other language-particular facts. *Coarticulation: Theory, data and techniques*, 179-198.

Martel-Sauvageau, V., & Tjaden, K. (2017). Vocalic transitions as markers of speech acoustic changes with STN-DBS in Parkinson's Disease. *Journal of communication disorders*, *70*, 1-11.

Matthies M, Perrier P, Perkell JS, & Zandipour M. (2001). Variation in anticipatory coarticulation with changes in clarity and rate. *Journal of Speech, Language, and Hearing Research*, 44(2):340-53.

McClean, M. D. (2000). Patterns of orofacial movement velocity across variations in speech rate. *Journal of Speech, Language, and Hearing Research*, *43*(1), 205-216.

McClean, M. D., & Tasko, S. M. (2003). Association of orofacial muscle activity and movement during changes in speech rate and intensity. *Journal of speech, language, and hearing research : JSLHR*, *46*(6), 1387–1400.

McNeil MR, Robin DA, Schmidt RA. (1997). Apraxia of speech: definition, differentiation, and treatment. In: McNeil MR, ed. *Clinical Management of Sensorimotor Speech Disorder*s. New York: Thieme. 311–344.

Mefferd, A. S., & Dietrich, M. S. (2019). Tongue-and jaw-specific articulatory underpinnings of reduced and enhanced acoustic vowel contrast in talkers with Parkinson's disease. *Journal of Speech, Language, and Hearing Research*, 62(7), 2118-2132.

Meyer, A., Roelofs, A., and Levelt, W. J. M. (2003). Word length effects in object naming: the role of a response criterion. *Journal of Memory and Language*. 48, 131–147.

Meynadier, Y. (2002). Interaction entre prosodie et (co) articulation linguopalatale en français. *Travaux Interdisciplinaires du Laboratoire Parole et Langage d'Aix-en-Provence (TIPA)*, (21), 236-243.

Meynadier, Y. (2004). Articulation et coarticulation aux frontières prosodiques en français. In *XXVè Journées d'Etude sur la Parole* (pp. 381-384).

Michel Lange, V., & Laganaro, M. (2014). Inter-subject variability modulates phonological advance planning in the production of adjective-noun phrases. *Frontiers in psychology*, *5*, 43.

Milborrow, S. 2021. Earth: Multivariate Adaptive Regression Splines. R Package version 5.3.1.

Mildner, V. (2018) Aspects of Coarticulation. *Challenges in analysis and processing of spontaneous speech.* MTA Nyelvtudományi Intézet, Budapest, pp. 27-48.

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.

Mok, P. P. (2010). Language-specific realizations of syllable structure and vowel-to-vowel coarticulation. *The Journal of the Acoustical Society of America*, *128*(3), 1346-1356.

Mok, P. P. (2013). Does vowel inventory density affect vowel-to-vowel coarticulation? *Language and Speech*, 56(2), 191-209.

Moon, S. J., & Lindblom, B. (1994). Interaction between duration, context, and speaking style in English stressed vowels. *The Journal of the Acoustical society of America*, *96*(1), 40-55.

Mücke, D., Thies, T., Mertens, J., & Hermes, A. (2020). Age-related effects of prosodic prominence in vowel articulation. In *12th International Seminar on Speech Production (ISSP)* (126-129).

Mücke, D., Hermes, A., & Tilsen, S. (2020). Incongruencies between phonological theory and phonetic measurement. *Phonology*, 37(1), 133-170.

Munhall, K. G., Ostry, D. J., & Parush, A. (1985). Characteristics of velocity profiles of speech movements. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(4), 457.

Nicolaidis, K. (1999). The influence of stress on V-to-V coarticulation: An electropalatographic study. *Proceedings of ICPHS* 14, 1087–1090.

Nijland, L., Maassen, B., Meulen, S. V. D., Gabreëls, F., Kraaimaat, F. W., & Schreuder, R. (2002). Coarticulation patterns in children with developmental apraxia of speech. *Clinical Linguistics & Phonetics*, *16*(6), 461-483.

Nittrouer, S., Studdert-Kennedy, M., & Neely, S. T. (1996). How children learn to organize their speech gestures: Further evidence from fricative-vowel syllables. *Journal of Speech, Language, and Hearing Research*, *39*(2), 379-389.

Noiray, A., Cathiard, M. A., Ménard, L., & Abry, C. (2011). Test of the movement expansion model: Anticipatory vowel lip protrusion and constriction in French and English speakers. *The Journal of the Acoustical Society of America*, *129*(1), 340-349.

Noiray, A., Abakarova, D., Rubertus, E., Krüger, S., & Tiede, M. (2018). How do children organize their speech in the first years of life? Insight from ultrasound imaging. *Journal of Speech, Language, and Hearing Research*, *61*(6), 1355-1368.

Noiray, A., Wieling, M., Abakarova, D., Rubertus, E., & Tiede, M. (2019). Back from the future: Nonlinear anticipation in adults' and children's speech. *Journal of Speech, Language, and Hearing Research*, *62*(8S), 3033-3054.

Oh, E. (2008). Coarticulation in non-native speakers of English and French: An acoustic study. *Journal of phonetics*, *36*(2), 361-384.

Öhman, S. E. (1966). Coarticulation in VCV utterances: Spectrographic measurements. *The Journal of the Acoustical Society of America*, *39*(1), 151-168.

Öhman, S. E. (1967). Numerical model of coarticulation. *The Journal of the Acoustical Society of America*, *41*(2), 310-320.

Oliveira, C., Valente, A. R., Albuquerque, L., Barros, F., Martins, P., Silva, S. S., & Teixeira, A. J. (2021). The Vox Senes project: a study of segmental changes and rhythm variations on European Portuguese aging voice. In *IberSPEECH* (pp. 135-38).

Ostry, D. J., & Munhall, K. G. (1985). Control of rate and duration of speech movements. *The Journal of the Acoustical Society of America*, *77*(2), 640-648.

Perkell, J. S., Zandipour, M., Matthies, M. L., & Lane, H. (2002). Economy of effort in different speaking conditions. I. A preliminary study of intersubject differences and modeling issues. *The Journal of the Acoustical Society of America*, *112*(4), 1627-1641.

Pernon, M., Trocello, J. M., Vaissière, J., Cousin, C., Chevaillier, G., Rémy, P., ... & Woimant, F. (2013). Le débit de parole du patient wilsonien dysarthrique peut-il être amélioré en condition de double tâche? *Revue neurologique,* 169(6-7), 502-509.

Pernon, M., Levêque, N., Delvaux, V., Assal, F., Borel, S., Fougeron, C.,... & Laganaro, M. (2020). MonPaGe, un outil de screening francophone informatise d'évaluation perceptive et acoustique des troubles moteurs de la parole (dysarthries, apraxie de la parole). *Rééducation orthophonique*, *281*, 171-97.

Perrier, P. (2012). Gesture planning integrating knowledge of the motor plant's dynamics: A literature review from motor control and speech motor control. *Speech planning and dynamics*, 191-238.

Perrier, P. (2014) "GEPPETO": A target-based model of speech production including optimal planning and physical modeling. In *Adventures in Speech Science*.

Perrier, P., & Ma, L. (2008). Speech planning for V1CV2 sequences: Influence of the planned sequence. In *ISSP 8th International Seminar on Speech Production (ISSP)* (69-72).

Ramig, L. A. (1983). Effects of physiological aging on speaking and reading rates. Journal of communication disorders, 16(3), 217-226.

Rastatter, M. P., & Jacques, R. D. (1990). Formant frequency structure of the aging male and female vocal tract. *Folia Phoniatrica et Logopaedica*, *42*(6), 312-319.

Recasens, D. (1984). Vowel-to-vowel coarticulation in Catalan VCV sequences. *The Journal of the Acoustical Society of America*, *76*(6), 1624-1635.

Recasens, D. (1989). Long range coarticulation effects for tongue dorsum contact in VCVCV sequences. *Speech Communication*, *8*(4), 293-307.

Recasens, D. (2002). An EMA study of VCV coarticulatory direction. *The Journal of the Acoustical Society of America*, *111*(6), 2828-2841.

Recasens, D. (2015). The effect of stress and speech rate on vowel coarticulation in catalan vowel–consonant–vowel sequences. *Journal of Speech, Language, and Hearing Research*, *58*(5), 1407-1424.

Recasens, D. (2018). Coarticulation. In *Oxford Research Encyclopedia of Linguistics*.

Recasens, D., Pallarès, M. D., & Fontdevila, J. (1997). The DAC model of lingual coarticulation. Evidence from VCV coarticulatory patterns. *Journal of the Acoustical Society of America*, *102*, 544-561.

Repp, B. H. (1986). Some observations on the development of anticipatory coarticulation. *The Journal of the Acoustical Society of America*, *79*(5), 1616-1619.

Robert, V., Wrobel-Dautcourt, B., Laprie, Y., & Bonneau, A. (2005, July). Inter speaker variability of labial coarticulation with the view of developing a formal coarticulation model for french. In *5th Conference on Auditory-Visual Speech Processing-AVSP 2005*.

Roland, V., Delvaux, V., Huet, K., Piccaluga, M., Haelewyck, M. C., & Harmegnies, B. (2016). Dynamique phonétique et contrôle moteur dans la maladie de Parkinson: analyse du contrôle de la production des glides (Speech dynamics and motion control in people with Parkinson's disease: analysis of glides' production). In *Actes de la conférence conjointe JEP-TALN-RECITAL 2016. volume 1: JEP* (pp. 211-219).

Rosenbek, J.C., Kent, R.D., & LaPointe, L.L. (1984). Apraxia of speech: An overview and some perspectives. In J.C. Rosenbek, M.R. McNeil, & A.E. Aronson (Eds.), *Apraxia of speech: Physiology-acoustics- linguistics- management* (1–72) San Diego: College-Hill Press.

Rosenblum, S., Engel-Yeger, B., Fogel, Y. 2013. Age-related changes in executive control and their relationships with activity 568 performance in handwriting. *Human movement science*, 322, 363-376.

Rubertus, E., & Noiray, A. (2018). On the development of gestural organization: A cross-sectional study of vowel-to-vowel anticipatory coarticulation. *PloS one*, *13*(9), e0203562.

Saltzman, E., & Kelso, J. A. (1987). Skilled actions: a task-dynamic approach. *Psychological review*, *94*(1), 84.

Saltzman, E. L., & Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. *Ecological psychology*, *1*(4), 333-382.

Saltzman, E., & Byrd, D. (2000). Task-dynamics of gestural timing: Phase windows and multifrequency rhythms. *Human Movement Science*, *19*(4), 499-526.

Saltzman E, Nam H, Krivokapic J, Goldstein L. 2008. A task-dynamic toolkit for modeling the effects of ′ prosodic structure on articulation. In *Proceedings of the 4th International Conference on Speech Prosody,* ed. PA Barbosa, S Madureira, C Reis, pp. 175–84. Boston: ICSA

Scarborough, R., & Zellou, G. (2013). Clarity in communication: "Clear" speech authenticity and lexical neighborhood density 570 effects in speech production and perception. *The Journal of the Acoustical Society of America,* 134(5), 3793-3807.

Schnur, T. T. (2011). Phonological planning during sentence production: beyond the verb. *Frontiers in Psychology,* 2:319.

Schnur, T. T., Costa, A., and Caramazza, A. (2006). Planning at the phonological level during sentence production. *Journal of Psycholinguistic Research,* 35, 189–213.

Schriefers, H., and Teruel, E. (1999). Phonological facilitation in the production of two-word utterances. *European Journal of Cognitive Psychology*, 11, 17–50.

Southwood, M. H., Dagenais, P. A., Sutphin, S. M., & Garcia, J. M. (1997). Coarticulation in apraxia of speech: A perceptual, acoustic, and electropalatographic study. *Clinical linguistics & phonetics*, *11*(3), 179-203.

Staiger, A., & Ziegler, W. (2008). Syllable frequency and syllable structure in the spontaneous speech production of patients with apraxia of speech. Aphasiology, 22, 1201–1215.

Tabain, M. (2003). Effects of prosodic boundary on/aC/sequences: acoustic results. *The Journal of the Acoustical Society of America*, *113*(1), 516-531.

Team, R. C. (2019). 2020. *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria: Available at: https://www. R-project. org/.[Google Scholar].*

Tjaden, K. (2000). An acoustic study of coarticulation in dysarthric speakers with Parkinson disease. *Journal of speech, language, and hearing research*, *43*(6), 1466-1480.

Tjaden, K. (2003). Anticipatory coarticulation in multiple sclerosis and Parkinson's disease. *Journal of speech, language, and hearing research*, 2003.

Tjaden, K., & Weismer, G. (1998). Speaking-rate-induced variability in F2 trajectories. *Journal of Speech, Language, and Hearing Research*, *41*(5), 976-989.

Tjaden K, & Wilding G. E. (2005). Effect of rate reduction and increased loudness on acoustic measures of anticipatory coarticulation in multiple sclerosis and Parkinson's disease. *Journal of Speech, Language, and Hearing Research,* 261-77.

Torre III, P., & Barlow, J. A. (2009). Age-related changes in acoustic characteristics of adult speech. *Journal of communication disorders*, *42*(5), 324-333.

Traunmüller, H. (1990). Analytical expressions for the tonotopic sensory scale. *The Journal of the Acoustical Society of America*, *88*(1), 97-100.

Tuller, B., & Story, R. S. (1988). Anticipatory and carryover coarticulation in aphasia: An acoustic study. *Cognitive Neuropsychology*, *5*(6), 747-771.

Turner, G. S., Tjaden, K., & Weismer, G. (1995). The influence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research*, 38(5), 1001-1013.

Turco, G., Fougeron, C., & Audibert, N. (2016). Que nous apprennent les gros corpus sur l'harmonie vocalique en français?(What can we learn from big speech corpora about French vowel harmony?). In *Actes de la conférence conjointe JEP-TALN-RECITAL 2016. volume 1: JEP* (571-579).

Turvey, M. T. (1990). Coordination. *American psychologist*, 45(8), 938.

van den Heuvel, H., Cranen, B., & Rietveld, T. (1996). Speaker variability in the coarticulation of/a, i, u. *Speech communication*, *18*(2), 113-130.

Van der Merwe, A. (1997). A theoretical framework for the characterization of pathological speech sensorimotor control. *Clinical management of sensorimotor speech disorders*, *2*, 3-18.

Van Der Merwe, A. (2021). New perspectives on speech motor planning and programming in the context of the four-level model and its implications for understanding the pathophysiology underlying apraxia of speech and other motor speech disorders. *Aphasiology*, *35*(4), 397-423.

Van Son, R. J., & Pols, L. C. (1992). Formant movements of Dutch vowels in a text, read at normal and fast rate. *The Journal of the Acoustical society of America*, *92*(1), 121-127.

Volenec, V. (2015). Coarticulation. *Phonetics: Fundamentals, Potential Applications and Role in Communicative Disorders*, 47-86. Nova: New York

Watanabe, S., Arasaki, K., Nagata, H., & Shouji, S. (1994). Analysis of dysarthria in amyotrophic lateral sclerosis--MRI of the tongue and formant analysis of vowels. *Rinsho shinkeigaku= Clinical neurology*, 34(3), 217-223.

Weismer, G. (1997). Motor speech disorders. In W. J. Hardcastle and J. Laver (Eds), *The Handbook of Phonetic Sciences*. Cambridge, MA: Blackwell, 191–219.

Weismer, G., & Berry, J. (2003). Effects of speaking rate on second formant trajectories of selected vocalic nuclei. *The Journal of the Acoustical Society of America*, *113*(6), 3362-3378.

Whalen, D. H. (1990). Coarticulation is largely planned. *Journal of Phonetics*, *18*(1), 3-35.

Whalen, D. H., & Chen, W. R. (2019). Variability and central tendencies in speech production. *Frontiers in Communication*, *4*, 49.

Whiteside, S. P., & Varley, R. A. (1998a). A reconceptualisation of apraxia of speech: A synthesis of evidence. *Cortex*, *34*(2), 221-231.

Whiteside, S. P., & Varley, R. A. (1998b). Coarticulation in apraxia of speech: an acoustic study of non-words. *Logopedics Phoniatrics Vocology*, *23*(4), 155-163.

Winkler, R., Ma, L., & Perrier, P. (2010, September). A model of optimal speech production planning integrating dynamical constraints to achieve appropriate articulatory timing. In *CPMSP2 2010-Cognitive and Physical Models of Speech Production, Speech Perception and Production-Perception Interaction-Part III Planning and Dynamics* (44-48).

Winter, B. (2019). *Statistics for linguists: An introduction using R*. Routledge.

Xue, S. A., & Hao, G. J. (2003). *Changes in the Human Vocal Tract Due to Aging and the Acoustic Correlates of Speech Production. Journal of Speech Language and Hearing Research*, 46(3), 689.

Zellou, G. (2017). Individual differences in the production of nasal coarticulation and perceptual compensation. *Journal of Phonetics*, *61*, 13-29.

Zharkova, N. (2017). Voiceless alveolar stop coarticulation in typically developing 5-year-olds and 13-year-olds. *Clinical linguistics & phonetics*, *31*(7-9), 503-513.

Zharkova, N., Hewlett, N., & Hardcastle, W. J. (2012). An ultrasound study of lingual coarticulation in/sV/syllables produced by adults and typically developing children. *Journal of the International Phonetic Association*, *42*(2), 193-208.

Ziegler, W. (2008). Apraxia of speech. *Handbook of clinical neurology*, *88*, 269-285.

Ziegler, W. (2009). Modelling the architecture of phonetic plans: Evidence from apraxia of speech. *Language and Cognitive Processes*, *24*(5), 631-661.

Ziegler, W., & Von Cramon, D. (1985). Anticipatory coarticulation in a patient with apraxia of speech. *Brain and Language*, *26*(1), 117-130.

Ziegler, W., & Von Cramon, D. (1986a). Disturbed coarticulation in apraxia of speech: Acoustic evidence. *Brain and Language*, *29*(1), 34-47.

Ziegler, W., & von Cramon, D. (1986b). Timing deficits in apraxia of speech. *European archives of psychiatry and neurological sciences*, *236*(1), 44-49.

Ziegler, W., Aichert, I., & Staiger, A. (2012). Apraxia of speech: Concepts and controversies. *Journal of speech, language, and hearing research*, 55, S1485-S1501

Ziegler, W., Lehner, K., Pfab, J., & Aichert, I. (2020). The nonlinear gestural model of speech apraxia: Clinical implications and applications. *Aphasiology*, *35*(4), 462-484.

Appendix A. Speech material for the **MSD** and **age** studies with English translation

Lundi, le chat, le loup et Papa vont à Bali. Les copains sont tout contents.	On Monday, the cat, the wolf and Dad go to Bali. The friends are very happy.
Mardi, Papi y va aussi. Il dit: "Je n'ai pas un sou! Qui va prendre soin de moi?" "Moi!" dit le chat, "moi!" dit le loup. "Vous?", Papi réfléchit.	Tuesday, Grandpa goes too. He says: "I don't have any money! Who will take care of me?" "Me!" says the cat, "me!" says the wolf. "You?", Grandpa thinks.
Mercredi, Papi dit: "Toi, le chat, tu es doux tu es chou, tu n'as pas de poux! Mais pas ce loup: il a une cape rouge et je n'aime pas ce gars-là!".	Wednesday, Grandpa says: "You, the cat, you are sweet you are nice, you don't have lice! But not that wolf: he has a red cape and I don't like that guy!".
Jeudi, le chat et Papi se baladent à Bali. Papa glisse! Aïe! Ouille! Son cou craque, son coulde claque, c'est la débâcle!.	“On Thursday, the cat and Grandpa go for a walk in Bali. Dad slips! Ouch! Ouch! His neck cracks, his elbow snaps, it's a debacle!
Vendredi, Papa a mal. Il pleure, il crie! "Toi, Papi, aide-moi, trouve le nain!" "Un nain? On n'en a jamais vu par ici?!".	“Friday, Dad is in pain. He cries, he screams! "You, Grandpa, help me, find the dwarf!" "A dwarf? We have never seen one around here?!".

Samedi matin, le chat va voir son ami le loup et lui dit: "Aide-moi à soigner Papa!".	Saturday morning, the cat goes to see his friend the wolf and says: "Help me to look after Dad!
Samedi soir, le loup lui donne sa recette magique: "Coupe un oignon, cache-le sous la souche, et lorsque le lilas fleurira, Papa sera guéri!"	Saturday night, the wolf gives him his magic recipe: "Cut an onion, hide it under the stump, and when the lilac blooms, Dad will be cured! Abracadabra, that's it, we did it!
Dimanche, le chat tout doux, le loup magicien, Papa et Papi quittent Bali. Les copains sont tout contents.	On Sunday, the soft cat, the wizard wolf, Dad and Grandpa leave Bali. The friends are very happy.”

Appendix B. Speech material for the **boundary** study. List of sentences per boundary type.

Condition	Target sequence	Sentence 1	Sentence 2	Sentence 3
Within word	*papi*	*Quand papi l'aura fait, il partira.* "When grandpa does it, he will leave"	*Quand papi l'aura vu, il nous croira.* "When grandpa sees it, he will believe us"	*Quand papi l'aura pris, il rentrera.* "When grandpa gets it, he will come home"
	papa	*Quand Papa s'en va au sport, il baille.* "When dad goes to the gym, he yawns"	*Quand papa s'en va courir, il râle* "When dad goes for a run, he moans"	*Quand Papa s'en va nager, il traine.* "When dad goes for a swim, he hangs around"
Word boundary	pa#pi	*Papa pilote un avion de ligne.* "Dad pilots an airliner"	*Papa pilote un bateau fluvial.* "Dad pilots a river boat"	*Papa pilote un hélicoptère.* "Dad pilots an helicopter"
	pa#pa	*Papa passe par chez toi en voiture.*	*Papa passe par chez toi en vélo.*	*Papa passe par chez toi en camion.*

		"Dad drives by your house"	"Dad passes by your house on his bike"	"Dad passes by your house in his truck"
Clause Boundary	pa##pi	*Papa, pilote à Marseille, rentre tard.* "Dad, a pilot in Marseille, comes home late"	*Papa, pilote à Bordeaux, rentre tard.* "Dad, a pilot in Bordeaux, comes home late"	*Papa, pilote à Toulouse, rentre tard.* "Dad, a pilot in Toulouse, comes home late"
	pa##pa	*Papa, passant par Toulon, l'a vu.* "Dad, passing by Toulon	*Papa, passant par Nanterre, l'a vu.* "Papa, passing by Nanterre, saw it"	*Papa, passant par Roskoff, l'a vu.* "Dad, passing by Roskoff, saw it"

Appendix C. Participants table for the perception experiment (**boundary study**)

Listener ID	**Age**	**Sex**	**Native language**	**Etude Ortho/Ling**	**Judged speaker**	**type**
AY	26	F	French	Yes	all	internal
DDA	31	F	Italian	Yes	all	internal
CF	51	F	French	Yes	all	internal
A3_1	20	F	French	Yes	Speaker5	external
A3_2	27	F	French	Yes	Speaker5	external
A3_3	36	M	French	No	Speaker5	external
A5_1	55	M	Kurdish	No	Speaker4	external
A5_2	23	F	French	Yes	Speaker4	external
A5_3	25	F	French	Yes	Speaker4	external
A3_1	23	M	French	No	Speaker3	external
A3_2	21	F	French	Yes	Speaker3	external
A3_3	41	F	French	Yes	Speaker3	external
A2_1	66	M	French	No	Speaker1	external
A2_2	22	F	French	Yes	Speaker1	external
A2_3	66	F	Polish	No	Speaker1	external
A1_3	36	F	French	No	Speaker2	external
A1_1	22	F	French	Yes	Speaker2	external
A1_2	24	F	French	Yes	Speaker2	external

Variations individuelles dans la coarticulation anticipatoire V-à-V : effet des Troubles Moteurs de la Parole, de l'age, de changements de tempo et du type de frontière.

La coarticulation anticipatoire se réfère à l'anticipation des mouvements articulatoires pour la réalisation de cibles de parole à venir et peut être considérée comme un indice de planification. Dans quatre études, la coarticulation anticipatoire V-à-V est étudiée dans différents Troubles Moteurs de la Parole, i.e. Apraxie de la Parole et Dysarthrie associée à la SLA, la maladie de Wilson, et la maladie de Parkinson (et comparée à la coarticulation C-à-V), chez des adultes âgés de 20 à 93 ans, et dans un groupe restreint de locutrices dans différentes conditions de parole : tempo lent, rapide et normal, dans un mot, à travers une frontière de mot et de proposition relative.Les résultats montrent une réduction de la coarticulation V-à-V dans l'Apraxie de la Parole et la Dysarthrie, qui pourrait être expliquée par des déficits spécifiques à ces pathologies. Une réduction non-linéaire de la coarticulation avec l'âge semble liée à un ralentissement du débit jusqu'à 70 ans, alors qu'une relation directe n'est pas trouvée pour les locuteurs plus âgés. Les différences inter-individuelles de coarticulation en réponse aux changements de tempo suggèrent que la relation entre la coarticulation et le débit articulatoire est spécifique au locuteur. Des variations inter-individuelles de coarticulation sont trouvées aussi en fonction du type de frontière et ne sont que partiellement expliquées par le phrasé prosodique. Ces résultats sont discutés selon deux axes, l'un traitant de la taille des unités de planification motrice dans la parole, et l'autre discutant comment peuvent êtres modélisées les variations de coarticulation en fonction du locuteur et de la population.

Mot clés : coarticulation, AoS, Dysarthrie, âge, planification motrice, débit

---

Interspeaker variations in V-to-V coarticulation: effects of Motor Speech Disorders, age, speech tempo changes, and boundary type.

Anticipatory coarticulation refers to the anticipation of articulatory movements for the achievement of forthcoming speech goals and can be considered an index of planning in speech. In four investigations, anticipatory V-to-V coarticulation is investigated in different Motor Speech Disorders, i.e. AoS and Dysarthria associated with ALS, Wilson Disease, and Parkinson's Disease (and compared to C-to-V coarticulation), in adults spanning 20 to 93 years old, and in a small set of speakers in different conditions: at a slow, fast, and habitual speech tempo, and within a word, across word boundary, and across clause boundary.The results show a reduction of V-to-V coarticulation in AoS and Dysarthria, which could be accounted for by disorder-specific impairment. A non-linear reduction of coarticulation is found with age. If this reduction can be seen in relation to a slowing of speech in speakers aged 20 to 70, with age-specific patterns of covariation, a relationship between these two factors is not found after 70 y.o.a. Individual responses to changes in speech tempo suggest that the relationship between coarticulation and articulation rate is speaker-specific. Interspeaker variations in V-to-V coarticulation are found also depending on boundary type and are only partially explained by specificities in the prosodic phrasing.These results are discussed along two axes, one discussing the size of the planning units in speech, i.e. units over which the movements for the achievement of a string of speech targets are planned, and the other discussing how speaker-specific and population-specific variations in coarticulation can be modeled in a coproduction account of coarticulation.

Keywords: coarticulation, AoS, Dysarthria, age, motor planning, rate

Université Sorbonne Nouvelle, Ecole doctorale 622 – Sciences du Langage

Maison de la Recherche, 4 rue des Irlandais, 75005 Paris