Vowel Nasalization in German

A real-time MRI Study

Esther Kunay

München 2022

Vowel Nasalization in German
A real-time MRI Study

Inaugural-Dissertation
zur Erlangung des Grades eines Doktors der Philosophie
der Ludwig-Maximilians-Universität
München

vorgelegt von
Esther Kunay

aus
Walsrode

2022

Erstgutachter:	Prof. Dr. Philip Hoole
Zweitgutachter:	Prof. Dr. Jonathan Harrington
Tag der mündlichen Prüfung:	09.11.2021

Acknowledgements

I would sincerely like to thank all those people who have contributed to this thesis and provided their support to me.

First of all I would like to thank my supervisor Phil Hoole for his advice and useful comments. I am grateful that he accompanied this project with his expertise and so many ideas. He also provided numerous resources that I was allowed to use for my research.

I also thank Jonathan Harrington for the fruitful discussions and thoughts. To both I would like to express my gratitude for the trust they showed me right from the beginning and throughout the project.

I would like to thank Jens Frahm for this unique collaboration with the Biomedical NMR research group at the MPI for Biophysical Chemistry and the flexibility that made this amazing number of MRI measurements possible. Thanks to Dirk Voit, Kurt Böhm and Jost Kollmeier for their patience and technical support and their readiness for testing new tools and ideas. Special thanks go to Arun Joseph, who conducted all the MRI measurements and who encouraged me in following this path.

My thanks go also to Stavros Skopeteas and Götz Keydana for their willing support at the Institute of Linguistics at the University Göttingen.

I am thankful to Chris Carignan and to Miriam Oschkinat for their technical help and tips. I would also like to thank Jasmin Mues and Pauline Sander for assisting me in recruiting our participants and helping me out with the data segmentation. I thank all participants who patiently took part in the MRI measurements and also in the perception test. Furthermore, I want to thank all people of the PhD colloquium at the Munich IPS for their critical and valuable comments. My thanks go also to Ulrike Vallender-Kalus for her advice and information on the formal issues.

Special thanks go to Garrett for patiently proofreading this work and for his mental support.

Besonderer Dank gebührt meinen Eltern, die mich stets unterstützt haben, meiner Schwester, die mich durch besondere Zeiten begleitet hat und meinen Brüdern, die mich ermutigt haben, diesen Weg zu gehen. Danken möchte ich auch Matthias für die unzähligen Gespräche und die offenen Türen sowie Martin, Sebastian, Flo und Maren, die mir eine große Hilfe waren. Ebenso bin ich Marie und Nico dankbar für die kleinen und großen Freuden in dieser Zeit.

Contents

List of Figures

List of Tables

List of Abbreviations

C	oral consonant
CT	computer tomography
CVCV	consonant-vowel-consonant-vowel
CVNC	consonant-vowel-nasal-consonant
CVNCC	consonant-vowel-nasal-consonant-consonant
CVNCV	consonant-vowel-nasal-consonant-vowel
CVNV	consonant-vowel-nasal-vowel
EMA	electromagnetic articulography
EMG	electromyography
EPG	electropalatography
F	fricative
f0	fundamental frequency
F1	first formant
F2	second formant
F3	third formant
FLASH	fast low angle shot (MRI sequence)
Fr.	French
FSE	fast spin echo (MRI sequence)
GRE	gradient echo (MRI sequence)
It.	Italian
Lat.	Latin
MR	magnetic resonance
MRI	magnetic resonance imaging
N	nasal consonant
NMR	nuclear magnetic resonance
OVL	overall velum lowering
Port.	Portuguese
PC1	first principal component
PCA	principal component analysis
RF	radio frequency
ROI	region of interest

rt-MRI	real-time magnetic resonance imaging
SE	spin echo (MRI sequence)
TE	echo time
TR	repetition time
V	vowel
Ṽ	nasalized vowel
VF	vowel-fricative
VLP	Vowel Length Parameter
VN	vowel-nasal
VNC	vowel-nasal-consonant
VNF	vowel-nasal-fricative
VP	velopharyngeal port
VT	vocal tract
VTA	vocal tract aperture function

Chapter I

Introduction

The aim of this thesis is to shed light on the basic principles of velum movements during vowels and vowel-consonant sequences in Standard German. Insights into the temporal and spatial adjustments of the velar behaviour during these specific sound sequences may contribute to our understanding of how the evolution of contrastive vowel nasalization is initiated. About one fifth of the world's languages exhibit nasal vowels that have a phonemic status (Maddieson, 1984, 2007): the presence or absence of nasality on the vowel changes the lexical meaning of a word, as is attested for e.g. several Romance and Hindi languages, numerous Chinese varieties such as the Min and Wu dialect families as well as many West African languages. In many cases, these nasal vowels have not always been part of the specific vocalic system, but developed from older language stages in which the vowel was followed or preceded by a nasal consonant. Over time, the nasal was lost, whereas the vowel attained contrastive nasality. Although much experimental research has been conducted to understand the basic function of the articulatory gestures involved during the production of these sounds, still not much is known about the articulatory and perceptual factors that are crucial for the initiation of this specific sound change. This thesis aims at contributing further insights into the interplay between the tongue and the soft palate, or velum, during vowels and nasal consonants in natural word productions. Studies on this issue are typically concerned with languages exhibiting contrastive nasal vowels in their phonemic systems or languages with strong coarticulatory nasality during pre-nasal vowels. In contrast, Standard German may not be famous for its vowel nasality, but it is still quite suitable for this research question, because it is in a stage that is probably far away from the evolution of contrastive nasal vowels[1]. This, in turn, allows for investigating the very basic principles of the spatial and temporal interactions between the velar and lingual gestures during vowel-nasal sequences followed by different consonantal contexts, from which conclusions may be drawn with respect to the articulatory and perceptual needs that are required to initiate a change from coarticulatory to contrastive nasalization.

Furthermore, most of the studies in prior research on velum movement patterns involved only a handful of participants, which is due to the high cost and effort of running these kinds of experiments, because investigating the soft palate commonly requires a highly invasive measurement technique that relies on patience and good will on the part of the participant. While every laboratory situation constitutes a more or less unnatural setting for the participant, the investigation of the soft palate is especially challenging. For example, measurement techniques like electromyography (EMG) or electromagnetic articulography (EMA) involve methods in which fine needles are inserted into the relevant muscles, or, respectively, small electrodes are glued to the articulators. Apart from the fact that such experiments are quite uncomfortable for the speaker, it is questionable how valid the speech data can be considered from a somewhat restricted articulatory organ when the slightest

[1] In some regions of Germany nasal vowels do exist, e.g. in a couple of Alemannic dialects. The study at hand, however, is explicitly concerned with Standard German.

movement differences might be essential during sound production. As an alternative to the invasive techniques, non-invasive imaging devices generally allow the participant to speak without any restriction, but often also come with several drawbacks. For example, while ultrasound is well suited to capturing the tongue movements, it is difficult to receive reliable data for velum movements because the air at the tongue surface reflects the ultrasound beam back to the transducer, such that the tissue beyond the tongue-air boundary cannot be appropriately displayed. Other imaging techniques may involve harmful radiation, for which reason they are not suitable for speech studies with long measurement times. During the past few decades, however, the development of magnetic resonance tomography (MRI) has proceeded dramatically, a non-invasive imaging technique that is not based on radiation. While some early speech studies used MRI to describe the configuration of the articulatory organs during isolated and sustained sounds, modern MRI techniques allow for imaging fluent speech in real time with up to 100 frames per second (fps). The drawbacks for speech studies, however, are twofold: on the one hand, MRI scanners are primarily used in hospitals for diagnostics, making it difficult to conduct a scientific study with many participants. For each measurement, skilled MRI personnel is required, and timing slots are rare. In addition, MRI measurements are usually expensive, so a study with more than a handful of participants might exceed budget limits. On the other hand, common MRI scanners require a supine position, which may confound the speech gestures during articulation due to gravity effects. With respect to the first concern, the data obtained for this thesis can be considered extraordinary because they involve articulatory data from more than 30 native speakers of Standard German, who were measured via real-time MRI with a high temporal resolution of 19.98 ms (50.05 fps) and a spatial resolution of 1.41x1.41 mm in-plane resolution (slice thickness 8 mm). This study was possible thanks to the Max Planck Institute of Biophysical Chemistry, Göttingen, Germany, which is specialized in the research and improvement of real-time MRI technology and has developed a groundbreaking technology that allows for fast imaging and rapid reconstruction (Uecker et al., 2010, 2012). The second concern, i.e. speaking in a supine position, has been sporadically considered by several phoneticians in the past with varying results. Some report small position changes of the tongue due to gravity effects (Kitamura et al., 2005; Perry, 2011b). More significant, however, is the finding that atypical articulatory patterns are more facilitated by unnaturally long sustained sounds than by slight gravitational effects in a supine position (Engwall, 2013). This issue is considered in more detail in section 1.5.1.

Overall, capturing the articulatory interactions of the lips, tongue, velum and larynx in a mid-sagittal plane during fluent speech offers the chance to examine across a relatively large number of participants whether specific movement patterns occur systematically and especially, how the gestures of the soft palate are affected by segmental factors such as vowel length, vowel quality or the post-vocalic context or by prosodic factors, such as stress and speaking rate. In addition to those parts considering the articulatory aspect, one chapter of this thesis is concerned with the perception of vowel nasalization by the same speakers who participated in the MRI study. The question is addressed to what extent German speaker-turned-listeners perceive vowel nasalization in a given set of artificial stimuli and also whether a link is present between the production and perception of vowel nasalization within one individual language user.

In the following, the nature of vowel nasalization and the general velopharyngeal mechanism are outlined as well as the predominant phonetic accounts regarding how sound change may be initiated by coarticulatory effects in production and perception. Some prior research on the segmental and prosodic parameters affecting the temporal and spatial movement patterns of the soft palate is reviewed, followed by a summary of the measurement techniques that are utilized to track the movements of the soft palate. Since the data at hand were acquired by means of real-time MRI, this technology is delineated in more detail. Furthermore, as all of the MRI experiments presented in this thesis refer to individual data subsets of the same overall corpus obtained via the same method, the procedure of image recording and analysis is explicitly outlined for all articulatory experiments discussed in the specific chapters.

1.1 Vowel nasalization

Elevating the soft palate, or velum, during speech and swallowing is the result of a complex interplay of the activation and control of numerous oropharyngeal and velopharyngeal muscles. In addition, during the act of speaking, velar gestures must be coordinated with other articulatory organs that involve muscle control of the lips, tongue and larynx as well as the respiratory musculature. During the production of sound sequences in or across words, the articulatory gestures for each sound are commonly not executed in isolation from each other, but are affected by the articulatory gestures involved in the production of the surrounding sounds. For example, the acoustic impression of a coronal stop in the German word *Klempner* 'plumber', often pronounced and perceived as [klɛmpt^{n}nɐ], is the result of an early coronal contact of the tongue blade for the nasal stop [n], which under normal circumstances is accompanied by the lowering of the soft palate to allow the air stream to escape through the nose. In the case of [klɛmpt^{n}nɐ], however, the gesture for the tongue tip and the velum do not occur simultaneously but are produced with a slight temporal shift: the soft palate is still raised during [p], while the tongue tip already induces the coronal closure for [n]. The consequence of these overlapping gestures during the transition from [p] to [n] is a closed velopharyngeal port with a simultaneously coronal closure. This articulatory configuration is typical for the oral coronal stop [t], which in this case is released by lowering the soft palate for [n].

In a similar way, such overlapping gestures of the velum and the tongue play a role during the production of sound sequences that involve vowels surrounded by nasal sounds:

these vowels typically become nasalized to some extent. In general, the phonetic term ‘nasalization’ describes the partial or exclusive usage of the nasal cavity during the production of speech sounds. Basically, when producing a nasal stop like [m], [n] or [ŋ], the velum is lowered and an oral closure is formed by the lips, tongue tip or the tongue back, such that the air stream escapes through the velopharyngeal port and solely passes the nasal cavity (fig. 1.1, right). In contrast to nasal sounds, oral sounds such as stops or fricatives usually require the velopharyngeal port to be closed by means of the raised mobile part of the velum. Raising the soft palate, on the one hand, allows an increase of the intra-oral air pressure that is necessary for the production of oral consonants (fig. 1.1, left). On the other hand, a closed velopharyngeal port prevents vowels and liquids from shifting their acoustic characteristics too much into the direction of nasality.

Figure 1.1: Position of the soft palate during the oral stop [t] (left) and the nasal stop [n] (right). The stops were uttered in isolation. Copyright: Jens Frahm / Max Planck Institute for Biophysical Chemistry.

In the case of coarticulatorily nasalized or contrastive nasal vowels, however, the velum is lowered as for the nasal stops, but no contact is induced by the tongue with the palate, such that the air stream may pass through both the oral and the nasal cavity. This is commonly referred to as vowel nasalization. During fluent speech, vowels adjacent to nasal consonants are always nasalized to some extent, which is often scarcely perceptible for the listener. Thus, the vowel in a word like *Bahn* [baːn] ‘train’ is affected by the articulatory configuration of the following nasal consonant: the vowel is partially produced with a lowered soft palate in anticipation of the following alveolar nasal stop, which requires an open velopharyngeal port for production. This case is demonstrated in the right image of fig. 1.2, which depicts the tongue and velum configuration during [aː] in the word *Aufnahme* ‘recording’. For comparison, the left side illustrates the corresponding configuration for [aː] surrounded by non-nasal sound context in *Wiedergabe* ‘replay’.

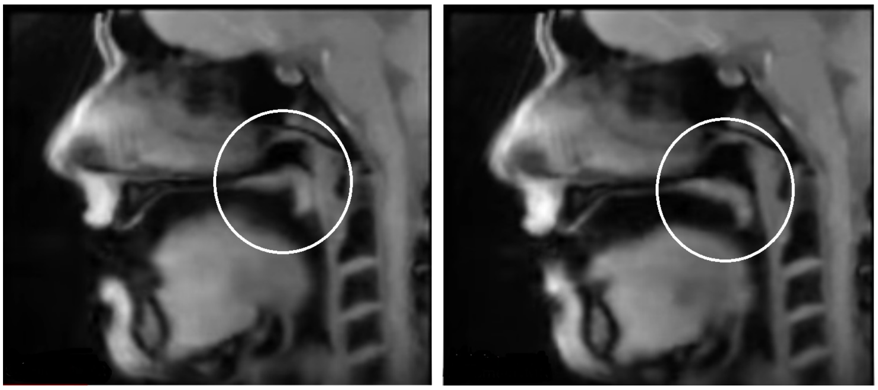


Figure 1.2: Position of the soft palate during [aː] in oral context (*Wiedergabe*, left) and nasal context (*Aufnahme*, right). When surrounded by nasal context, the soft palate is slightly lowered during the vowel. Copyright: Jens Frahm / Max Planck Institute for Biophysical Chemistry.

Languages have been found to vary with respect to the degree of non-contrastive anticipatory nasalization (Clumeck, 1976; Solé, 1992). For example, Northern Italian dialects show a clear tendency for a strong nasalization of the vowels accompanied by nasal loss, especially if the nasal is followed by a voiceless consonant (Hajek, 1997; Hajek and Maeda, 2000; Sampson, 1999). In Polish, nasal vowels preferentially occur before fricatives (Brooks, 1968); similarly, nasal vowels occur in Chilean Spanish vowel-nasal sequences in which the post-vocalic nasal is weakened especially if it is followed by a fricative. Thus, the word *naranja* 'orange' is commonly pronounced as [narãxa] (Delattre, 1946). In some Swabian dialects, the noun *Mann* 'man' is realized as [mãː] or [mɔ̃ː]. Moreover, American English is well-known for nasalizing vowels in pre-nasal position, especially if the nasal consonant is followed by a voiceless stop (Beddor, 2007, 2009; Malécot, 1960, *among others*). In some cases, the nasal is even completely lost, such that listeners solely rely on the nasality cue provided by the vowel to differentiate between words such as *camp* and *cap* (Malécot, 1960).

As indicated in the beginning, for some languages, vowel nasalization has become a contrastive property of the vowel[2]. In most of these languages, the nasal vowels have developed from vowel-nasal (VN) sequences of older stages (Chen, 1972; Ferguson, 1963; Hajek, 1997; Ruhlen, 1973; Schourup, 1973). It is widely assumed that during this process, coarticulatorily nasalized vowels progressively became more and more nasalized over time, while the nasal stop was collaterally weakened and finally lost in pronunciation:

[2]In many of these languages, the difference between oral and nasal vowels is also expressed by changes of the oro-pharyngeal shape, e.g. by alterations of the tongue position (e.g. Carignan, 2014; Carignan et al., 2015; Shosted et al., 2012, 2015).

Lat. tempus	> Fr. temps [tã]	‘time’
Lat. dens	> Fr. dent [dã]	‘tooth’
Lat. bonus	> Fr. bon [bɔ̃]	‘good’
Lat. veni	> Port. vim [vĩj̃]	‘I came’
Lat. manus	> Port. maõ [mɐ̃w̃]	‘hand’
Skr. danta	> Hindi [dãt]	‘tooth’

While it is generally accepted that nasal vowels evolved from VN sequences, the question of which specific articulatory and perceptual factors initiate such a process is more difficult to answer. The evolution of contrastive vowel nasalization has been the object of phonetic research for many decades, considering the rich variety of measurement techniques and analyzing methods that were used to illuminate the basic principles of velar function as well as the perception of nasalized sounds. Several of these experimental studies are outlined in more detail in the course of the thesis at hand. Considering the fact that there are still open questions about the basic mechanisms underlying the gestural interactions during articulation, this work aims at contributing further knowledge about the nature of coarticulatory vowel nasalization in production and perception for a language with no ongoing sound change known in terms of vowel nasalization.

1.2 Sound change

In addition to the large amount of experiments on lingual and velar behaviour, more general models of the phonetic and phonological principles are provided that were developed and extended to explain how sound change may be generally initiated and processed. Some of the basic models are outlined in the following.

Common to the phonetic models on sound change is the basic consideration that the production of speech can be regarded as a sequence of interleaved movements induced by the articulatory organs, such that the individual speech sounds overlap with each other in time: they are coarticulated. The effects that arise due to these context-dependent variations are systematic. For example, the fundamental frequency of a vowel is influenced by the voicing of the preceding stop: vowels following a voiceless stop show a slightly higher pitch than vowels following a voiced stop (e.g Brunelle and Kirby, 2016; Coetzee et al., 2014; Hombert et al., 1979; Kirby and Ladd, 2016). This is ascribed to the laryngeal differences between the voiced and voiceless stop and the subsequent variations during the transition into the vowel. The resulting effect of the systematically varying pitch height is assumed to be the basis of tonogenesis, as in many Asian and African languages.

Under normal circumstances, contextually conditioned coarticulatory variations do not automatically lead to the establishment of a new contrastive category, because listeners are experienced with these systematic effects and factor out the predictable variation. In terms of the example above, they usually do not hear a higher pitch in the vowel, because they perceptually attribute the pitch alteration to the preceding stop. Obviously, however, sound change does take place across languages, suggesting that at some level, listeners do not arrive at the interpretation of the sound material originally intended by the speaker. In turn, speakers may contribute to such misinterpretations, because they are able to vary their speaking style dependent on a given situation, such that when the lexical content is neglectable or predictable, the utterance may be produced with a higher degree of temporal and spatial gestural overlap. Considerations like these lead to different theoretic models that provide an exploratory approach for the initiation of sound change based on phonetic principles, with some of the accounts focusing on the listener's role and others emphasizing the contribution of the speaker.

Ohala (1981; 1993; 2005; 2012) develops a model that points to the role of the listener in a sound change process. As a starting point, he argues that some kinds of sound change cannot be explained by a pure articulatory account. For example, the common change from [ki] to [tʃi] (referred to as 'velar softening') as in Lat. *placeo* [plakeo]> It. *piaccio* [pjatːʃo] 'I like' involves a change of the place of articulation of the intervocalic stop: the palatal contact in [k] and [t] is induced by totally different parts of the tongue, which cannot easily be explained by articulatory variation on the part of the speaker. Although some degree of tongue back fronting before high vowels is to be expected, merging the dorsal gesture in [k] with the more front but still dorsal gesture in [i] would still not result in the apical gesture in [tʃ] (Ohala, 1993, p. 157). Moreover, Ohala, with reference to Blumstein (1986), notes that the direction of this change is asymmetrical, such that "dorsal stops become apicals near front vowels, but apicals never seem to become dorsals near back vowels" (Ohala, 1993, p. 157). Instead of assigning this change to the variability occurring during articulation, Ohala points to the spectral shapes between [t] and [k], which are found to be highly similar except for one specific acoustic feature (namely a spectral peak near 3 kHz during the burst of [k]). If this feature is missed by the listener, the velar stop, originally exhibiting this feature, sounds more similar to the apical stop. The asymmetry is explained by the assumption that the listener is more likely to miss this feature if present in the original sound rather than to insert a feature that is not present in the original sound (Ohala, 2012, p. 30). This and similar examples lead Ohala to postulate a sound change model in which the listener and not the speaker is assumed to play the crucial role: phonologization occurs due to the listener's failure to correct for the coarticulatory effect perceived from the speaker's produced utterance.

The basic idea of this model is illustrated by the often-cited example of a conceivable sound change from /ut/ to /yt/. This example is based on findings reported by Lindblom (1963), who observed spectral changes within vowels produced in CVC sequences with different speaking rates. The largest alteration was found for /u/ when surrounded by /d/, achieving a spectral pattern similar to that of /y/ in fast speech. This effect is used by Ohala to demonstrate his idea on sound change. Under normal circumstances, the speaker utters some sequence of /ut/, which, due to the rapid change of the tongue position from the back to the front, is acoustically distorted, such that it actually occurs more as [yt]. This outcome is also heard by the listener, but due to sufficient experience with this coarticulatory effect (i.e. a spectral shape more typical for [y] when /u/ is followed by an alveolar stop), it is factored out and assigned to the source /t/. Accordingly, the listener is successful in reconstructing the form /ut/, which was intended by the speaker.

However, if an inexperienced listener or a listener who did not detect the source of the coarticulatory effect hears the distorted form, the effect may not be factored out but instead is interpreted as inherent to the vowel. Subsequently, the listener-turned speaker relies on this new form and repeats it in their own production. This failure of correction on the part of the listener is referred to as *hypocorrection* (Ohala, 2012, p. 28). The development of contrastive vowel nasalization may be considered as an example of hypocorrection: when the speaker produces a nasalized vowel followed by a nasal consonant, the listener usually assigns the source of the coarticulatory nasalization to the nasal stop. However, if the listener fails to correct for this effect, perhaps due to an undiscovered conditioning environment, nasalization is associated with the vowel and reproduced as such by the listener-turned speaker (Ohala, 1993, p. 162f).

The converse case, i.e. *hypercorrection*, is also a possible scenario. Here, the speaker intends /yt/ and produces the form [yt], which is heard by the listener, but due to experience with back round vowels in apical contexts, the fronting character of [y] is ascribed to the apical stop instead of being interpreted as inherent to the vowel. This scenario may be involved in the process of dissimilation, as in the case of the English word *sword*, which in its present-day form is pronounced without the glide [w] (Ohala, 2012, p. 29).

That the listener is assumed to play the crucial role in the sound change process does not exclude those instances in which a change is obviously based on articulatory variation. For instance, Ohala considers the case of stop insertion (similar to the [klɛmpt$^{\text{n}}$nɐ] example in the beginning) in words like *youngster* [jʌŋkstɚ] or *warmth* [wɔɹmpθ], which in other languages has become phonologized (e.g. Lat. *cam(e)ra* > Fr. *chambre* ‘room’, cf. Ohala 1993, p. 160). Here, the stop emerges as a result of the interleaved gestures during oral closure and the modification of the velopharyngeal port opening: while the lips are still closed for /m/, the lowered velum raises for the upcoming fricative. Accordingly, during the transition from the nasal to the fricative, a sound emerges that combines velum closure with the place of articulation of the prior nasal, resulting in a short oral stop. Similarly, the case of tonogenesis mentioned above is taken as another example in which systematic variation is provided clearly on the part of the speaker. Importantly, however, Ohala emphasizes that

> "[c]ases like these support the notion that variation is found in speech due to what the speaker may do but it takes the listener to misinterpret or misparse the elements of pronunciation in order to produce a sound change." (Ohala, 2012, p. 31).

In many respects, Lindblom (1990) and Lindblom et al. (1995) follow Ohala's listener-based approach to sound change. However, they differ in two essential aspects: on the one hand, the role of misperception is de-emphasized, on the other hand, the role of the speaker is considered more significant. In their model, sound change can be seen as a three-step process: large phonetic variation emerges due to the listener-oriented adjustments made by the speaker, who is confronted with the conflict of sufficient intelligibility on the one hand and reduction of articulatory effort on the other hand, dependent on the conversational situation. While listeners are usually unaware of the variations produced by the speaker, it occasionally happens that they become aware of the phonetic variants in the signal, for example when the actual lexical content is less informative, predictable or redundant. The listener-turned speaker may accept this new variant and reproduce it. In the last step, the new form is either accepted or rejected by the speech community. This decision depends on a combination of social, perceptual and articulatory factors. Considering the first two steps, Lindblom and colleagues thus clearly argue that misperception alone, as suggested by Ohala, is not sufficient to trigger sound change: if a word is misperceived by the listener and subsequently reproduced by the listener-turned speaker, this presupposes that the listener had access to the lexical form of the word and thus to its actual pronunciation (Lindblom et al., 1995, p. 19). Accordingly, the listener is not forced to produce the new form but may choose between the two variants. Whether the old variant is replaced by the new one in the speech community depends on

> "how well it fits the social, communicative, articulatory and lexico-systemic criteria that speech community members tacitly apply to it when they encounter it as listeners, and when, and if, they try it out as speakers." (Lindblom et al., 1995, p. 19).

The three steps just outlined are to be delineated in more detail. The sound change model suggested by Lindblom et al. (1995) is based on the general principles of the *hypo-* and *hyper-speech model* (H&H) introduced by Lindblom (1990) (not to be confused with Ohala's model of hypo- and hypercorrection). According to this theory, the speaker is assumed to adapt to the listener's individual informational needs, such that the speaker provides the signal with "sufficient discriminatory power for making the correct lexical identifications, not necessarily that they be invariant" (Lindblom et al., 1995, p. 8). Hence, the signal perceived by the listener needs to contain "just enough [information] for that percept to emerge in interaction with the stored knowledge" (Lindblom et al., 1995, p. 8). The speaker is thus assumed to adapt to the social and communicative needs considered for each specific speech situation, along a continuum from more energetic articulations (*hyper-speech*) to shapes that exhibit less energetic forms (*hypo-speech*), i.e. a decrease of gestural amplitude and an increase of temporal overlap. This interplay between the speaker's demand for intelligibility and articulatory simplification determines the phonetic form finally produced by the speaker, such that the alterations between hypo- and hyper-speech (and all its intermediate steps) lead to the large phonetic variation observed.

Under normal circumstances, the listener is unaware of this variability, as the signal is transformed to the form stored in the listener's knowledge. Occasionally, however, the listener becomes aware of the actual acoustic form, for example when lexical information is superfluous, as may be the case in high frequency words or words with low meaning content (e.g. greetings, exclamations, overlearned phrases; Lindblom et al., 1995, p. 14). In these cases, the listener actively stores "unprocessed phonetic patterns captured in sporadic moments of acoustic/auditory 'truth' " (Lindblom et al., 1995, p. 17). To become accepted as a new form, this phonetic variant is evaluated by the speech community with respect to the phonetic shape, among others. The evaluation is made without awareness and, as mentioned above, involves social, communicative and articulatory criteria. Thus, this sound change scenario reflects the general principles presumed in H&H theory: the speech community evaluates the perceptual contrast versus articulatory simplicity. Whether a variant that is 'easier' to articulate will be accepted depends on its potential of perceptual confusion with words that exhibit a similar structure. Conversely, a variant exhibiting a more contrastive shape than the old form, facilitating perception, may nonetheless be rejected if it is too complex in articulation (Lindblom et al., 1995, p. 20). In this sense, sound change is seen as a process of phonetic variation and phonological selection.

The sound change model developed by Beddor (2009, 2012) takes some aspects from the above models into account but fundamentally differs in its perspective on the listener's role. In this model, sound change is not initiated by a perception error on the part of the listener but instead because the listener is highly sensitive to the fine phonetic detail that is provided by the acoustic signal. These phonetic cues "assist listeners in determining what speakers are saying and in making linguistic decisions" (Beddor, 2009, p. 787). Essential to the model is the hypothesis that listeners are attentive perceivers of the acoustic consequences of coarticulation, but still may arrive at a different representation than intended by the speaker. This is due to the large variation of the temporal and spatial extent of coarticulation, which arises from the varying interactions between the coarticulatory effect and its source and which, on the part of the listener, leads to different perceptual weightings of the acoustic properties. Thus, the grammars listeners arrive at are listener-specific, because the perceived acoustic signal is compatible with different phonological analyses.

Beddor exemplifies her model by a series of experiments on vowel nasalization in production and perception, considering the question of how coarticulatory nasalization in a pre-nasal vowel becomes more associated with the vowel rather than the actual source that gives rise to it. As one basic assumption, at some stage during the process of phonologization, the duration of the coarticulatory effect is inversely related to the duration of its source: the longer the vowel is nasalized, the shorter the nasal stop. This implies that the duration of the velum lowering gesture itself is roughly stable, whereas its alignment with the oral articulators is variable.

Beddor finds evidence from production experiments that were run with native speakers of American English: in words with VNC sequences (with C being either a voiced or voiceless stop), longer portions of vowel nasalization correlated with shorter nasal consonants. Although this correlation held both across and within the voicing contexts, the extent of the observed covariation was found to be linked to the voicing of the post-nasal segment (also noticed in prior research by e.g. Busà, 2003, 2007; Cohn, 1990; Hattori et al., 1958; Raphael et al., 1975): when followed by a voiceless segment, the nasal was especially short and the vowel was nasalized to a greater extent. This observation is generally explained by aerodynamic and auditory factors, suggesting that while voicing is compatible with some velar leakage, voiceless obstruents require a closed velopharyngeal port for a sufficient buildup of air pressure. This may involve an earlier velum closing gesture, resulting in a shorter nasal stop (Ohala and Ohala, 1991; Beddor, 2009). The finding that the duration of vowel nasalization was related to the extent of the nasal stop is accounted for by the concept of a constant-sized velum lowering gesture that is variably aligned. According to this idea, the duration of the overall velum lowering gesture is roughly stable across the voicing contexts, but is shifted more into the vowel in the voiceless condition, with the effect that the nasal stop is shortened and the post-nasal segment is temporally increased (Beddor, 2009, p. 789).

The findings from Beddor's production experiment illustrate the concept of covariation between the coarticulatory source and its effect as well as the idea that articulatory (here: velar) gestures are roughly stable in duration but differently aligned depending on the context. Consequently, these variations are assumed to affect the perception of nasals and vowel nasalization, such that listeners may interpret this covariation in different ways. In this sense, perception is seen as a process "that takes advantage of the variation afforded by the acoustics" (Beddor, 2009, p. 798). With respect to vowel nasalization in English, the listener is confronted with a range of ṼN variants (where Ṽ means a nasalized vowel) across and within different voicing contexts and accordingly needs to formulate a phonological representation that satisfies this range of variants. One possibility, and central to the model, is that at least some listeners use the covarying information in the signal in determining the representation. These listeners are thus assumed to use the information of 'nasal' rather than /ɛ̃/ or /n/ to differentiate between a word pair like *bent* and *bet*, i.e. they are insensitive to the accurate distribution of nasalization but rely on the general information of whether nasality is present. Consequently, "such listeners would be accurate perceivers of the input signal, yet would have arrived at a phonological representation that differs from that of the speaker who intended /bent/." (Beddor, 2009, p. 798).

Evidence for the hypothesis of such perceptual equivalence comes from two further experiments in which listeners were exposed to different ṼN patterns. In the first one, listeners heard variants of the nonsense word [gã(m)ba], which was modified in terms of the duration of the nasal and nasality on the vowel. The words were presented in pairs that differed in three ways: in the first type, the pair members only differed in the duration of the nasal stop. In the second type, the words clearly differed in the overall nasality that extended to both the vowel and the nasal. In the third type, overall nasality was similar in both words but was distributed differently on the nasal and the vowel. Listeners were asked to decide whether the words of a pair sounded the same or different. As predicted, most difficulties occurred in discriminating those pairs with similar nasality, suggesting that listeners treated nasality on the vowel and nasal consonant as perceptually equivalent.

The second experiment tested whether the patterns of perceptual equivalence are context-dependent under the assumption that listeners are exposed to systematic ṼN variation in voiced and voiceless contexts. This question was addressed by conducting a discrimination and identification task. The discrimination task was structured in a comparable way to the first experiment. As stimuli, the continua of [bet]-[bẽnt] and [bed]-[bẽnd] were tested separately. While the results for the [bẽ(n)d] continua mainly reflected those of the [gã(m)ba] task, findings from the [bẽ(n)t] stimuli suggested that those pairs differing in the duration of the nasal segment were most difficult to discriminate, which was not in agreement with the prior findings. As an explanation to this discrepancy, Beddor points to the listeners' context-dependent sensitivity to vowel nasalization: if the vowel is heavily nasalized and the nasal is shortened in production, especially in VNC sequences when C is voiceless, listeners have experience with this pattern, such that vowel nasalization in these contexts is expected to serve as a reliable cue for at least some perceivers. Moreover, closer inspection revealed that individual listeners systematically differed in their performance: while some participants consistently exhibited a pattern of perceived equivalence (judging the similar-nasality pair members as the same), others showed an equal performance when rating pairs with different and similar nasality, suggesting that these listeners were highly sensitive to the nasalization occurring on the vowel.

In the identification task, listeners were exposed to a four-choice test with stimuli of /bed, bend, bet, bent/ that exhibited varying degrees of vowel nasalization and nasal duration. Listeners were supposed to indicate which of the four words they heard. As predicted, they required a longer nasal stop in the voiced than in the voiceless context to identify a word as CVNC. Moreover, in both voicing contexts, listeners needed a shorter nasal stop to elicit CVNC responses the more the vowel was nasalized. In the voiceless context, however, this effect was especially conspicuous: across listeners, a heavily nasalized vowel was sufficient to give a CVNC response, with no nasal stop present at all. The results from the identification test reflected those from the discrimination task: a trade-off relation was evident for the nasalized vowel and the nasal stop, suggesting perceptual equivalence. The finding that a heavily nasalized vowel in voiceless context was sufficient to elicit the perception of a nasal stop indicates that in this context, listeners strongly relied on vowel nasalization to make their decision. However, the response patterns of individual listeners revealed large listener-specific differences: while for some, a stimulus with a short nasal counted as /bet/, others identified the same stimulus as /bent/. Under the assumption that such results achieved under laboratory conditions point to specific acoustic properties listeners use in natural speech, the perceptual differences across participants are taken "as evidence of their different phonological grammars" (Beddor, 2009, p. 812f).

In summary, Beddor's account comprises a production and a perception aspect: on the one hand, phonologization involves a stage in which the duration of the coarticulatory source and its effect are inversely related. On the other hand, listeners perceive the covarying information and may arrive at a phonological interpretation of the signal that is different from the interpretation intended by the speaker. The variably realized cues in the input signal allow for different perceptual choices, because different perceptual weightings of the coarticulated effect and the source are compatible with the input. Hence, listeners can be categorized as more conservative or more innovative perceivers (Beddor, 2012, p. 53): while the conservative listeners rely on the coarticulatory source in making perceptual decisions, the innovative listeners primarily weight the coarticulatory effects. These innovations are then manifested through the listener's own productions or through expectations about the coarticulatory patterns occurring during speech. Accordingly, sound change is seen as a process that is driven by those listeners who arrive at innovative representations due to their "close but selective attention to the dynamic coarticulated signal" (Beddor, 2012, p. 53).

1.3 The velopharyngeal mechanism

The fact that across (and even within) languages vowel nasalization shows a highly varying pattern with respect to its temporal and spatial extent suggests that speakers have some control over the fine-tuned gestures of the soft palate during speech. For a better understanding of the velopharyngeal mechanism responsible for the opening and closing gesture the physiological structures of the soft palate and its connected tissue are outlined in the following as well as its function in the gestural coordination of speech sounds.

During normal breathing, the nasal cavity is coupled with the oropharyngeal cavity via the velopharyngeal port (henceforth VP), which allows the airstream to escape from the lungs through the nose. The VP comprises the passage between the pharyngeal wall and the soft palate. An open VP requires a lowered velum, while a closed VP is achieved by elevating the soft palate. One primary task of the velopharyngeal musculature is to smoothly ensure the different lowering and raising phases during the swallowing act. As the bolus is transported towards the esophagus, the nasal cavity is separated from the oropharyngeal space by elevating the velum to prevent saliva or food entering the upper airway. The accurate mechanisms of the velopharyngeal gestures during speech as well as their participating muscles have been the object of various interdisciplinary fields of research, such as speech therapy, investigation of dysphagia and phonetics. The soft palate consists of structures of connective tissue and mucous membrane and is based on the palatine aponeurosis, which forms the mobile process of the hard palate and into which the palatal muscles insert. The velopharyngeal musculature primarily consists of muscle bundles that are responsible for opening and closing the velopharyngeal port. These muscles include the levator palatini, tensor veli palatini, musculus uvulae, superior pharyngeal constrictor, palatopharyngeus, palatoglossus and salpingopharyngeus. Figure 1.3 shows a sketch of the muscles that are most relevant for the velum adjustments from a dorsal perspective.

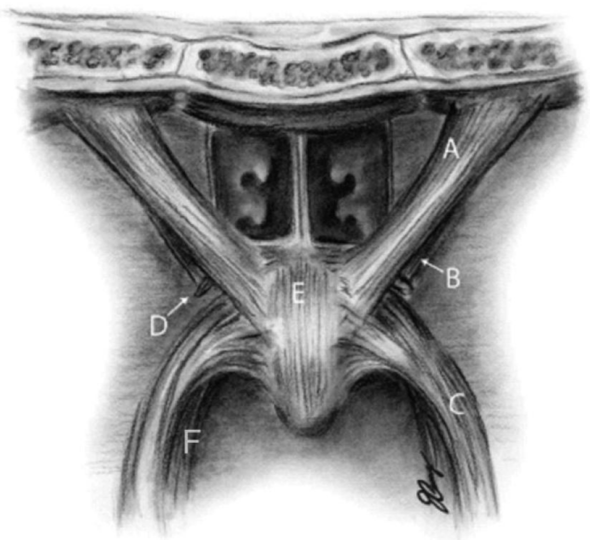


Figure 1.3: Dorsal view of the nasal surface of the velum and the velopharyngeal muscles. A: levator veli palatini; B: tensor veli palatini; C: palatopharyngeus; D: hamulus; E: musculus uvulae; F: palatoglossus. From Perry, 2011a, p. 89. Copyright: Thieme Medical Publishers, Inc.

It is widely agreed that the levator palatini muscle is primarily responsible for the elevating gesture of the soft palate (e.g. Bell-Berti, 1973; Lubker, 1968). This muscle originates at the petrous bone which is part of the temporal bone and inserts from both sides of the skull into the palatine aponeurosis.

The course of the tensor veli palatini muscle is quite similar to that of the levator palatini. It originates at the spina angularis of the sphenoid bone and ends in a tendon that winds around the hamulus pterygoideus, a hook-like process of the medial pterygoid plate of the sphenoid bone. The tendon finally inserts into the palatine aponeurosis. The primary function of the tensor veli palatini muscle is to open the tuba auditiva during contraction to equalize the pressure in the tympanic cavity.

The musculus uvulae is located within the velum and therefore does not have any external attachments. It originates at the posterior nasal spine and palatine aponeurosis and inserts in the tip of the uvula. Its main function is to strengthen the contact with the posterior pharyngeal wall during the closure of the velopharyngeal port.

The process of velopharyngeal closure does not only involve muscles retracting and elevating the velum but additionally includes pharyngeal muscles that assist in establishing sufficient sealing of the nasopharynx. The superior pharyngeal constrictor muscle is part of the upper pharyngeal wall and consists of four parts with different osseous and muscular origins. During contraction, the superior pharyngeal constrictor narrows the nasopharyngeal area by forming a muscular bulge, the *Passavant's ridge*. In addition to the process of velum retraction, the Passavant's ridge contributes to sealing the nasal cavity from the oral cavity during swallowing and speech.

Similarly, the palatopharyngeal muscle originates in the palatine aponeurosis and is divided into two bundles by the levator palatini and the uvular muscle. The fiber bundles of the palatopharyngeal muscle run behind the palatine tonsil towards the lateral pharyngeal walls and the back rim of the thyroid cartilage, where they insert. Depending on the location of the bundles, they have different functions: the upper transverse fibers support the constriction gesture of the lateral pharyngeal walls and also contribute to the formation of Passavant's ridge, while the vertical fibers determine the velum position in general.

For the production of nasal sounds, moving the soft palate in the downward direction is crucial to facilitate VP opening. In theory, this could be achieved either by actively pulling down the soft palate by means of increased muscle activity of at least some of the velopharyngeal muscles or by relaxing the levator palatini. Research on this issue provides evidence for both scenarios, often reporting highly speaker-specific patterns. Some studies suggest that the velum lowers in a passive way due to relaxation of the levator palatini (Bell-Berti, 1973, 1976), while others report an increased activity of the palatoglossus muscle during the velum lowering process (Lubker et al., 1970). In fact, the palatoglossus muscle constitutes a link between the palatal and the lingual musculature: its muscle fibers originate in the palatine aponeurosis and run laterally towards the fiber bundles of the lingual musculature, where they insert in the lateral margins of the tongue. The palatoglossus muscle supports the process of swallowing by raising the root of the tongue and narrowing the isthmus faucium. The fact that speakers vary in using the palatoglossus for actively lowering the soft palate might be related to individual anatomical conditions, as the specific location of the origins of the muscle attachments slightly differs for each speaker (Perry, 2011a, p. 89). Furthermore, the palatoglossus is also suggested to play a crucial role for the lower velum position consistently reported for the low vowel /a/ compared to other vowels. As this muscle contracts for narrowing the faucial isthums during low vowels, it might cause a pull-down effect on the soft palate, suggesting that the tongue and velum position may directly influence each other (Moll and Shriner, 1967). Other studies, in contrast, found increased palatoglossus muscle activity primarily during lingual instead of velar gestures and, in turn, reported some weaker contraction of the levator palatini for low vowels compared to high vowels, which suggests that the velum is lowered in a more passive rather than active way (Bell-Berti, 1973, 1976; Lubker, 1968).

1.4 Linguistic factors affecting velum behaviour

The spatial and temporal extent of the velum lowering gesture during speech is influenced by various linguistic parameters that involve intrinsic articulatory properties of the individual sounds and their surrounding contexts as well as prosodic factors such as speaking rate and stress. Importantly, the temporal amount of the lowering gesture does not necessarily correlate with the degree of spatial displacement, such that extensive temporal vowel nasalization may occur even with a relatively small degree of VP opening (Clumeck, 1976).

Vowel height Across languages, it has been widely attested that low vowels in nasal contexts are produced with a lower velum position than high vowels. This pattern is even observed in non-nasal contexts (Amelot and Rossato, 2006, 2007; Bell-Berti, 1973, 1976; Clumeck, 1976; Lubker, 1968; Moll and Shriner, 1967; Rossato et al., 2003) and is commonly accounted for by both phonetic-mechanical and perceptual principles. On the articulatory side, a lowered velum during a low tongue position may be constituted by the specific muscular connection between the soft palate and the tongue, the palatoglossus muscle, which is assumed to induce some pull-down effect on the velum when the tongue is in a low position (Moll and Shriner, 1967). In addition, the levator palatini, which is primarily responsible for raising the soft palate, is found to exhibit some more increased contraction in high vowels compared to low vowels (Bell-Berti, 1973, 1976; Lubker, 1968). On the perceptual side, high vowels are reported to tolerate less velar port opening before they are perceived as nasalized: listeners perceive high vowels as nasalized when these are provided with only little velum lowering, whereas low vowels require much more VP opening to be rated as nasalized (House and Stevens, 1956; Maeda, 1989, 1993; Ohala, 1975). Conflicting findings, however, were also reported (Ali et al., 1971; Lintz and Sherman, 1961). Based on the results from these perception studies, it is suggested that the differences in velum lowering between low and high vowels are related to the speaker's intention of preventing high vowels from being unintentionally perceived as nasalized by constricting the VP to a higher degree. Low vowels, in turn, allow for larger changes of the frequency spectrum before they are judged as nasalized. Thus, velum height and perceived nasalization are not correlated equally for each vowel; instead, each vowel has its own scope of velum lowering before it is perceived as nasalized (Maeda, 1989, 1993; Ohala, 1975).

Vowel length Another intrinsic property affecting at least the temporal extent of vowel nasalization is vowel length. From an articulatory perspective, only a handful of studies are concerned with physiological data on the timing of velum lowering within pre-nasal vowels of different durations. However, it is remarkable that across languages the low vowel /a/ is often longer than mid or high vowels in a comparable context (Clumeck, 1976; Hajek, 1997; Toivonen et al., 2015), suggesting that low vowels are frequently affected by vowel nasalization not because they are low but because they are long (Hajek, 1997; Hajek and

Maeda, 2000). Further evidence for the role of vowel length is provided by the finding that no case is reported in which short vowels have become contrastively nasalized without the long vowels being affected first: if a language has short contrastive nasal vowels, it always also has long nasal vowels. The reverse, however, is not the case, suggesting that long vowels are preferentially nasalized, which provides the basis for the emergence of contrastive vowel nasalization (Hajek and Maeda, 2000). More evidence for vowel length as a relevant factor comes from various perception studies, indicating that increased vowel length is accompanied by increased perceived vowel nasalization, even if the degree of VP opening is controlled and left unchanged for different vowel lengths (e.g. Delattre and Monnot, 1968; Hajek and Watson, 1998; Whalen and Beddor, 1989). Thus, at least on the perceptual side, vowel length is likely to play a fundamental role in the establishment of contrastive nasal vowels.

Post-vocalic context The temporal and spatial extent of vowel nasalization is further affected by the post-vocalic context. As illustrated by fig. 1.2 in the beginning, vowels preceding a nasal stop generally show a greater degree of velum lowering than vowels surrounded by oral context. However, as just indicated, velum lowering is also evident during oral vowels, especially during the low vowel /a/ (Amelot and Rossato, 2006; Bell-Berti, 1973, 1976; Bell-Berti et al., 1979; Clumeck, 1976). In fact, velum lowering and also velum raising occur in fine-grained stages; these gestures are more complex than simply exhibiting a two-level closed vs. open mode (Bell-Berti and Krakow, 1991; Solé, 1992). Even if the velum is sufficiently raised to allow the production of oral consonants, it may still continue elevating up to anatomical restriction, depending on the consonantal context (Amelot and Rossato, 2006; Kuehn, 1976). However, the degree of velum lowering during the vowel is not only affected by the immediate post-vocalic segment, but also by the context that follows a VN sequence. For example, nasals followed by fricatives are especially prone to becoming weakened and finally deleted, leaving a sound sequence that consists of a vowel or nasal vowel immediately followed by the fricative consonant (Busà and Ohala, 1995; Busà, 2003). This phenomenon is accounted for by the acoustic similarities between the transition cues of vowels followed by a nasal and vowels followed by a fricative, such that listeners might ascribe the nasal portion in a vowel-nasal-fricative sequence to the transition cues otherwise occurring in vowel-fricative sequences (Busà and Ohala, 1995). Further contextual effects include the voicing of the post-nasal obstruent. For American English, as outlined earlier in this chapter, the nasal is found shorter before voiceless stops (and sometimes is even not present) than before voiced stops (Beddor, 2007, 2009; Malécot, 1960) and in addition, vowel nasalization is found to be inversely related to the length of the nasal: the more nasalization on the vowel, the shorter the nasal, which is especially distinct when the nasal is followed by voiceless consonants.

Stress With respect to the non-segmental factors showing some impact on the nasality of vowels, diachronic cross-linguistic reports indicate a tendency for stressed vowels to be more affected by nasalization than unstressed vowels (Schourup, 1973). However, phonetic studies on the velar and lingual behaviour during pre-nasal stressed and unstressed vowels have produced inconsistent results. In general, speakers show large variation with respect to velum lowering during stressed versus unstressed vowels: for some speakers, both low and high vowels show a lower velum position in a stressed context, while for others, the intrinsic velum position during the vowel is enhanced, such that the velum is lower in stressed /a/ but higher in stressed /i/ compared to the unstressed counterparts (Krakow, 1993). Similar findings are reported for the tongue gesture (De Jong, 1995; Farnetani and Vayra, 1996; Kent and Netsell, 1971; Straka, 1963). Furthermore, vowel duration is highly affected by stress: stressed vowels are generally longer than unstressed vowels. Similarly, vowel nasalization tends also be to extended, which, however, is accompanied by some considerable variation across speakers (Busà, 2003). In contrast, studies on perceived nasalization of synthetic stressed versus unstressed vowels with different lengths and degrees of VP opening point to a secondary role of stress: while unstressed vowels with no VP opening are perceived as more nasalized than the stressed analog, the opposite is found with a maximally open VP (Hajek and Watson, 1998). In summary, while the diachronic data suggest a major role of stress, the physiological findings from phonetic experiments are not consistent.

Speaking rate Another prosodic factor that affects velum lowering patterns is the speaking rate. During fast speech, the gestures required for articulating the sound sequences can be modified in various ways. On the one hand, they may be produced with an increased velocity that enlarges temporal overlap, on the other hand, the gestures may be spatially reduced in their degree of displacement. These two options for achieving the gestural targets are found to be variably used across speakers and may also co-occur. With respect to the soft palate in particular, more temporal overlap, i.e. fewer single-stage lowering gestures during sound sequences are found with rapid speech (Bell-Berti and Krakow, 1991). Moreover, fast speech can affect both the spatial amount and the velocity of the lowering gesture, which, however, is implemented in various ways: some speakers reduce the range of the velum movement extremes, but exhibit a roughly stable velocity, while others show a similar extent of the movement amplitudes independent of the speaking rate but increase the velocity of the velar movements in fast speech (Kent et al., 1974). Moreover, for those speakers reducing the amount of the movements, this strategy can affect both the raising and the lowering gesture (Moll and Shriner, 1967) or only one of these (Kuehn, 1976). Data from two different speakers showed that when the velum was raised higher than necessary for closing the VP in moderate speech, it was raised to a lesser extent in faster speech but not affected in the low position. In contrast, a speaker with a less extreme high position in moderate speech showed a stable position when the velum was raised but a reduced lowering gesture in fast speech when the low position was required (Kuehn, 1976).

1.5 Measurement techniques

Investigating the soft palate during natural speech is far from trivial, as the velum is located in the back of the oro-pharyngeal area. In phonetic research, highly invasive methods have often been used to obtain data on the velar movements during fluent speech or isolated sounds. With respect to the data quality, it is questionable whether the velum can move in a natural, unaffected way when such devices are applied. In principle, non-invasive imaging methods provide a good alternative for tracking the velum, although not all imaging techniques are suitable for long measurements with a higher number of participants (a few X-ray studies on velum behaviour do exist, though). With the technical progresses of the imaging techniques primarily utilized for medical diagnostics and research, MRI and real-time MRI recordings have also been sporadically used for several phonetic studies, presenting analyses of sustained sounds or slow speech. Nowadays, modern real-time MRI recordings can be obtained with a high temporal resolution and a concurrent good image quality. This section provides a short overview of the conventional measurement techniques for tracking speech gestures and velum movement in particular. The review is followed by a more detailed characterization of the MRI operating mode and the specific real-time MRI technology applied for the experiments in this thesis.

X-Ray Early experiments using X-ray imaging provided much information about the vocal tract and the configuration of gestures during speech. With X-ray, images are generated by sending high energy electromagnetic beams from a source on one side through the body to a receiver on the other side. Since the affected tissues absorb the beams to different extents, the output indicates differences in tissue density. While in early single-frame X-ray experiments, static articulator positions were imaged, the modified technique of cineflourography allowed for observing dynamic speech, in which the images were projected on a fluoroscope and filmed with high speed (Kent and Moll, 1969; McClean, 1973; Moll, 1962; Moll and Daniloff, 1971). However, the attractiveness of imaging the vocal tract during dynamic speech with X-ray is accompanied by two major drawbacks. First, the electron beams pass all tissues, which are reflected on the images. Thus, structures with high density like bones and teeth partly block the view of softer and less dense tissues such as the tongue or soft palate. Second, long-term exposure to X-ray beams constitutes a health risk for the participants due to an increase of cumulative radiation dose. It should be noted, though, that the cineflourography technique was developed and improved over the years, such that continuous radiation was obviated and exposure was reduced to a certain extent.

Electromyography Electromyography (EMG) is a technique that allows for recording electrical signals during muscle contraction. Small electrodes are placed near the muscles under investigation to detect the weak electrical waves which occur during a transmitted pulse to the interface between the nerve and the muscle. EMG electrodes can either be applied on the tissue surface or intramuscular, depending on the muscle position and constitution. For investigating the velopharyngeal muscles, only intramuscular electrodes are appropriate due to the muscular arrangements and deeply located fibers that otherwise would induce erroneous signals if surface electrodes were utilized. Provided with fine needles, the intramuscular electrodes are inserted into the muscles under examination, which, in case of the soft palate, usually requires the application of a topical anesthetic in advance. In most EMG studies concerned with the velopharyngeal musculature, hook-wired electrodes were used, as they are particularly thin and flexible (e.g. Bell-Berti, 1973, 1976; Lubker, 1968). Early EMG studies provided a lot of important information about the activity of the specific velopharyngeal muscles during speech in different languages (Bell-Berti, 1975; Bell-Berti and Hirose, 1973; Seaver and Kuehn, 1980). Most of them involved only a small number of participants due to factors such as time, effort and discomfort.

Fiberoptics Another method for investigating the laryngeal structures and the velopharyngeal mechanism involves the application of flexible fiberoptic endoscopes. For examination, a thin bundle of glass fibers is inserted into the nose up to the required position, usually above the velopharyngeal port. The fibers contain a light source to illuminate the naso-pharyngeal cavity as well as a lens or camera to transmit the image. Fiberoptic imaging thus allows for the direct observation of the velopharyngeal behaviour during speech (Bell-Berti et al., 1979; Bell-Berti and Hirose, 1975; Henderson, 1984; Matsuya et al., 1974). However, its invasive nature leads to at least some restrictions of normal breathing and the production of nasal sounds on the part of the participant.

Photodetection Similar to the fiberoptic devices, photodetection involves a light source and a light detector, capturing the relative amount of light passing the port under examination. Based on the voltage transduced, time-varying information about the altering degrees of opening is received. As such, photodetection is suitable for investigating the opening changes of the velopharyngeal port during speech. Several photodetection devices were constructed for phonetic research, such as the nasograph (Ohala, 1971) and similar instruments (Dalston, 1982; Kuenzel, 1977). The nasograph consists of a thin flexible tube that contains both a light source and a detector. To examine velar movements, the tube is inserted through the nostrils into the nasal cavity until the light source is positioned below the VP and the detector is located above the VP. More recent studies have used a similar principle but a slightly different setup with the light source above the velum and the detector in the pharyngeal cavity (Amelot et al., 2006). Photodetection in these setups provides detailed information about velar movements during speech, leading for example to the finding that oral vowels surrounded by a nasal context often stay nasalized (Amelot et al., 2006). However, as with the fiberoptic devices, the question remains to what extent natural velar behaviour is disturbed by a tube that is positioned in this sensitive area.

Airflow measurements Access to velar behaviour is also possible by indirect measurement methods, such as the measurement of the nasal airflow. The amount of nasal airflow provides information about the degree of VP opening. To obtain nasal airflow data, the participant wears a special mask covering the nose, sometimes with two tubes inserted into the nostrils and material sealing the rims. During speaking, the air passing the nasal cavity is channeled into a pressure measurement tool. The measurement signal is obtained by converting the pressure into a time-varying electric signal. In addition, a split mask is often used, which covers both the nose and the mouth and allows for measuring the airflow separately by means of an integrated divider. The isolated airflow data, however, are not always a precise indicator for velum lowering, as the amount of air emitted through the nose also depends on other factors, such as the constriction within the oral tract or variations in glottal aperture (Krakow and Huffman, 1993). For example, due to the high and front position of the tongue during /i/, the oral cavity is much more constricted than during the low vowel /a/, inducing a higher intra-oral pressure. Thus, presupposing a scenario with identical VP opening for both sounds, more nasal airflow pressure would be registered for /i/. Similarly, changes of the overall airflow amount may occur due to the sub-glottal pressure in accentuated utterances or due to an altering glottal aperture during aspirated stops (Krakow and Huffman, 1993). Therefore, to ensure a reliable interpretation of the airflow signal, this method is often combined with other instrumental techniques such as acoustic recordings or direct tracking of the articulators (Basset et al., 2001; Birch et al., 2002; Chi et al., 2015).

Spectral analysis Acoustic analyses are an attractive tool for the investigation of nasalization as an effect of velum lowering. Providing a safe and non-invasive technique, natural speech can be recorded and analyzed in terms of acoustical parameters. However, the identification of specific spectral characteristics of nasal sounds and especially of nasalized vowels is not trivial, as the coupling of the oral and nasal cavities results in complex and interacting resonances (Maeda, 1993; Ohala and Ohala, 1993). Nonetheless, a few general spectral differences between oral and nasal vowels can be summarized (further details in chapter V). With nasal coupling, additional formants and anti-formants occur in the frequency spectrum (Krakow and Huffman, 1993), inducing changes of the bandwidth and frequency of the first formant (F1), which is reported to shift upwards when the nasal tract is coupled (Fujimura and Lindqvist, 1971; House and Stevens, 1956). Furthermore, resonance and anti-resonance pairs are found in the region of F2 and F3. The overall formant amplitude is generally reduced due to the damping effect of the nasal cavity surface and narrow nostril aperture. However, although spectral analysis may be meaningful for some specific questions on nasals and nasalization, they nonetheless provide poor information about the fine-detailed lowering stages during velum movement.

Electromagnetic Articulometry Electromagnetic articulometry (EMA) constitutes a point-tracking measurement method. Instead of imaging the full vocal tract, specific tissue points of a determined articulator are tracked and the received signal is converted for further analysis. The EMA method works with alternating magnetic fields to register the movements of small receiver coils that are fixed on the articulator under examination. Three (in a two-dimensional system, Perkell et al., 1992) or six (in a three-dimensional system, Hoole et al., 2003) transmitter coils are placed around the participant's head, with each generating alternating magnetic fields at different frequencies. These magnetic fields induce an alternating signal in the receiver coils, generating an electric signal that is inversely related to the distance between the transmitter coils and receiver coils: the larger the distance, the weaker the signal. By means of these distance values, the location of the receiver coils can be calculated as a function of time. The general advantage of EMA is the potential to track multiple articulators simultaneously during speech at a rapid sampling rate. This technique is applied to explore various research questions (e.g. Byrd et al., 2005; Fuchs et al., 2005; Kühnert and Hoole, 2004). In addition, concurrent acoustic recordings can be easily made. However, with respect to the velum in particular, tracking the soft palate has proved challenging because gluing a sensor on the soft palate is not successful for every participant as the coil may not stay fixated until the measurement is completed. Despite these challenges, while most EMA experiments consider tongue movements, some studies are also concerned with velum tracking (e.g. Amelot and Rossato, 2006; Jaeger and Hoole, 2011; Katz et al., 1990).

Magnetic Resonance Imaging Magnetic resonance imaging (MRI) is an imaging method that is based on nuclear magnetic resonance (NMR) and creates sectional images of a measured object by taking advantage of the physical characteristics of protons. MRI provides a high potential for tissue discrimination without being invasive and additionally lacks the risk of ionizing radiation, which makes this method quite attractive for speech studies. In conventional acquisition, only images of static vocal tract configurations can be recorded, which in earlier applications was very helpful in providing important insights regarding the velum position during different sustained vowels (Whalen, 1990). Since its invention in the early 1970s, however, MRI technology has been continuously improved and the acquisition time has been dramatically reduced, such that modern MRI technologies allow for data acquisition in real time with a frame rate up to 100 fps. However, as MRI machines are predominantly used in the medical sector and generate high costs, they are usually not easily accessible for studies from external research fields, even less so if these studies involve a larger number of participants. Hence, no phonetic MRI study has been carried out so far that involved more than a handful of speakers (e.g. Byrd et al., 2009; Carignan et al., 2013; Demolin et al., 2003; Engwall et al., 2006; Proctor et al., 2013). Another issue concerns the supine position of the participants during MRI measurements.

This point has been addressed several times (e.g. Engwall, 2013; Kitamura et al., 2005; Perry, 2011b; Whalen, 1990) and is discussed at the end of the following section.

In summary, each of the measurement techniques presented above provides advantages and disadvantages for examining the movement patterns of the soft palate. However, the observation of the velum during natural speech, unaffected by technical instruments, is solely enabled by non-invasive measurement techniques. Since the data analyzed in this thesis were obtained by means of modern real-time MRI, the following section presents some basic key points on the operating mode of this technology as well as some background information about the particular measurement sequence applied.

1.5.1 MRI: From signal to image

Invented in the early 1970s by Paul C. Lauterbur and Peter Mansfield, the MRI method has been continuously modified and developed. This section outlines the basic function of the MRI procedure (primarily referring to Elster, 2021 and Weishaupt, 2014) as well as the Fast Low Angle Shot method (FLASH), which is commonly applied today in the broad field of scientific and medical research (e.g. Iltis et al., 2015; Niebergall et al., 2013; Uecker et al., 2010; Zhang et al., 2012).

Basic physics A conventional MRI scanner consists of the following basic components:

- the magnet, which creates a strong magnetic field; the measurement unit is indicated in tesla (T). Modern MRI magnets usually provide 1.5 T or 3 T fields for diagnostics, while in research, magnets with 0.1-9.4 T are also used.
- gradient coils that create additional magnetic fields to perform the spatial encoding of the signal and thus the image formation; gradient coils are used to generate a gradient of the magnetic field along different physical axes.
- the transmitting radio frequency (RF) coil for applying a radio frequency pulse to excite the tissue under examination
- receiving coils, located near the examined part of the body to receive the emitted signal from the excited tissue
- external computers for controlling and data processing

MRI technology is based on the physical properties of the nuclei of hydrogen atoms. These protons have a certain spin, which generates a magnetic moment. When exposed to a strong external magnetic field, the magnetic moments are aligned in parallel and anti-parallel directions with respect to the field, with more spins aligning along the parallel direction. These extra spins are what generate the MR signal. When aligned along the field, the sum of the magnetic moments leads to a macroscopic magnetization in an equilibrium state. To attain the necessary resonance signal for generating an image, a radio frequency (RF) pulse is transmitted into the stable system to flip the spins out of the static field alignment to a pre-defined angle, the flip angle. As the macroscopic magnetization is flipped, it begins to precess around the static field with a characteristic frequency (*Larmor frequency*), inducing a voltage which is received as a signal by the receiving coils.

Spatial Localization When selecting a particular slice for measuring, two important properties of the Larmor frequency are utilized. First, the Larmor frequency is proportional to the magnetic field strength: the stronger the field, the higher the frequency. Second, the spins are excited only if the RF pulse has the same frequency as their Larmor frequency. To receive a signal from only one particular slice, additional superposing gradient magnetic fields are required. As the Larmor frequencies of the proton spins are proportional to the magnetic field strength, they adapt to the particular magnetic strength at a specific point. The gradient coils can be systematically adjusted within the three directions referred to as x, y and z, analogue to the three axes of a coordinate system. For slice selection, the gradient along the z-axis (which usually is along the body positioned in the scanner) creates a linearly changing magnetic field. By transmitting a particular RF pulse, only those protons with their frequency identical to the pulse are affected, while the neighbouring spins are not influenced due to the slightly different precession frequencies caused by the gradient coil.

To determine the spatial encoding within the slice, so-called frequency and phase encoding is used. Frequency encoding is achieved by affecting the precession frequencies within the selected slice by means of another gradient field along one of the remaining axes of the xy-plane, for example the x-axis. Thus, each column of the slice can be identified by its particular frequency. Since the MR scanner receives the signal from the whole excited slice, a mathematical transformation, the *Fourier transformation*, is applied to identify different frequencies in the signal that belong to different positions in the image along the x-axis.

Information from the third axis is received by phase encoding, which is applied by inducing phase differences between the proton spins. For doing this, another gradient field is transiently switched on, causing the spins to precess with higher frequencies on one end of the gradient field and with lower frequencies on the other end, which in effect causes phase shifting. When the gradient is turned off, the spins maintain their phase shifts while the precession frequency re-aligns to the external field. As different phase shifts are required to encode a 2D image, the measurement sequence is repeated several times with different phase-encoding gradients. The raw data obtained are sampled and digitized in a data matrix (*k-space*), in which the x-axis usually represents the frequency encoding direction and the y-axis correspondingly reflects the phase encoding direction. Each measurement of excitation and relaxation of spins is sampled as a line in the k-space. To complete the k-space, these measurements have to be repeated until all lines are filled. Different parts of the filled k-space correspond to different spatial frequencies: data near the center correspond to low frequencies, reflecting more general contours in the resulting image, while information about more precise details is obtained from high frequencies in the periphery of the k-space. Generating an image is achieved by applying an inverse Fourier transformation, in which each point of the k-space contains information about each voxel in the image and conversely, each voxel in the image contains information about all individual points in the k-space. Data sampling, or filling the k-space, can be realized with different trajectories, such as rectilinear (*Cartesian*), spiral or radial methods.

Repetition Time TR and Echo Time TE During an MRI measurement, the process of proton excitation and signal acquisition is repeated many times. The time between two succeeding excitation pulses is referred to as the repetition time (TR). The longer the TR, the more time is available for all excited spins to re-align to the external magnetic field. Thus, if the image is to be T1-weighted, a short TR is required to detect the differences between the relaxation times of the distinct tissues. In contrast, a long TR permits all types of tissue protons to equally re-align, resulting in reduced signal differences.

For a T2-weighted image the echo time (TE) of the signal is essential. This parameter indicates the time span that passes after the excitation pulse until the echo signal is measured. The echo signal refers to the signal arising from the transverse magnetization at the time of phase coherence after excitation. With a short TE, the excited spins are in phase coherence, emitting a good signal but showing less contrasts of the single tissues. The longer the echo time, the more distinctly the differences of T2 relaxation times turn out.

T1 and T2 relaxation: image contrast After the RF pulse has been applied, the transverse magnetization decreases primarily because of two processes. On the one hand, the magnetisation re-aligns along the static external field, a process which is referred to as longitudinal relaxation characterized by the time constant T1. This time constant is dependent on both the field strength and the composition of the tissue. On the other hand, the transverse relaxation (T2) describes the dephasing process of the transverse magnetization: immediately after excitation, one part of the proton spins is in phase coherence, i.e. this fraction shows synchronous precession. Due to their inherent magnetic properties, weak local changes in magnetic field strength occur, which lead to alterations in their precession frequency and provokes a decay of the transverse macroscopic magnetization. Both T1 and T2 are crucial factors for the image contrast: depending on the weighting of these parameters, the image is either T1- or T2-weighted. As different types of tissue exhibit variable proton characteristics, the time constants will differ for each kind of tissue, which produces distinctive signal strengths that are responsible for the image contrasts.

MRI: Basic sequences The individual steps described so far constitute the basis for an MRI measurement sequence, which consists of slice exciting RF pulses, switching on and off various gradients for localization as well as the reception of the emitted electromagnetic signal. Depending on the image requirements, multiple repetitions of a particular sequence are necessary. Common to all sequences is that at some point the MR signal needs to be excited and subsequently acquired, for which several types of sequences are available. In the following, the spin echo and gradient echo sequences are outlined in more detail.

Spin echo and Fast spin echo The spin echo (SE) sequence was one of the earliest methods used in the 1970s and still is widely applied in the form of the modified fast spin echo (FSE) sequence. The conventional SE method works with slice selective RF pulses that flip the macroscopic magnetization by 90° into the transverse plane. As phase decay proceeds very quickly, there is only a short time window of synchronous precession, such that the signal is difficult to detect. To solve this issue, a second RF pulse is transmitted with a flip angle of 180°, which reverses the dephasing effects: by rotating the protons by 180°, the precession direction remains the same, but the faster spins ahead and the slower spins that 'fell behind' refocus. Thus, the RF pulse is applied at half of the echo time, such that detectable transverse magnetization is re-assembled when the spins are in phase again. This is referred to as the echo signal.

The advantage of SE is its insensitivity to static field inhomogeneity, resulting in high image quality. The disadvantage of this technique is its speed: spin echo sequences are highly time-consuming due to a long TR. Before transmitting a new 90° pulse, the longitudinal magnetization needs to recover, which takes up to two or three seconds. However, the construction of one image requires multiple phase-encoding steps. When generating an image out of 256 repetitive measurements, each taking two seconds, the whole procedure takes about nine minutes.

Fast spin echo sequences are based on the same principle as SE but use a series of 180° pulses instead of one single pulse following the 90° excitation. For each of the resulting echos the phase-encoding gradient is changed such that multiple phase-encoding steps are achieved within one TR. Thus, by obtaining multiple lines in k-space within one TR, image acquisition time is considerably reduced compared to the SE technique.

Gradient echo and Fast Low Angle Shot As an alternative to the spin echo method, signal acquisition can be also achieved by applying a gradient echo sequence (GRE). Instead of transmitting multiple RF pulses during one TR, the GRE involves one single excitation and the application of a gradient with reversed polarity to trigger the echo signal. A MR pulse-timing diagram for a GRE sequence is depicted in fig. 1.4. First, a RF pulse causes the macroscopic magnetization to flip towards the transverse direction. In contrast to SE, however, the flip angle is much lower than 90°, which leaves the longitudinal magnetization largely unaffected and reduces the repetition time to a considerable extent. Next, the frequency-encoding gradient is switched on with negative polarity to accelerate the dephasing process in the transverse plane; subsequently, the polarity of the gradient is reversed. This induces a rephasing process of the spins, such that the transverse magnetization is re-aligned and hence re-built, received as the echo signal.

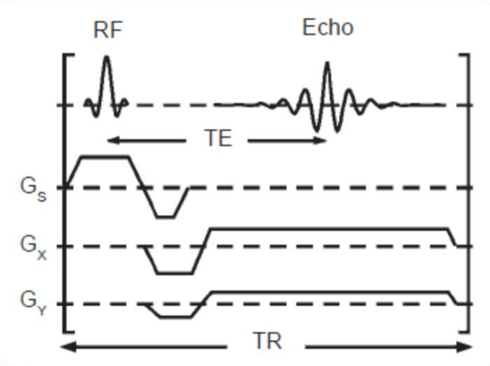


Figure 1.4: Basic gradient echo sequence (GRE). RF: radio frequency pulse, Echo: gradient echo, TE: echo time, TR: repetition time, G_S: slice-selection gradient, G_X, G_Y: frequency-encoding gradients. From Uecker et al., 2012, p. 463.

The work at hand involves analyses from MRI data acquired via the gradient echo sequence *FLASH*, which is described in more detail in the following. Invented and applied in the mid 1980s by Frahm and colleagues (Frahm et al., 1986; Matthaei et al., 1985), the formerly prevalent spin-echo method was replaced by the concept of a much lower flip angle induced by the RF pulse (about 10° instead of 90°), which was combined with a reversed gradient field. This sequence was named *Fast Low Angle Shot* (FLASH). In addition to the low angle, the repetition time of FLASH is usually much shorter than the relaxation time of the protons, which leads to a steady state after several excitations and a constant longitudinal magnetization. Thus, multiple rapid single recordings are possible and image acquisition time is dramatically reduced to the scope of seconds.

Real-time MRI with FLASH2 Since the invention of FLASH, further research aimed at reducing the acquisition time and providing a method for processing images in real-time. The term ‘real-time imaging’ in this context refers to "the rapid and continuous acquisition of image data sets followed by image reconstruction and visualization – preferably without noticeable delay." (Uecker et al., 2012, p. 461). Thus, besides the specific technical parameters and equipment, the second essential factor for carrying out imaging in real time is the mathematical condition of a fast reconstruction process.

The mathematical demands for image reconstruction are fulfilled by the FLASH2 method, developed to combine radial acquisitions with a specific mathematical approach to reconstruct highly undersampled data (Uecker et al., 2010, 2012). The basis for fast image reconstruction is an extensive reduction of the data amount required to generate one single image, i.e. data undersampling. With FLASH2, this is achieved by using radial instead of rectilinear trajectories for data space coverage. Relying on the frequency-encoding gradient alone, radial trajectories cross the center of the k-space and therefore capture both low and high frequencies in the data space. Data undersampling is thus provided by only a small set of equally distributed spokes that are turned for each successive image by a given angle, which leads to a temporally interleaved arrangement of the spokes (fig. 1.5).

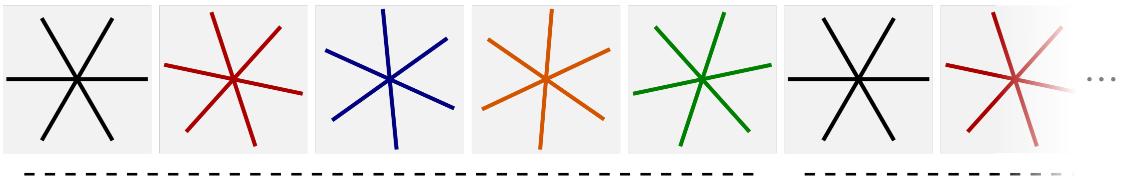


Figure 1.5: Undersampled temporally interleaved spoke arrangement. Radial trajectories are turned by a constant angle for each successive frame and repeated after an odd number of turns. From Kollmeier, 2021, p. 7.

Image reconstruction from the undersampled data is obtained by an iterative algorithm that defines the image as the solution to a nonlinear inverse problem: instead of calculating each new image from the raw data, an image is estimated by comparison to the raw data and the inclusion of the information from the previous image of a time series. This prior knowledge is used as an initial estimate for the iterative reconstruction process.

In summary, real-time MRI with FLASH2 is achieved by radial data acquisition with a high degree of data undersampling combined with an iterative nonlinear image reconstruction. This combination allows for high-quality MR imaging with a temporal resolution up to 100 frames per second.

MRI: phonetic research on velum function MRI and real-time MRI have been used in a number of phonetic studies to obtain insights about the gestural interactions during speech (see Ramanarayanan et al., 2018, for a review of phonetic MRI experiments and general opportunities for MRI data processing in speech studies). These studies include investigations of the articulatory coordination during vowels and coarticulatory effects on vowels (Demolin et al., 1997; Proctor et al., 2015), liquid consonant and fricative production (Lee et al., 2015; Proctor and Walker, 2012) as well as studies on the tongue position during nasal versus oral vowels (Carignan et al., 2013, 2015; Engwall et al., 2006; Silva and Teixeira, 2015; Silva et al., 2013) and velar coordination during nasals (Byrd et al., 2009; Carignan, 2018; Carignan et al., 2019, 2020; Proctor et al., 2013, see also Carignan et al., 2021). While all these studies provide valuable contributions to the basic understanding of articulatory coordination, only a handful specifically focused on the temporal and spatial behaviour of the soft palate during speech production. Some of them are considered in the following.

In an early MRI study on velar function during vowels, Whalen (1990) demonstrated clear variations in velum position for different English vowels: while both /a/ and /i/ were produced with a closed VP, the soft palate was found higher for /i/ than for /a/; a finding which is in line with former findings on velar function (e.g. Bell-Berti, 1973; Lubker, 1968; Moll, 1962). Similarly, the interplay between velum behaviour and tongue position was investigated by Demolin et al. (1998, 2003), who used the transversal plane to examine the velum opening process during four sustained French nasal and oral vowels. The nasal vowels were found to be produced with a more retracted tongue compared to their oral counterparts, but the front vowels were accompanied by a more open VP than the back vowels, with the narrowest VP registered for the nasal vowel /ɔ̃/. Serrurier and Badin (2005, 2008) utilized static MRI and computer tomography (CT) data from one French speaker to develop a three-dimensional linear articulatory model of the velum and pharyngeal wall. Tobin et al. (2006) and Byrd et al. (2009) presented results on varying temporal coordination processes of lingual and velar gestures during /n/ in different syllable positions and stress patterns by applying real-time MRI at a temporal resolution of ≈11 fps. Near-synchronicity of the tongue tip and velum opening gesture was found in onset position, while in the coda, velum lowering started before the tongue tip gesture. Proctor et al. (2013) investigated French and English data with respect to velum timing patterns at ≈33 fps. They found the velum lowering patterns in French nasal vowels differing from those of the oral counterparts, such that during oral vowels in CVN sequences, the velum remained in a high position until the tongue started moving towards the following nasal. In contrast, velum lowering in English started early in the vowel, similar to the pattern observed for French nasal vowels. Carignan et al. (2019) and Carignan et al. (2021) investigated velum movement patterns in German /Vnd/ and /Vnt/ sequences with respect to the question of how the post-nasal context affects the temporal and spatial amount of the velum lowering gesture. They found that, unlike in American English, the overall velum lowering gesture was not predominantly shifted more into the vowel in /Vnt/ compared to /Vnd/, but instead was above all slightly truncated in time and space. This study was based on the same corpus utilized for the thesis at hand.

MRI: The effect of gravity on velum behaviour As indicated previously, one limitation of applying MRI measurements in speech studies is the necessity to obtain data from participants who are in a supine position, a posture in which natural speech is easily possible but not typical in everyday conversation. Thus, referring to speech material produced in a supine posture to draw conclusions about natural speech production may be inappropriate to some extent, and the legitimate question arises whether gravity has an impact on the location and behaviour of the mobile articulators and in particular on the soft palate.

Some EMG studies reported an increased activation of the genioglossus posterior when speaking in a supine posture, probably to ensure an unrestricted airway. These activity patterns were assumed to have some effect on the lingual modifications during speech (Otsuka et al., 2000; Sauerland and Mitchell, 1975). In an ultrasound study, Stone et al. (2007) found considerable inter-speaker variability, with seven of thirteen participants showing a tendency for a more retracted tongue in supine posture. In a real-time MRI experiment, Engwall (2013) compared the gestural behaviour during sustained sounds with those from natural speech and additionally considered the effect of gravity on articulation by measuring himself in face upward and face downward position. Small differences were found, with a more retracted tongue and hence a more narrowed pharyngeal cavity in the face upward posture. While these gravity effects were rated as "moderate" (Engwall, 2013, p. 312), the author postulated that sustaining vowels for an unnatural long time span was decidedly more problematic for drawing conclusions about natural speech. The effect of a more retracted tongue in supine posture was also observed by Kitamura et al. (2005), who placed their participants into an open-type MRI scanner that enabled measurements in both a sitting and a supine posture. A tendency was found especially for back vowels to be produced with a more retracted tongue in supine position. However, the three speakers showed strong individual differences with respect to gravity. Similarly, with a special focus on the velopharyngeal structures, Perry (2011b) tested four participants in a sitting and supine posture by means of an open-type MRI scanner. For the sustained sounds /i/ and /s/, minimal changes were found for the two postures. However, during the production of /i/ in supine position, the soft palate showed a slight but significant difference in its height, exhibiting a lower position than in the upright posture.

In summary, most studies found small differences between the upright and supine posture with respect to the tongue position and sometimes also for the soft palate. However, results primarily referred to sustained sounds, not to dynamic speech. In addition, a high amount of variability was evident across the speakers with respect to the impact of gravity on the articulatory gestures. Along the lines of Engwall, who concluded that moderate gravity effects are less problematic than drawing conclusions from sustained sounds, real-time MRI is considered a powerful tool for obtaining data on natural speech, i.e. data on the interplay of the tongue, lips, larynx and the soft palate, which is otherwise difficult to access.

1.6 This thesis

The thesis at hand provides basic findings about the velum movement patterns during vowels and consonants produced by native speakers of Standard German. Data were originally acquired from 36 participants who were measured via state-of-the-art real-time magnetic resonance tomography (rt-MRI). As outlined above, velum lowering and the extent of vowel nasality are affected by different linguistic factors. Therefore, the speech corpus under investigation involves sound sequences of various segmental contexts with differences in vowel length, vowel height and the nature of the following consonantal context; in addition, the influence of the speaking rate and altering focus patterns is explored. Moreover, the question is considered to what extent German listeners are sensitive to coarticulatory vowel nasalization in perception and in particular, whether there is a relationship between the usage of coarticulatory nasalization in production and the perceptual sensitivity within one individual. The findings are finally discussed in terms of current theories of sound change based on phonetic principles.

1.6.1 Speech corpus

All experiments outlined in this thesis are based on one overall speech corpus originally obtained from 36 German participants. However, the actual results involve data from only 33 speakers due to issues with image registration for three of the participants and subsequent problems with generating the velum signal from their images. Depending on the specific research question, each chapter refers to a different subset of the overall corpus. The individual subsets are specified in the respective chapters. As the procedure and the analysis methods are identical for all experiments, the following sections provide some general information about the target words, the rt-MRI and acoustic recordings as well as the analyses of the images and acoustic data.

Target items The overall corpus consists of 152 German lexical items (see appendix table A.1) embedded in carrier phrases with varying prosodic conditions. Each speaker read out ≈350 stimuli. The items consist of monosyllabic and disyllabic natural words of German, which, if necessary, were inflected to achieve the required sound sequence. The target items were constructed in a way that allowed for exploring the final corpus with respect to different research questions. The overall corpus comprises the following groups of sound sequences (N=nasal stop, C=oral stop, F=fricative):

(1) *Speech material groups*

- **vowel height**: e.g. *Biene* [biːnə] 'bee' vs. *bahne* [baːnə] 'channel'; *biete* [biːtə] 'offer' vs. *bat* [baːt] 'asked for'
- **vowel tensity** (partly accompanied by vowel quality contrasts): e.g. *Toner* [toːnɐ] 'toner' vs. *Tonne* [tɔnə] 'ton'; *Saate* [zaːtə] 'seed' vs. *satte* [zatə] 'saturated'
- **N vs. C** e.g. *Sahne* [zaːnə] 'cream' vs. *Saate* [zaːtə] 'seed'
- **voicing contrast**: *Bunde* [bʊndə] 'league' vs. *bunte* [bʊntə] 'colorful'
- **C vs. NC**: e.g. *Dieter* [diːtɐ] 'Dieter' vs. *diente* [diːntə] 'served'
- **N vs. NC**: e.g. *lohne* [loːnə] 'worth it' vs. *lohnte* [loːntə] 'was worth it'
- **NF vs. F**: e.g. *schienst* [ʃiːnst] 'appeared to be' vs. *schießt* [ʃiːst] 'shoots'
- **NC vs. NF**: e.g. *Windeln* [vɪndəln] 'diapers' vs. *winseln* [vɪnzəln] 'whimper'

Carrier phrases and prosodic conditions The target items were placed into carrier phrases that basically exhibited two structures:

(2) *Carrier phrases for target words*

a) *Wieder* [target word] *gesagt/gedacht/gesehen/gehört/erkannt/erzählt*
'Said/thought/seen/heard/recognized/told...again'

b) *Bis er* [target word] *sagt/schreibt/erkennt/erklärt*
'Until he says/writes/recognizes/explains...'

To avoid adaptation effects and to obtain a consistent CVCV structure across the target word and the following verb, the final verbs of the carrier phrases varied in their structure and content. Depending on the final segment of the target word, the following verb exhibited either a word-initial oral consonant or a vowel followed by an oral consonant to ensure that velar closure was elicited after the target word was uttered. The target words were presented in three different prosodic conditions, presented in (3). Condition 1 contained a target word that was nuclear accented, such that the pitch accent occurred on the syllable with primary lexical stress (in our corpus always on the initial syllable). This condition refers to a stress pattern that is often used in a broad focus context, for which the carrier phrase (2a) was used in this study. Each of the 152 target words was recorded in this broad focus environment. The carrier phrase in (2b) served as environment to indicate contrastive focus on the target word versus contrastive focus on the final verb. Thus, the target word was either pronounced with focal accent or, when focal accent was on the final verb, the target word was pre-focal unaccented. To facilitate the correct intonation patterns uttered by the participants, either the target word or the final verb was underlined, as exemplified in (3.2) and (3.3):

(3) *Prosodic conditions for the target words*

1. broad focus: *Wieder Bunde gesagt.*

2. focal accented: *Bis er* <u>*Bunde*</u> *sagt.*

3. pre-focal unaccented: *Bis er Bunde* <u>*sagt.*</u>

Conditions 2 and 3 primarily involved words with contrastive voicing of the post-nasal consonant as well as word pairs with VNF vs. VF sequences. However, as it turned out, speakers had difficulties in condition 3 with consistently pronouncing the target word as focal unaccented, i.e. more damped and with a lower pitch than the final verb. Although the participants were instructed in a training session and corrected in their stress patterns if necessary (see section 1.6.3), many of them had problems in assigning contrastive focus to the verb instead of the target word. Therefore, the analysis of focus effects (chapter IV) does not include condition 3, but instead provides a comparison between the contrastive focus and broad focus conditions. In addition, the sentence structure of condition 1 was also used to create stimuli that were read with an increased speaking rate. This fast rate condition comprised target words with lax versus tense vowels followed by a nasal versus oral stop.

1.6.2 Speakers

Data were acquired from 36 monolingual native speakers of Standard German (22 female), recruited from the local university and a local sports team. Participants were aged between 19 and 35 years, the mean age was 24.36 years (SD=4.22). To ensure comparability with respect to the pronunciation patterns, only speakers with no substantial dialect impact were accepted. Detailed demographic information can be looked up in the appendix (table C.1). In addition, participants gave written information about the town and region in which they grew up and went to school. All participants reported normal hearing and speaking function and each participant filled out additional forms determining compatibility for an MRI measurement. Persons with fixed braces, non-removable piercings or older tattoos and older surgical metal were excluded as well as claustrophobic persons and pregnant women. All speakers gave written consent before the MRI measurement and were paid for their participation.

1.6.3 Procedure

Preparation Before the actual MRI recording, each participant was instructed in a separate preparation meeting, which took place in the same week as the MRI measurement. During preparation, participants filled out the required forms and gave written consent before they made themselves familiar with the reading task. Speakers were asked to read out the stimuli (i.e. the complete carrier phrases), which were presented in blocks on a notebook screen. Each block comprised 13 to 14 consecutive slides for the broad, focal and pre-focal accent conditions and 19 to 20 slides for the fast speech stimuli. The slides switched automatically after four seconds for all conditions except for the fast stimuli, for which the switching intervals were reduced to two seconds. Each block contained only one type of the four conditions. This was decided based on preliminary tests, in which participants showed difficulties in producing the correct prosodic pattern when conditions varied too quickly. Furthermore, each block started with one dummy sentence, such that participants had the chance to adjust to the specific condition. During the instruction session, participants sat in a quiet room and were asked to read out the stimuli sentences loudly. For the purpose of comparison to their later performance in the MRI machine, they were also recorded during their practice via the audio editor Audacity (version 2.2.2) with a sampling rate of 44100 Hz. For recording, the M-Audio M-Track 2x2 audio-interface (inMusic GmbH) was used. Participants were engaged with one specific prosodic condition until their intonation was deemed appropriate. After practicing all four conditions, at least one block of each condition was presented again to repeat the respective intonation pattern and to give speakers the chance to familiarize themselves with most of the stimuli. In total, the preparation session took about one hour.

MRI recording procedure Participants were measured via real-time MRI, which was carried out in the same week as the preparation session. The measurements took place at the Max-Planck-Institute for Biophysical Chemistry in Göttingen, Germany. Before entering the MRI machine, participants filled out further consent forms at the institute and were checked again for MRI compatibility. The measurement started with obligatory localizing scans to generate scout images for determining the plane. All images were obtained from a mid-sagittal slice with a thickness of 8 mm. In total, 25 reading blocks containing four prosodic conditions were recorded per participant. The sentences appeared on a screen projected onto a mirror just above the head coil. Depending on the exact number of sentences, one block took about 60 seconds of recording time. With the temporal resolution of 19.98 ms, ≈2800 images were acquired per block, which resulted in a total of 70.000–80.000 images per participant. While the order of the prosodic conditions was the same for all speakers, the stimuli within the blocks and the blocks within their specific condition were randomized to avoid habituation effects on the part of the speakers. At least two blocks of the same condition were consecutively presented before another condition was introduced. In addition, synchronous acoustic recordings were received by an optical microphone with integrated software for adaptive noise cancelling (Dual Channel-FOMRI, Optoacoustics, or Yehuda, Israel). The microphone was positioned central to the lips before the beginning of the measurement; the acoustic recordings were triggered by the specific measurement sequence (FLASH), which ensured synchronicity to the image recordings. After the measurement, the acoustic data were processed automatically using the Matlab software, such that the single measurement blocks were separated to facilitate further analysis. In addition, the audio system allowed for an easy communication between the instructor and the participant during the breaks to ensure that the speaker felt comfortable and to give specific instructions. Moreover, the speaker's performance was monitored during measurements, such that erroneous pronunciations were noted and the respective slides were presented once again after the final block. The overall measurement procedure took one and a half hours per participant.

Image acquisition For image acquisition, a 3 T MRI system was used (Magnetom Prisma Fit, Siemens Healthineers, Erlangen, Germany). Participants were measured in supine position via a 64-channel head coil with the RF-spoiled FLASH sequence (section 1.5.1). An in-plane resolution of 1.41 x1.41 mm and a slice thickness of 8 mm was applied. The field of view was 192x192 mm, covered by images of 136x136 pixels. The low flip angle of 5° allowed for the short repetition time TR of 2.22 ms and an echo time TE of 1.47 ms. Individual images were obtained from a single set of 9 spokes, which resulted in a temporal resolution of 19.98 ms or 50.05 fps.

1.6.4 Data analysis

Image analyses The images were processed in MATLAB (The Mathworks Inc., 2017, details in Carignan et al., 2020 and Carignan et al., 2021). For each speaker's data set, the images were first registered by pre-creating a region of interest (ROI) that covered the upper portion of the head. By this method, each image was aligned to the first image of the measurement such that small movements of the head that occurred during the recording were compensated. To create a velum signal, a second ROI was manually defined for each speaker around the spatial range boundaries of the velum opening and closing gestures. This ROI comprised approximately 600 voxel sites, which were defined as dimensions in a principal component analysis (PCA). As there was only one primary degree of freedom associated with the opening and closing gesture, the first principal component (PC1) necessarily referred to the velum movement and explained 52.7% (SD=9.4) of the data variance (cf. Carignan et al., 2021). To create the velum signal, the scores from PC1 were logged for each individual image, which is exemplified in fig. 1.6.

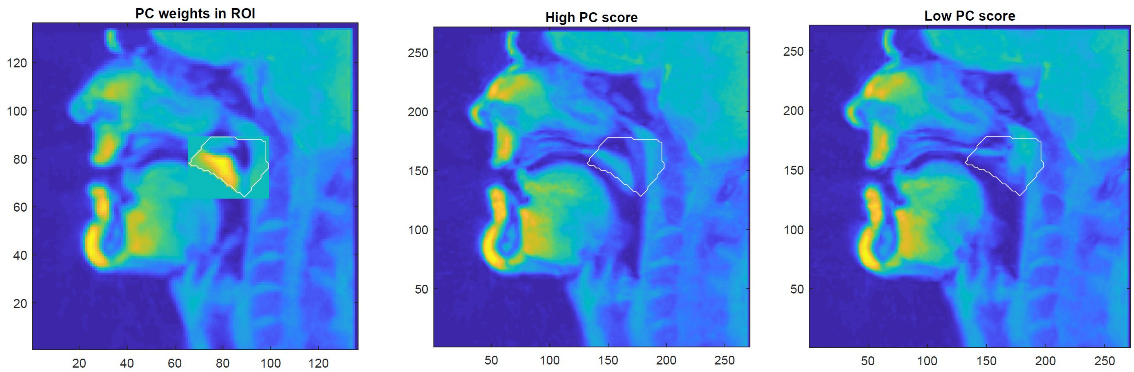


Figure 1.6: Image analysis by PCA. PCA loadings were estimated within a ROI (left); positive loadings are represented by the bright pixels. PC1 scores represent the correlation coefficient between each frame and the loadings. High scores are achieved if the velum closely resembles the positive loadings (middle). Low scores result from an image comprising a raised velum (right).

For each single image, a specific score was logged, which allowed for obtaining a time-varying signal with a sampling rate of 50.05 Hz (fig. 1.7, middle panel). Participants' individual morphology was taken into account by the fact that each speaker's data set was registered individually. The PC1 scores were scaled between 0 and 1, referring to the minimum and maximum PC1 scores for each speaker's data set. Hence, the values can be interpreted as follows: high values correspond to a low velum position and low values indicate a raised velum. The analyses referring to experiments I and V consider the velum signal at a specific time point, i.e. at the acoustic vowel midpoint. Experiments IV and

VI are based on the maximum velum signal (and also on the minimum signal in exp. VI) within the acoustic time span of a consonantal segment or sequence. For segmentation based on the acoustic analysis, see below.

Moreover, the velocity of the velar opening and closing gestures was considered, which was derived from the basic velum signal. Kinematic analyses were performed to determine several key time points during the overall velum movement gesture (fig. 1.7, lower panel). Where the velum movement signal reached 20% velocity thresholds, these points were determined as the onset and offset of the velum gesture (Kroos et al., 1997). However, the kinematic analyses provided in this thesis (experiments II, III) refer to the points of the maximum velocity during the opening and closing gesture, because by using these points, conceivable intermediate stages of lowering are excluded that might have occurred during the initial movements towards the actual opening or closing gesture.

To characterize the temporal aspects of the velum movement gesture, basically two temporal metrics were defined:

OVERALL VELUM LOWERING (OVL): the interval from the point of maximum velocity during velum opening to the point of maximum velocity during velum closing

VOWEL NASALIZATION: the interval from the point of maximum velocity during velum opening to the acoustic vowel offset

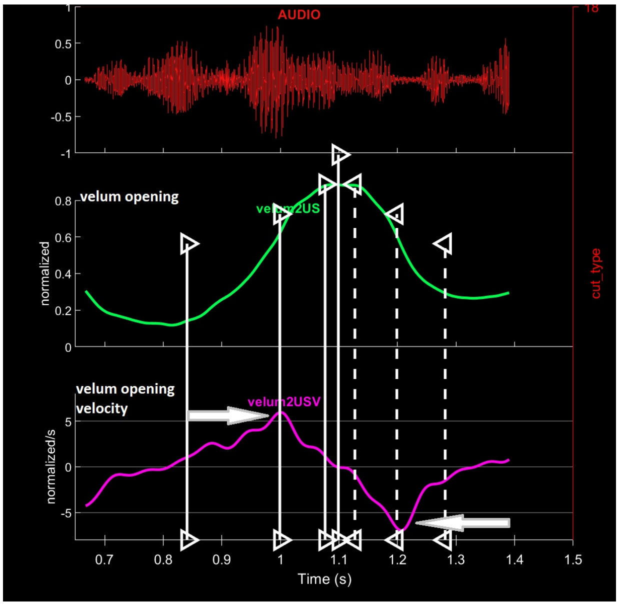


Figure 1.7: Time-varying signal obtained from PC1 scores (middle panel) for the sentence 'Wieder Sahne gesagt.'. High values represent a lowered velum, low values indicate a raised velum. The lower curve represents the velocity derived from the time-varying signal. The white lines indicate velocity key points. The arrows point to the maximum velocity during the opening and closing gesture.

The duration of the plain nasal segment is not considered, which is due to the following consideration. Vowel nasalization is defined as the interval between the point of maximum velocity during velum opening and the acoustic vowel offset. If nasal duration were to be defined in our data, this interval would refer to the difference between the acoustic vowel offset and the point of maximum velocity during velar closing. One research question might consider whether the duration of the nasal is correlated with the duration of vowel nasalization. However, the fact that these two intervals share one common boundary would automatically lead to the result that they were negatively correlated. Such an erroneous negative correlation is induced by inevitable measurement errors in manual analysis: if one interval is slightly increased, then the adjacent interval is necessarily slightly decreased (Ohala and Lyberg, 1976). As the correlation in our data would be calculated based on single pairs that all include small measurement errors, the result would inevitably be overall negative. This point will be reconsidered in chapter III.

In addition to the analysis of the velar movements, a second method was applied to capture the lingual and pharyngeal movement patterns during the target words, which is here referred to as *vocal tract aperture function* (VTA) (procedure described in Carignan et al., 2020). After image registration, a semi-polar grid consisting of 28 lines was applied semi-manually to the vocal tract, reaching from the glottis up to the alveolar ridge (fig. 1.8 left). This was achieved by manually selecting the locations of the glottis, velopharyngeal port and alveolar ridge as well as a location of air. The midpoint of the line from the alveolar ridge to the glottis was accordingly located within the genioglossus muscle in all subjects and served as the origin for the semi-polar grid. The closest grid line to the air selection was used to compute a threshold of pixel intensity, with low pixel intensities representing air and high intensities indicating flesh. The threshold was defined as 25% of the pixel intensity range along this grid line and further used for the estimation of the vocal tract constriction for each single grid line. Next, the posterior and superior edges of the vocal tract were detected semi-automatically based on the degree of pixel intensity change. Thus, the grid lines were terminated by the posterior or superior boundary (fig. 1.8 right). To quantify the lingual and pharyngeal movement patterns over time, the pixel intensities were scaled between 0 (i.e. all pixel intensities below the 25% threshold) and 1 (all intensities above the threefold of this threshold) for each participant individually, such that vocal tract constriction values could be determined by calculating the sum of the pixels between the 25% threshold and the threefold of this boundary along each grid line. By establishing the upper boundary, a plateau was created that did not further differentiate between the pixel intensities beyond the threshold, which helps to maximize the contrast between the presence and absence of tissue. For the purpose of interpretability, the grid lines were grouped into five different articulatory regions, i.e. the alveolar, palatal, velar, hyper- and hypopharyngeal area. These areas were covered by a specific number of grid lines dependent on the individual speaker (alveolar region: 3-4 lines, palatal: 6-7 lines, velar: 6-7 lines, hyperpharyngeal: 5-6, hypopharyngeal: 6-7 lines). For each region, the means of the scaled intensities were calculated across the respective grid lines. Accordingly, high values refer to high constriction and low values indicate less or no constriction of the respective articulator.

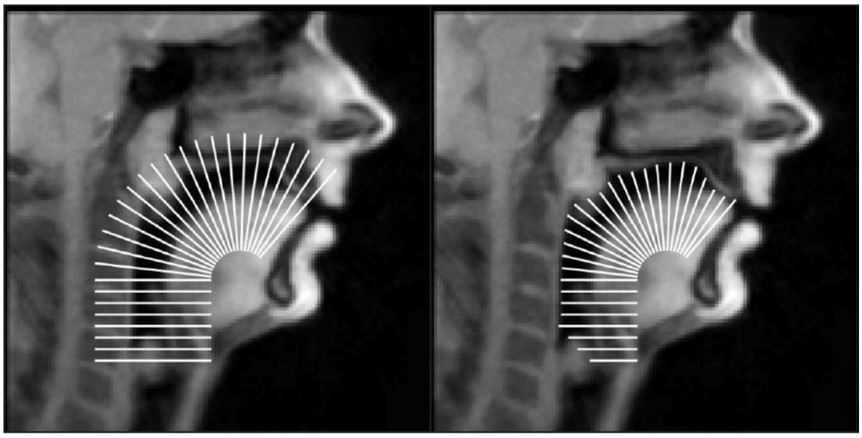


Figure 1.8: Semi-polar grid lines for the estimation of vocal tract aperture values. Left: basic grid. Right: terminated grid lines by estimated tissue-air boundaries.

Acoustic analysis The acoustic data were processed via MATLAB (version 9.3.0.713579, R2017b) to achieve further noise cancelling of the scanner tone. Acoustic analyses were performed manually by means of the PRAAT software (Boersma, 2017) as follows. Each audio track containing one complete sentence was labelled on four tiers: the sentence, the target word, the target vowel and the target coda (i.e. the post-vocalic consonants). Figure 1.9 illustrates the labelled tiers of the stimulus sentence *Bis er Küste sagt* 'Until he says coast', with the target word *Küste* [kʏstə]. The acoustic boundary of the sentence-initial word onset was defined as the point of stop release in the case of *Bis* [bɪs] or as the point where spectral and acoustic energy was identified for the fricative in *wieder* [viːdɐ]. The offset boundary of the sentence-final word coincided with the transition into silence. The onset of the target word was determined by identifying spectral changes in the transition from the preceding vowel into the initial consonant of the target word (second tier 'w' in fig. 1.9). The word offset boundary was defined either at the end of the stop closure in case of a final stop, omitting the aspirated portion, or, in case of a final nasal or fricative, at the point where spectral changes were observed from low amplitude (or high frequency noise, respectively) to clear formant structures. The onset of the target vowel was defined either as the point of release of the preceding stop (i.e. pre-consonantal aspiration counted as part of vowel) or, if preceded by fricatives, as the transition changes from high frequencies into clear formant structures (i.e. abrupt modifications in F1, F2, F3). The vowel offset boundary was defined either at the oral stop closure or at the transition into clear spectral frequency changes (third tier 'v' in fig. 1.9). The coda onset coincided with the vowel

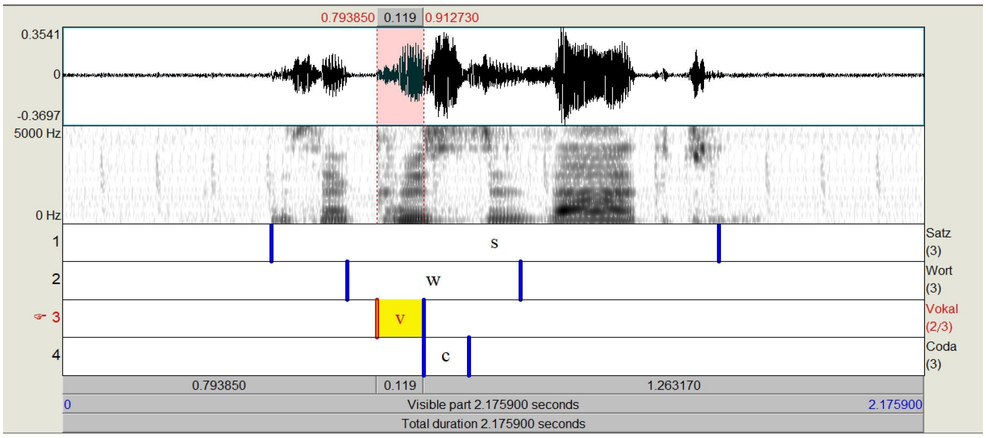


Figure 1.9: Illustration of the acoustic analysis of the stimulus sentence *Bis er Küste sagt*. The audio track was labelled on four tiers: total stimulus sentence (s), target word (w), target vowel (v), target coda (c).

offset. When the target word was monosyllabic, the coda offset and the word offset were identical. In case of disyllabic target words ending in a vowel, all consonantal segments in between vowels (e.g. [st] in [kʏstə]) were defined as coda, omitting the aspirated portion (tier 'c' in fig. 1.9). In addition, the acoustic energy apparent in the oscillogram was used for validation.

1.6.5 Structure of this thesis

This thesis is structured as follows. Each of the following chapters provides a specific analysis of the velum movement patterns with respect to one particular linguistic aspect. Some chapters additionally consider the tongue position. Chapter II is engaged with the relationship between velum height and vowel height in Standard German. After a review of the main findings from prior research on other languages, data are presented for the velum position in different tense and lax vowels that are followed either by a nasal stop or an oral stop. Chapter III explores the parameter of vowel length affecting the temporal extent of velum lowering, suggesting that the duration of vowel nasalization is related to the duration of the overall velum lowering comprising both the vowel and the nasal. This assumption is tested for a data set containing tense and lax vowels that are followed by a nasal-vowel or a nasal-stop-vowel sequence. In addition, the impact of the post-nasal context on vowel nasalization and the overall extent of velum lowering is investigated. Chapter IV is concerned with prosodic factors that are expected to show some impact on the velum during vowels and on tongue tip and velum movement patterns during nasal-oral stop sequences. Target words produced in the broad focus condition are compared to words provided with contrastive focus. Moreover, the effect of the speaking rate on the lingual and velar movements is considered in detail. Chapter V focuses on the perceptual aspect of vowel nasalization in German and provides results from an adaptive discrimination experiment, in which the speaker-turned-listeners were tested for their perceptual sensitivity to vowel nasalization when presented in synthetic stimuli. It is further explored whether for individual language users the perceptual sensitivity is related to the extent of coarticulatory vowel nasalization in production. Finally, the findings are summarized and considered in a general discussion to address the question of whether they can be associated with factors that might be involved in the initiation of sound change.

Chapter II

The effect of tongue position on velum height

Abstract

Velum position is investigated as a function of vowel height in tense and lax vowels preceding a nasal or oral stop. The primary aim is to investigate whether vowel height and velum position are interdependent, i.e. whether high vowels are produced with a velum that is raised to a greater extent than in mid-high and low vowels. Previous research on this topic has inconsistent findings: while some studies report a strict relationship between velum position and vowel height, others cannot confirm this connection in the strict way postulated. Common to most of them, though, is the finding that the low vowel /a/ exhibits the lowest velum position in nasal contexts, and some studies find velum lowering even in oral contexts. With respect to German, no study to date has investigated velum lowering patterns, so our findings may be seen as a contribution to further insights into velar behaviour during different vowels in a language without contrastive vowel nasalization. Our results show that pre-nasal vowels are generally produced with a lower velum than vowels preceding an oral stop, as expected. Within the nasal context, the velum is lowest for /a/, but otherwise no strict correlation is found for the mid-high and high vowels tested. In addition, tense and lax vowels differ in terms of the degree of velum lowering, such that tense vowels show more distinct differences in velum lowering than lax vowels. With respect to the oral context, no significant differences in velum lowering are found in tense or lax vowels.

2.1 The effect of tongue position on velum height

2.1.1 Introduction

Reports about the relationship between vowel height and velum position have been an object of research dating back to the 19th century. Changes of velum position during speech were already described by Brucke (1876) and Passavant (1869) and were later verified for vowels in nasal and oral contexts by multiple experiments using various languages. Fundamental contribution to the knowledge about velar function during speech was provided by numerous studies in the period from the 1960s up to the early 1990s, when researchers used diverse measurement techniques to illuminate velar movements during speech production in general and during specific vowels in varying environments. The most relevant studies are outlined below, followed by some more recent experiments dealing with velum position and vowel height.

In an early production study using cineflourographic films, Moll (1962) investigated velopharyngeal closure in vowels as a function of varying consonantal contexts for ten speakers of American English. In this experiment, as expected due to coarticulatory effects, vowels in nasal environments were found to exhibit velopharyngeal opening to a considerable amount, but even in oral consonantal contexts velum lowering was observed. Sustained in isolation, the low vowels /a/ and /æ/ were produced more often with an open velopharyngeal port (VP) and a lower velum than the high vowels /i/ and /u/. Within these vocalic height groups, though, no appreciable differences in velar height were found, i.e. lingual fronting effects on the velum position were not observed. Considering the pre- versus post-vocalic effects of /n/ on the vowel, a greater mean of distance between the velum and the pharyngeal wall was found in pre-nasal vowels than in post-nasal vowels.

The finding that the tongue height and the position of the velum are related has been frequently reported in subsequent research (Bell-Berti, 1973; Bell-Berti et al., 1979; Clumeck, 1976; Kuehn, 1976; Lubker, 1968; Moll and Shriner, 1967; Rossato et al., 2003). In a cineflourographic study, Moll and Shriner (1967) found higher velum positions during sustained /u/ than during /a/ and also a slightly more raised velum in nasal consonants compared to rest position. Based on these observations, they suggested a model in which velum activity was viewed as an on-off switch with only two stages: one for nasal sounds and utterance breaks and one for all other sounds. The intermediate stages observed during utterances, such as velum height differences between /u/ and /a/, were accounted for by mechanical forces of the surrounding lingual and pharyngeal muscles rather than by controlled multiple-staged activity levels.

The general tendency for low vowels to be nasalized to a higher degree was further observed for six languages by Clumeck (1976), who used a nasograph[1] for data acquisition. The findings from this study showed that for all six languages nasalization of low vowels preceding a nasal consonant was generally induced by an earlier lowering gesture compared to high vowels. However, languages differed in the onset of the lowering gesture, with speakers of American English and Brazilian Portuguese showing early velum lowering, while speakers of Hindi, French, Swedish and Amoy Chinese initiated the lowering gesture at a later point. Language-specific differences were also observed with respect to the lowering amplitude. In languages that exhibited velum position differences in the vowels, a lower velum was consistently found for low vowels compared to high vowels. In addition, during vowels in non-nasal contexts, only speakers of American English showed velum lowering, which was not reported for the other languages in this study.

Further insights into the velopharyngeal mechanism were obtained by other measurement techniques that provided more detailed information about muscle activity during speech, such as the method of electromyography (EMG, see section 1.5). In an experiment with EMG recordings and simultaneous cineflourographic films, Lubker (1968) demonstrated a high correlation between velar positioning and velar electromyographic activity as well as a strong relationship between tongue position and velum position: again, high vowels were found to be produced with a higher velum compared to low vowels. Electromyographic activity varied systematically with velar height: high velum positions were accompanied by increased EMG potentials[2]. In contrast to the on-off mode of the velopharyngeal activity suggested by Moll and Shriner (1967), the author postulated a complex multi-staged pattern of velopharyngeal muscle activation. Moreover, he argued for a perceptual approach explaining the differences in velum position rather than relying merely on anatomical constraints (cf. the exploratory approaches below).

Related to the observation of a correlated lingual and velar behaviour is the question of whether velum lowering – analog to velum raising – is a process induced by an interplay of the velopharyngeal muscles or by muscle relaxation. As outlined in section 1.5, several EMG experiments concerned with this issue showed that especially the levator palatini and palatoglossus muscles turned out to be of particular interest (Bell-Berti, 1973, 1976; Lubker et al., 1970; Lubker, 1968). Their findings, though, showed inconsistencies in some aspects.

The levator palatini is a muscle that has been identified to be primarily involved in raising and retracting the velum during speech. For example, by testing one Swedish

[1] A nasograph utilises a thin tube with a light source and a light sensor. The tube is inserted into the participant's nasal cavity up to the VP so that the light source is positioned below the soft palate and the light sensor within the nasal cavity. The lower the velum, the more light is captured by the sensor and amplified as voltage for analysis. See also section 1.5.

[2] The strict correlation between velum height and muscle activity was not confirmed by Seaver and Kuehn (1980). Here, velum height was found to be associated with vowel height; however, muscle activity for the levator palatini differed across participants.

speaker, Lubker et al. (1970) found voiceless oral consonants to be generally produced with a higher levator palatini activity than voiced consonants. In contrast, during the nasal stop /n/, little or no activity of levator palatini was registered during nasal production. With respect to vowels, the authors found increased levator palatini participation in vowels that were preceded by a nasal compared to other contexts. When preceded by an oral consonant, palatini activity was higher for /a/ when following voiced /d/ rather than voiceless /t/ or /s/. Also, there was a strong tendency for higher EMG potentials in /i/ rather than /a/. Based on these findings, Lubker and colleagues concluded that palatal levator activity was more based on phoneme categories than on single phonemes, with a high variability within these groups, such as voiced versus voiceless consonants, high versus low vowels and nasal consonants. The authors interpreted these groupings as "predictable under the assumption that palatal levator activity is dependent upon where the palate is, and where it must go to prevent excessive nasal coupling.“ (Lubker et al., 1970, p. 19).

Similar to this study, Bell-Berti (1973) investigated velopharyngeal muscle activity in vowels and consonants for three American English speakers by means of EMG measurements. Generally, among the muscles under investigation, she found the most consistent activity patterns for the levator palatini for all three participants, who primarily used this muscle for velar closure during oral articulation. More precisely, levator palatini activity was found highest before the production of oral consonants, while no activity was detected before or during nasal consonants. For vowels in oral consonantal contexts, levator activity was found to be lower than for the consonants themselves. In oral contexts or following a nasal, the high vowels /i/ and /u/ tended to show higher EMG potentials for the levator palatini compared to /a/. In addition, when following a nasal consonant the EMG potential signal was registered earlier for high vowels compared to the low vowel. During vowels preceding a nasal stop, however, levator activity was suppressed, regardless of the vowel quality. The finding that levator suppression occurred shortly after the beginning in vowels preceding a nasal stop was also reported in other experiments (Bell-Berti, 1976; Bell-Berti and Hirose, 1975; Fritzell, 1969; Henderson, 1984).

As a direct antagonist to the levator palatini, the palatoglossus muscle has often been suggested to affect the velum position by supporting the lowering gesture during the production of nasals. As delineated in section 1.3, the palatoglossus constitutes a connection between the palatal aponeurosis and the lateral margins of the tongue. However, as Perry (2011a) notes with reference to Moon et al. (1994), muscle activity during speech is less consistent across individuals. This might be due to anatomical variations of the local muscular attachments on the soft palate (Kuehn and Azzam, 1978; Perry, 2011a). Indeed, several EMG studies have indicated variability in palatoglossus activity for individual participants and languages (Bell-Berti, 1973; Bell-Berti and Hirose, 1973; Lubker et al., 1970; Seaver and Kuehn, 1980). For example, in their data from one Swedish participant, Lubker et al. (1970) found increased palatoglossus activity before the production of the nasal stop, leading the authors to assume that this speaker used the palatoglossus muscle to lower the velum actively. Similar muscle activity was only found during the production of /u/ and concomitantly with strong levator palatini activity. Based on these results, the authors suggested that the observed EMG potentials might have occurred either to assist the tongue in raising for the high back vowel or due to an antagonistic contraction during levator activity (Lubker et al., 1970, p. 18).

In contrast to these findings, the data presented in Bell-Berti (1973) did not confirm palatoglossus participation before or during the production of consonants, with the exception of the oral and nasal velar stop. However, two of the three participants showed EMG potentials for the palatoglossus for all vowels, the highest being for the low vowel /a/, while the third participant exhibited greatest EMG potentials for velar consonants in back vowel contexts. The author proposed that the tongue body movement and the tongue position were the relevant factors for inducing palatoglossus activity, rather than the manner of articulation (i.e. nasal versus oral sounds). Hence, in combination with her findings for the levator palatini, Bell-Berti suggested

> "that speakers of American English do not use increased activity in any muscle to produce nasal articulation, but rather produce such articulation by decreasing the activity in those muscles which are responsible for oral articulation." (Bell-Berti, 1973, p. 172).

On the other hand, further EMG studies on palatoglossus tension revealed a close relationship between the levator palatini and palatoglossus activity. By testing one speaker of Hindi, Dixit et al. (1987) found differentiated results with respect to front and back vowels. Palatoglossus activity was reported to be higher for central and back vowels than for front vowels, which held both for phonemically nasal and non-nasal vowels. In non-nasal vowels, an increase of palatoglossus activity and strong levator activity occurred simultaneously, while a more differentiated pattern was observed for the nasal vowels: for front vowels, a decrease in levator activity was apparent with a synchronous increase in palatoglossus activity, whereas the central and back vowels showed a much earlier increase in palatoglossus activity relative to levator palatini decline. Based on these data, the authors suggested that the palatoglossus muscle performed different functions for different vowel categories: during front nasal vowels, it served to actively lower the velum, while for central and back vowels, it was mainly associated with the tongue body movement. Velar lowering during back nasal vowels could then be seen as a side effect of continuous activity of the palatoglossus during levator palatini suppression (Dixit et al., 1987, p. 224).

Later experiments on velum height and tongue position largely confirmed the earlier findings. Demolin et al. (1998, 2003) examined velum opening processes during four sustained French nasal and oral vowels. The four participants showed inter-speaker variation with respect to velar behaviour during the vowels except for /ɔ̃/, which overall had the narrowest VP. For the other vowels, no strict order of VP opening differences could be determined across the participants: while one speaker showed the largest VP opening in /ɑ̃/, others produced /ɛ̃/ and /œ̃/ with a more open VP. In addition, individual anatomical structures did not allow for the definition of one specific VP value to create the acoustic effect of nasalization, as has been suggested by e.g. Maeda (1993). For oral vowels, /a/ again was found to be produced with incomplete velar closure, in contrast to the other vowels.

Rossato et al. (2003) presented EMA data (electromagnetic articulography, see section 1.5) of oral and nasal vowels and consonants from one French speaker. In accordance with Demolin et al. (2003), they found the nasal vowel /ɔ̃/ to be produced with the highest and /ɛ̃/ with the lowest velum position compared to /ɑ̃/ and /œ̃/. For the oral vowels, the soft palate was lowest for /a/ again, higher for the mid vowels and highest for the high vowels. Interestingly, the range of velum lowering was partially similar for oral vowels and nasal consonants and even to a small extent for oral and nasal consonants, suggesting that nasal consonants sometimes were produced with a relatively high velum and oral consonants with a slightly lowered velum.

In a follow-up study, Amelot and Rossato (2006, 2007) provided similar EMA data from one French and one Belgium French speaker. Velum height was investigated for all nasal and oral sounds of the French phonemic system. Although clear differences were found between the nasal and oral sounds, again a small amount of height range overlap was noted. Regarding the nasal sounds, nasal vowels were found to be produced with a lower velum than nasal consonants, while no such difference was reported for the oral sounds. For oral vowels, the lowest velum was found for /a/, but no strict relation for the other vowels was observed. Furthermore, in a nasal environment, the distinction between the oral and nasal vowels was still evident, i.e. during oral vowels showing coarticulatory nasalization the velum was never as low as in nasal vowels. In addition, oral vowels showed more nasalization when following a nasal consonant rather than preceding it, an observation which was suggested to be language-specific.

In summary, the fundamental studies outlined above allow for the following general statements on the relationship between tongue position and velum height:

- high vowels are generally produced with a higher velum than low vowels (Bell-Berti et al., 1979; Clumeck, 1976; Kuehn, 1976; Lubker, 1968; Moll, 1962; Rossato et al., 2003)
- levator palatini activity is often higher for high vowels than for low vowels (Bell-Berti, 1973; Lubker, 1968, but see Seaver and Kuehn (1980) for different results)

- during low vowels preceding a nasal, the velar lowering gesture begins earlier than during high vowels in the same context (Clumeck, 1976)
- during vowels following a nasal, levator palatini activity for raising the velum occurs later in low vowels than in high vowels (Bell-Berti, 1973)
- for front vs. back vowels, no significant differences are reported regarding velar height (Clumeck, 1976; Lubker, 1968), although varying activities for palatoglossus and levator palatini are observed for nasal vowels (Dixit et al., 1987)
- findings for the role of the palatoglossus muscle are inconsistent: while some studies reported palatoglossus participation during velum lowering in nasal sounds and during tongue retraction (Dixit et al., 1987; Moll and Shriner, 1967), others could not establish a strict relation (Bell-Berti, 1973)

Exploratory approaches

From an articulatory perspective, the relation between velum height and tongue position has been partly explained by anatomical constraints as indicated above: the palatoglossus muscle is involved both in raising and retracting the posterior part of the tongue body and sometimes also in lowering the soft palate (Dixit et al., 1987; Moon et al., 1994). Thus, during the low vowel /a/ the palatoglossus contracts and induces some kind of mechanical pull-down effect on the velum. In agreement with this assumption, higher palatoglossus EMG potentials were found for central and back vowels than for front vowels (Dixit et al., 1987). However, this does not explain why the high back vowel /u/ is produced with a more narrow VP than /a/, since /u/ involves high palatoglossus activity as well but often less velum lowering than /a/. As an alternative, the differences in velum lowering might not be due to palatoglossus activation itself but to mechanical lingual constraints, as assumed by Moll and Shriner (1967), who proposed to

> "attribute the differences in velar elevation between high and low vowels to changes in the degree of restriction on velar movement by tongue position than to contend that the speaker adjusts velar muscle activity to achieve the degree of velopharyngeal closure required for producing a given vowel without nasal quality."(Moll and Shriner, 1967, p. 65f).

However, the aspect of "producing a given vowel without nasal quality" has turned out to be a remarkable issue when velum adjustment in different vowels is illuminated from a perceptual perspective. It has been frequently reported that (usually synthesized) low vowels tolerate a higher degree of velopharyngeal opening before they are perceived as nasalized, while for high vowels only a small amount of velopharyngeal opening is required (House and Stevens, 1956; Lubker, 1968; Maeda, 1993; Ohala, 1975). Small changes of nasal coupling were found to have a strong impact on the acoustics of high vowels, but much less on low vowels (House and Stevens, 1956; Maeda, 1993). In general, nasal and nasalized vowels exhibit a complex acoustic spectrum consisting of both oral and nasal formants from the coupled oro-nasopharyngeal cavities (see chapter V). When nasal coupling is induced, the first formant F1 decreases for low vowels (House, 1957; Serrurier and Badin, 2008), but increases for high vowels (Delvaux et al., 2002; Fujimura and Lindqvist, 1971), and the second formant F2 is reported to decrease for non-back vowels (Delvaux et al., 2002). Generally, F1 amplitude is lowered while its bandwidth is increased (Delvaux et al., 2002; Fujimura and Lindqvist, 1971; House and Stevens, 1956). However, these changes have different impacts on the frequency spectra of the specific vowels: while the high vowel /i/ is affected with minimal nasal coupling, the low vowel /a/ tolerates much more coupling before significant changes in the spectrum become apparent (House and Stevens, 1956; Maeda, 1993). Essentially, the spectral behaviour of nasalized vowels can be summarized as follows: "[T]he lower is the F1 of a segment, the less will it tolerate nasalization [...]; if two segments have the same F1, the one with the lower F2 will be less tolerant of invading nasalization." (Ohala, 1975, p. 301).

The perceptual approach thus presumes that speakers are aware of these fine-detailed spectral changes related to velum lowering and that they are capable of controlling the soft palate gestures such that higher vowels are produced with a higher velum to prevent nasal coupling which otherwise would distort the acoustic characteristics of the specific vowel.

2.1.2 Experiment I: Velum height as a function of tongue position

Predictions

The main goal of this chapter is to present data for velar lowering patterns during German tense and lax vowels in nasal and oral contexts. Based on the findings from the prior studies outlined above, the following assumptions are proposed for the German data:

Hypotheses: Velum lowering in vowels of CVNV and CVCV sequences in German

H1 Pre-nasal vowels exhibit a significantly higher degree of velum lowering than vowels preceding oral consonants.

H2 In nasal contexts, /a/ shows a higher degree of velum lowering than all other vowels.

H3 In oral contexts, /a/ shows a higher degree of velum lowering than all other vowels.

H4 In both contexts, vowel height directly corresponds to velum height: the higher the vowel, the higher the velum position.

H5 Front vowels are produced with a lower velum than back vowels due to less mechanical constrictions of the tongue in the velar area.

Speech materials

For the current analysis, the speech material consists of a subset of the overall speech corpus. The subset material is listed in the appendix (table A.2). Stimuli with tense and lax vowels in CVN(V) contexts (n = 19) and CVC(V) contexts (n= 19) were surveyed, with the target vowel referring to the first vowel in these sequences. In total, the subset included 1264 items from 38 target words, divided into 631 CVN(V) items and 633 CVC(V) items and embedded into carrier phrases that were read out with no specific focus on the target word (prosodic condition 1, see section 1.6.1). The target vowels consisted of /aː, a, iː, ɪ, oː, ɔ, øː, œ, uː, ʊ, yː, ʏ/. The structure of the items involved either monosyllabic CVN and CVC words or disyllabic CVNV and CVCV words with the primary accent on the first vowel. The second vowel was either /ə/ or /ɐ/. As mentioned in the introduction, the original overall corpus was designed to cover a broad spectrum of various questions on vowel nasalization in different contexts in natural words. For this reason, the data considered here are not perfectly balanced. For the analysis in this chapter, oral lax /ʊ/ and tense nasal /uː/ are without a lax nasal counterpart and nasal lax /ʏ/ and tense oral /yː/ miss an oral lax counterpart.

Participants and procedure

Detailed information about the participants and measurement procedure are outlined in section 1.6. The MR images were processed with the method sketched in section 1.6.4. As the images were synchronized with the acoustic recordings (section 1.6.4), this allowed for determining the velum signal at a specific acoustic point of time during the vowel. The analyses in the next section refer to the velum opening signal at the temporal midpoint of the vowels preceding a nasal or oral consonant. Statistical analyses were carried out in the programming environment *RStudio* (version 1.2.5033) by applying linear mixed models with the *lmer* function from the *lmerTest* package. To test the hypotheses previously outlined, the velum opening signal was selected as the dependent variable in all models run. The context (nasal, oral), vowel (/a, i, o, ø, u, y/) and tensity (lax, tense) were selected as the fixed effects and the random effects included the speaker and word onset. Where analyses referred to the nasal or oral context alone, the fixed effects included the vowel and tensity.

Results

Nasal vs. oral context Figure 2.1 shows the mean velum lowering differences for vowels in nasal (CVNV) versus oral (CVCV) contexts per participant, separated by tense and lax vowels[3]. The vertical axis shows the normalized values for the velum opening signal, where

[3]The advantage of showing differences rather than the raw data is to avoid a misleading impression of overall tendencies, disregarding variability within and between the speakers. This is because on the one hand, there is an unequal number of tokens per vowel and on the other hand, the distribution of the raw data might not reflect speakers' individual relation between the lowering degrees in nasal versus oral contexts. The mean differences shown here allow for identifying of whether a tendency is evident across

increased values correspond to a more lowered velum. Accordingly, lower values refer to a more raised velum. As expected, nearly all instances are positive, indicating that nasal vowels show decidedly more velum lowering than their oral counterparts. Moreover, tense /a/ in particular exhibits highly different signal values between the nasal and oral contexts. This observation is also captured by figure 2.2, which includes the raw data for each vowel and tensity across the speakers.

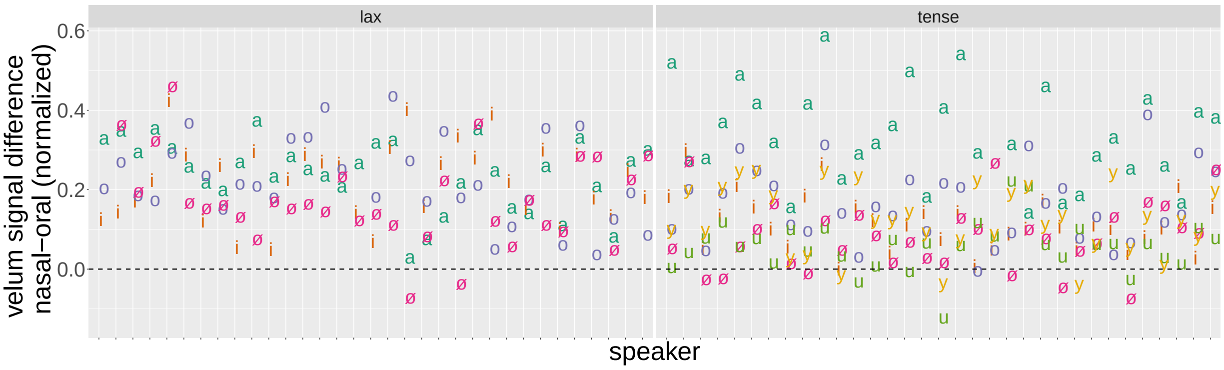


Figure 2.1: Mean differences of the velum lowering extent at the vowel midpoint for nasal-oral contexts. Values refer to differences for each vowel per participant, separated by tense and lax vowels. The ticks on the x-axis refer to speakers S03–S38, omitting speakers S11, S25 and S30 due to registration issues (see section 1.6.1).

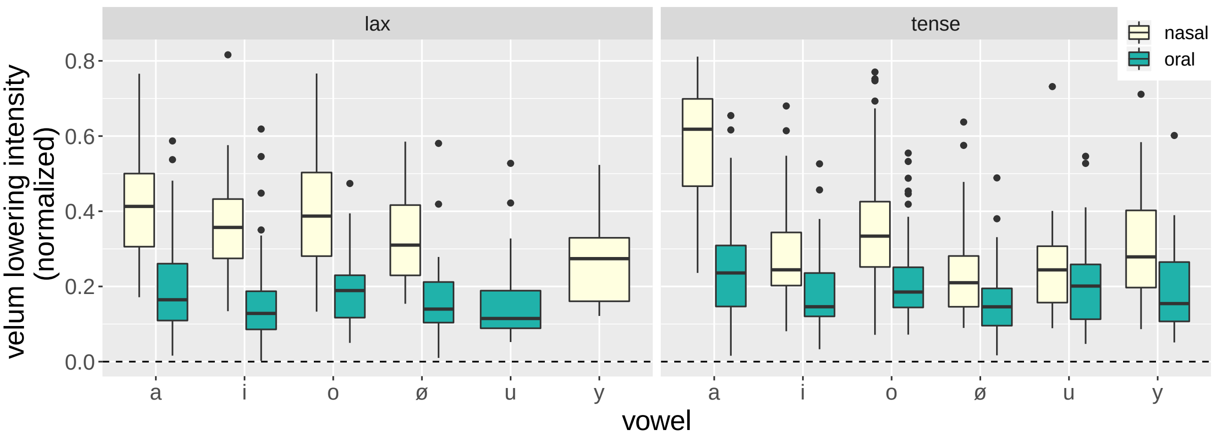


Figure 2.2: Velum lowering extent at the vowel midpoint in nasal and oral contexts across speakers. Velum intensity values are depicted for lax and tense vowels separately.

Statistical analysis confirmed the impression received from the plots: a significant main effect was found for the context ($F[1, 11] = 110.9$, $p<0.001$) as well as an interaction between the context and the vowel $F[5, 26] = 15.9$, $p<0.001$) and a three-way interaction between the context, vowel and tensity ($F[3, 14] = 10.1$, $p<0.001$). Post-hoc corrected bonferroni tests showed significant differences for all tense nasal versus tense oral vowels, though with various levels of significance (for /a, i, o, ø, i/: p<0.001, for /u/: $p<0.01$). With respect to the lax vowels, /a, o/ ($p<0.001$) and /i/ ($p<0.05$) showed significant differences in velum

individual speakers' velum lowering patterns.

lowering for the nasal versus oral contexts, but not /ø/ (p=0.0967). Lax /u/ and /y/ had no respective nasal and oral counterpart and thus were not taken into account.

Nasal context As might be assumed from the figures above, velum lowering was variably affected by vowels of different heights, which, however, did not hold equally for vowels in nasal versus oral contexts. Figure 2.3 shows the velum opening signal as a function of the vowels in nasal contexts only, with the velum signal means for tense and lax vowels per speaker (fig. 2.3, left) and the mean differences between tense and lax vowels per vowel category and speaker (fig. 2.3, right). Considering the overall lowering means, the low vowel /a/ generally appears to be produced with a significantly lower velum, more than the other vowels. Moreover, the differences between tense and lax vowels indicate that tense /a/ shows overall higher values than lax /a/, while the reverse tendency is apparent for /ø/, /i/ and /o/. Figure 2.4 depicts these findings for the velum signal values across the speakers. Note that this figure considers explicitly the nasal data that have already been depicted in fig. 2.2. The left side shows the overall signal values per vowel across speakers when pre-nasal tense and lax vowels are combined, while on the right side, data are given for tense and lax vowels separately.

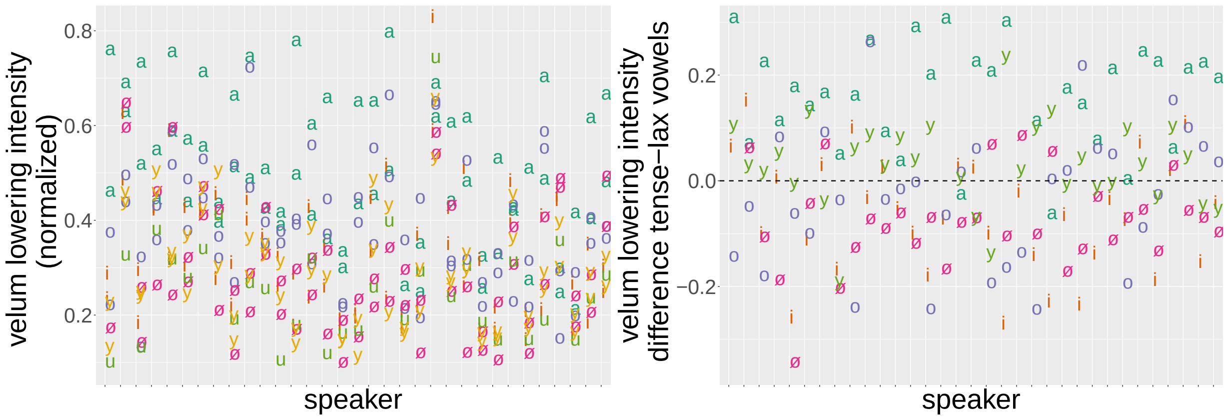


Figure 2.3: Velum lowering extent at the midpoint of the pre-nasal vowel in CVNV contexts. Left: separate mean signal values for tense and lax vowels per speaker, i.e. the left plot depicts both tensity categories per speaker. Right: velum lowering differences between tense and lax vowels (normalized). The ticks on the x-axis refer to speakers S03–S38, omitting speakers S11, S25 and S30.

A linear mixed model was applied to the data comprising solely the nasal context. Results revealed a significant main effect for the vowel (F[5, 22]= 13.4, p<0.001) and a significant interaction for vowel and tensity (F[4, 16] = 15.1, p<0.001). Pre-nasal lax and tense vowels apparently behaved differently with respect to the velum position: while for tense vowels, /a/ was produced with a decidedly lower velum than the other vowels (p<0.001), this was only reported for the contrast of lax /a/-/y/ (p<0.01). Further significant velum lowering differences in tense vowels were only observed for /o/-/y/ (p<0.01) and /o/-/u/ (p<0.05). Similar findings applied to the lax vowels, in which (besides /a/-/y/) only /o/ showed more velum lowering than /y/ (p<0.01). Otherwise, no significant differences were found between

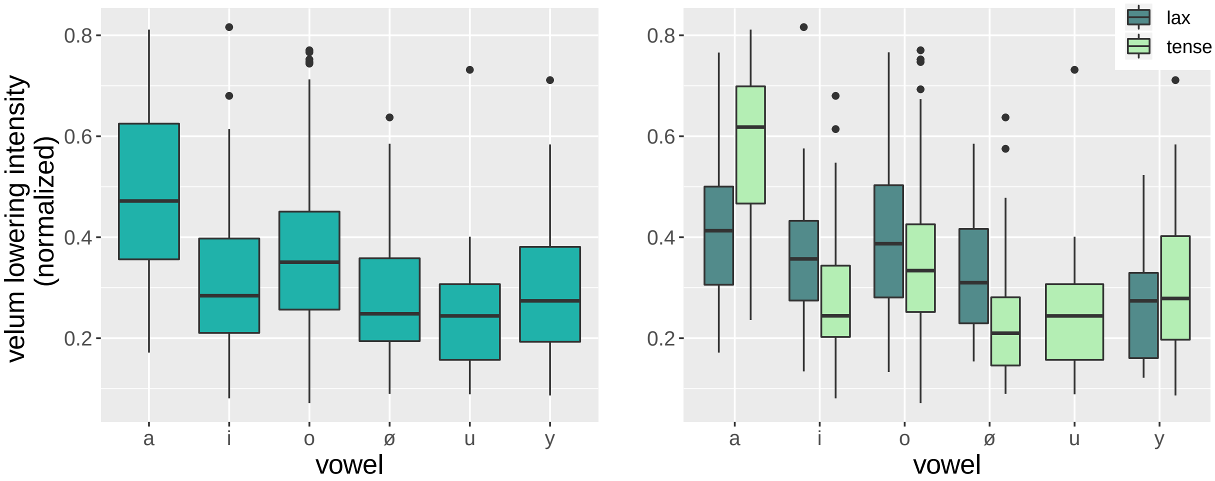


Figure 2.4: Velum lowering extent at the midpoint of the pre-nasal vowel in CVNV contexts. Left: overall signal values per vowel combining tense and lax vowels across speakers; right: velum lowering values for tense and lax vowels separately.

the remaining vowels. Considering the effect of tensity, the only contrast between lax and tense vowels was found for /a/, showing a lower velum in the tense vowel (p<0.001).

Oral context With respect to the degree of velum lowering in vowels followed by an oral obstruent, fig. 2.5 (left) scarcely reveals any distinct velum lowering differences between the vowels. This is in direct contrast to the pattern found for the nasal context, where the lowest velum position was clearly found for /a/. In the CVCV context, though, only a few speakers showed more extended lowering for /a/ compared to the other vowels. Considering velum differences between tense and lax vowels (fig. 2.5, right), some overall difference is evident with slightly higher values for the tense vowels. Very few speakers exhibited clear differences between tense versus lax /a/. Fig. 2.6 considers oral data that were also captured by fig. 2.2. Across speakers, the velum lowering patterns scarcely differ for the vowels in general (fig. 2.6, left), with tense vowels exhibiting slightly higher values than lax vowels (fig. 2.6, right).

Considering the statistical analysis, a linear mixed model applied for the oral context revealed an interaction between the vowel category and tensity ($F[4, 499] = 3.9$, p<0.01), but no main effect for the vowel alone. Post-hoc corrected bonferroni tests reported a significant velum signal difference only for tense /a/ versus /ø/ (p<0.01). The remaining vowel comparisons within their respective tensity category did not significantly differ. Regarding the effect of tensity, significant differences were found for /a/ (p<0.001), /i/ (p<0.01) and /u/ (p<0.001), with higher signal values for the tense vowel.

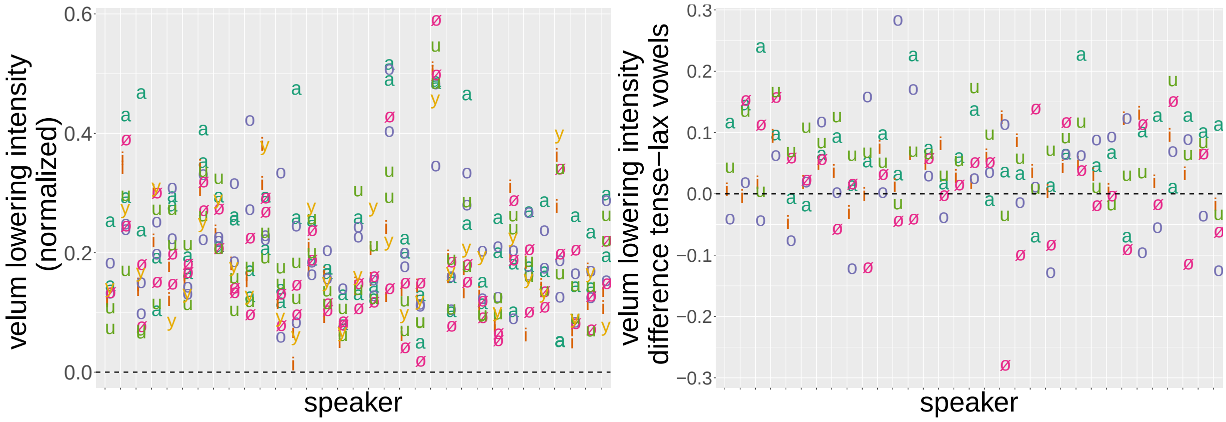


Figure 2.5: Velum lowering extent at the midpoint of the pre-consonantal vowel in CVCV contexts. Left: mean signal values for each tense and lax vowel per speaker, i.e. the left plot depicts both tensity categories per speaker. Right: velum lowering differences between tense and lax vowels per speaker (normalized). The ticks on the x-axis refer to speakers S03–S38, omitting speakers S11, S25 and S30.

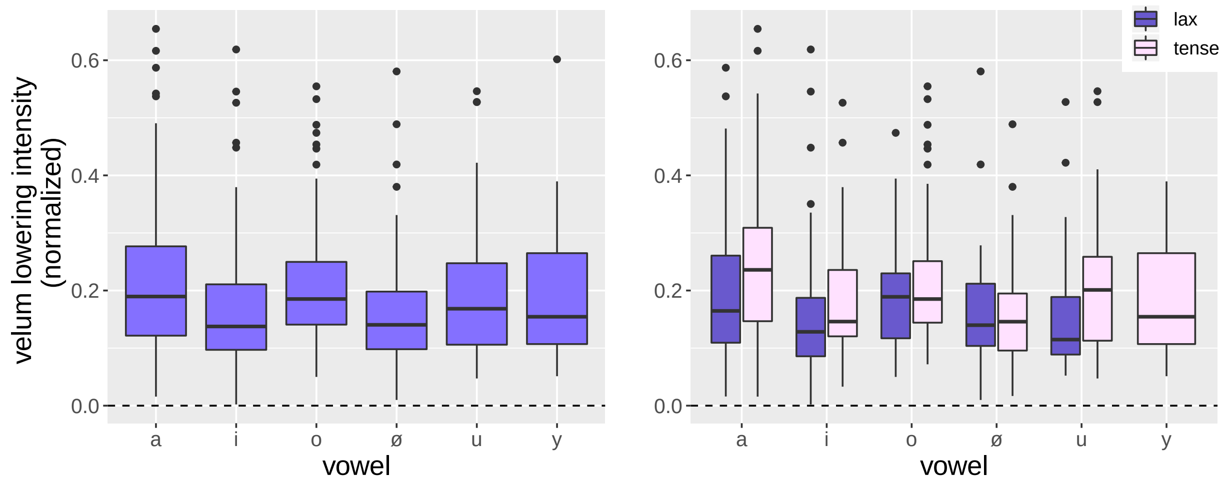


Figure 2.6: Velum lowering extent at the midpoint of the pre-consonantal vowel in CVCV contexts. Left: overall signal values per vowel combining tense and lax vowels across speakers; right: velum lowering values for tense and lax vowels separately.

Tongue position To illuminate the relationship between the tongue position and velum lowering in our data, the differences in tongue position are surveyed in more detail. If a systematic relation existed, it should become visible more clearly when the articulatory gestures are executed with a higher amplitude. Thus, the following analysis includes only tense vowels because on the one hand, tense vowels have been found to be generally articulated more in the periphery than the more centralized lax vowels (e.g. Hoole and Mooshammer, 2002) and on the other hand, the results above showed that the most distinct differences of velum lowering occurred between tense vowels.

To investigate the lingual behaviour, the values for the tongue position at the vowel midpoint were determined using the vocal tract aperture function (VTA), as described in section 1.6.4, which is based on changes of pixel intensities of consecutive images along a specific grid line. Figure 2.7 shows the values for the tongue signal of different regions: the palatal, velar and hyperpharyngeal area (for details on how these areas were conceived, see section 1.6.4). During vowel production, the tongue movements were captured in these regions, such that high values refer to high pixel intensity signals representing the tongue tissue during movements towards the particular area. Thus, fig. 2.7 illustrates the tongue position as a function of vowel height at the vowel midpoint. These values

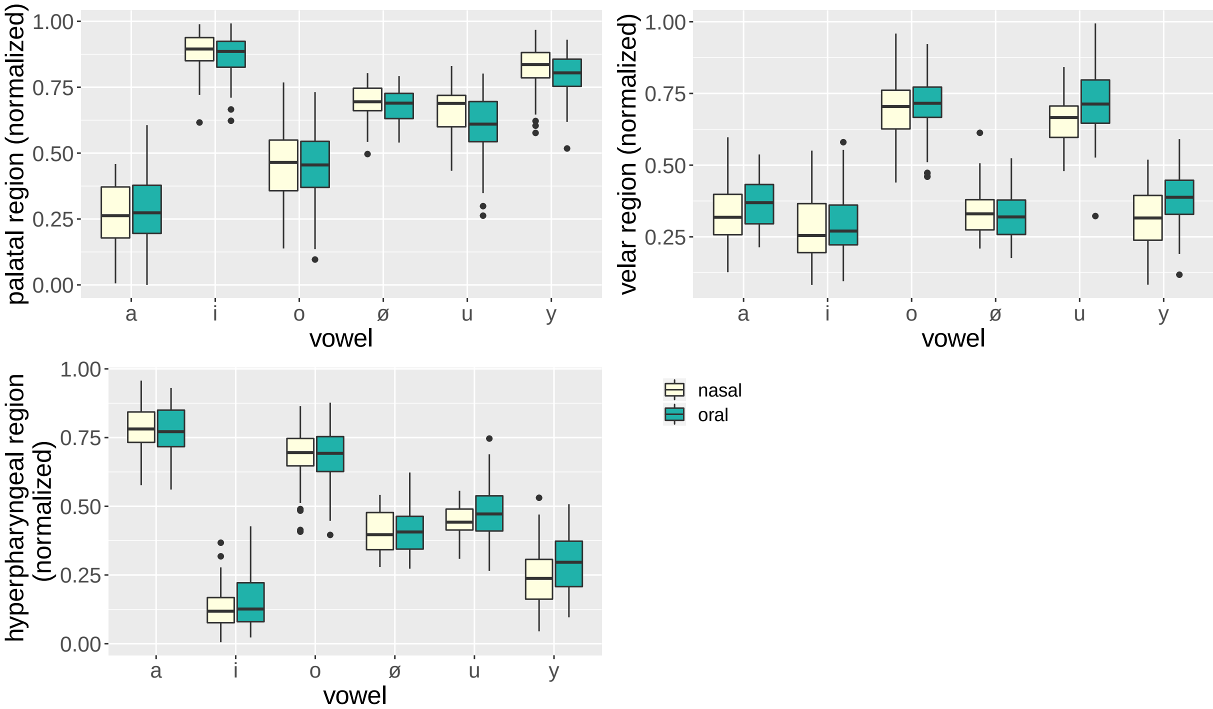


Figure 2.7: Tongue signal values at the vowel midpoint in tense vowels for three different regions (palatal, velar, hyperpharyngeal). The values correlate with pixel intensities reflecting the presence of lingual tissue.

basically demonstrate that the VTA method is well-suited to uncovering even fine details in tongue configuration. For example, considering the palatal region, /a/ and /o/ show low values, indicating that little or no tongue tissue was present around this area, as expected. However, even fine differences are revealed for /i/ and /y/, which both are traditionally classified as front high vowels. As indicated by the data (and confirming former reports, as e.g. Hoole and Mooshammer, 2002), /y/ was articulated with a slightly lowered and retracted tongue compared to /i/. Similarly, data in the velar region show higher values for /o/ and /u/, illustrating that these vowels were articulated with a high retracted tongue in contrast to the other vowels. In the hyperpharyngeal region, the highest values are related to /a/ and /o/, while /i/ and /y/ show the lowest values.

Considering the tongue position in terms of oral versus nasal contexts, one might expect some noticeable differences, as findings from other languages suggest some systematic variation of the oro-pharyngeal shape and tongue position in nasal vs. oral vowels (Carignan, 2014; Carignan et al., 2015; Shosted et al., 2012, 2015) or in vowels in nasal vs. oral environment (Mielke et al., 2017). The German data at hand, however, do not indicate any systematic differences, which suggests an equivalent range for the tongue movements during the vowels independent of the nasality context.

The tongue height differences between vowels and contexts were surveyed by statistical analysis. As the dependent variable, the tongue position values were selected for each region separately, resulting in three linear mixed models. For all three models, the fixed effects included the vowel (a, i, o, ø, u, y) and context (nasal, oral) and the random effects referred to the speaker and word onset.

The values for the palatal region indicate high signals for the high front vowels, but also for the high back vowel /u/. Statistical results revealed an interaction between the context and the vowel ($F[5, 598]= 3.65$, $p<0.001$). Post-hoc corrected bonferroni tests indicated that all vowels exhibited significantly different signal values compared to each other in both nasal and oral contexts ($p<0.001$) except for /ø/-/u/. Regarding the context effect, higher values were found in the nasal context only for /y/ ($p<0.001$).

For the velar region, results indicated a weak interaction between the context and the vowel ($F[5, 563]= 2.78$, $p<0.05$): all vowels in both contexts showed significantly different values compared to each other ($p<0.001$), with the only exception of /o/-/u/. In addition, oral /a/ and /y/ exhibited slightly higher values than their nasal counterparts ($p<0.01$ and $p<0.05$, respectively).

With respect to the hyperpharyngeal region, an interaction was found between the context and the vowel ($F[5, 520] = 4.46$, $p<0.001$). With respect to the differences between the vowels, all vowels significantly differed from each other in both contexts, except for /ø/-/u/ (nasal /a/-/o/ $p<0.01$, oral /a/-/o/ $p<0.05$, all other vowel pairings $p<0.001$). Nasality differences only affected /i/ and /y/, with higher values in the oral context ($p<0.01$ and $p<0.001$, respectively).

In summary, the data suggest that in German, unlike in languages with contrastive or at least more pronounced vowel nasality, no systematic variation of the tongue position is evident during vowels in oral vs. nasal environment.

2.2 Summary and discussion

The study and findings outlined above provide further insights into the basic principles of velar participation during vowels preceding nasal and oral contexts. Results are in general agreement with hypothesis H1 previously outlined: vowels preceding nasal consonants were commonly produced with a more lowered velum than vowels preceding oral stops. This is in accordance with findings from past studies in this field. However, while the effect was clearly demonstrated for velum lowering patterns in tense vowels, it did not hold for all lax vowels with statistical significance, although a tendency was apparent.

Vowel height Besides the analysis of the general effect of nasal versus oral contexts on the soft palate, the main issue was to investigate how vowel height differences affect the velum position. Data for nasal and oral contexts were surveyed separately. For vowels preceding a nasal stop, results indeed revealed differences in velum position: the low vowel /a/ was produced with a more lowered velum than all the other vowels, and a higher velum opening signal was found for the mid-high back vowel /o/ compared to /u/ and /y/. However, this relation was evident only for tense vowels. Pre-nasal lax vowels were all produced with a similar degree of velum lowering, with the only significant difference between /a/ and /y/.

Under the assumption that vowel height has some systematic impact on the position of the velum, it might not be surprising that the vowels under consideration did not significantly differ from each other, except for tense /a/. The remaining vowels /i, u, y/ and also /ø/ are generally produced with different tongue positions in the front versus back of the oral cavity and also with varying degrees of lip protrusion, but common to all of them is a relatively high tongue position during articulation. Therefore, if velum lowering is related to the tongue position in terms of tongue height, no differences would be expected to occur between vowels of similar height. In agreement with this assumption, velum lowering differences were found between the pre-nasal mid-high back vowel /o/ and the high vowels /u/ and /y/, at least for the tense vowels: tense /o/ showed the second highest signal values, which seems to be compatible with a relation concept between tongue height and velum position. On the other hand, no significant contrast was found between /o/ and the high front vowel /i/, which is not in total accordance with the strict correlation claimed by hypothesis H4.

With respect to lax vowels in nasal contexts, no unambiguous velum position contrast was apparent between the vowels except for /a/ versus /y/ and /o/ versus /y/. Although only marginally, /a/ and /o/ were again the only vowels with at least one significant contrast to a high vowel. The fact that lax /a/ did not differ from the other vowels as clearly as its tense counterpart is a remarkable point. While the relation among the other vowels roughly reflected the ratio seen for the tense vowels (with /o/ tending to show higher signal values than /i/, /ø/ and /y/), the pattern decidedly differed for lax /a/. One might have expected the opposite, i.e. clear velum lowering in pre-nasal lax vowels, because German lax vowels are generally shorter than tense vowels[4] and are therefore followed 'earlier' by the nasal consonant. Thus, an earlier lowering gesture during the vowel may be induced to ensure a sufficiently open velopharyngal port in time for the nasal stop. However, the data at hand do not support this scenario. Perhaps, it is just the short duration that contributes to a less extreme low position: it might take a certain time for changes in inherent muscle activity programming and execution (whether activation or relaxation) to occur, such that some anticipatory lowering during lax /a/ might be apparent but does not reach the lowering extent of the tense vowel. If measured just before the vowel offset, the velum position in lax /a/ might have been similar to tense /a/. In fact, as will be outlined in chapter III, velum lowering most often started after the vowel midpoint except for tense /a/, where the lowering gesture was initiated before the midpoint. Overall, results from the nasal data are

[4]Duration of /a/ in the current data: lax: mean = 101.14 ms, sd= 26.92 ms; tense: mean = 175.84 ms, sd= 26.56 ms.

compatible with prediction H2, but only with respect to the tense vowels: the low vowel /a/ is produced with a lower soft palate than the other vowels.

In terms of velum behaviour during oral vowels, no significant differences were evident in the data, which goes against prediction H3. Prior studies reported inconsistent results with respect to velum lowering patterns in oral contexts, where some of them found /a/ to exhibit the lowest velum position of the oral vowels surveyed (Amelot and Rossato, 2006, 2007; Demolin et al., 2003; Rossato et al., 2003), but others did not (Clumeck, 1976). However, where velum lowering in oral contexts was reported, results mostly referred to data from American English (Bell-Berti, 1973; Moll, 1962) and French (Amelot and Rossato, 2006, 2007; Demolin et al., 1998, 2003). Findings from other studies concerned with Swedish, Hindi, Amoy and also French did not show significant velum lowering patterns for vowels in oral contexts, which also included the low vowel /a/ (Clumeck, 1976). In addition, many authors reported large variations across and within the speaker groups of one language. Considering the low numbers of participants in these studies, the variation aspect should not be disregarded. The German data at hand from 33 participants reveals that, at least in this language, velum position is rather unlikely to be systematically affected in oral contexts by vowel height.

Front versus back vowels The results of the data presented also go against prediction H5: no systematic correlation was evident for velum lowering patterns of front versus back vowels. Considering the nasal context in which velum lowering was generally evident, the velum signal did not differ for tense /i/, /y/ and /u/, i.e. for vowels that commonly are produced with a high tongue position but varying front and back locations. Possibly, due to a more raised velum for high vowels, front versus back position effects might have been obscured. However, a similar pattern was found for mid-high tense /o/, which in fact showed higher velum signals than /y/ and /u/ but did not significantly differ from /ø/, a vowel that is also traditionally classified as mid-high. Lower tense vowels were not tested with respect to this question because the front versus back contrast in low vowels is not part of the German vowel system. On the other hand, the finding that in our data front and back vowels did not show any larger differences in velum lowering is in accordance with prior investigations in which no systematic relation across speakers was found as well (Moll, 1962; Seaver and Kuehn, 1980).

The remaining issue is a reiteration of the question previously asked with respect to the exploratory approaches: why is the low vowel /a/ predestined to be produced with a lower velum than the other vowels? Although this observation only applied to tense /a/ in our data, it is still worth some consideration. First, the palatoglossus muscle has been suggested to affect the velum position in specific vowels by mechanically pulling down the soft palate to some extent (Dixit et al., 1987; Moll and Shriner, 1967). This muscle runs from the lateral margins of the soft palate and inserts onto the lateral tongue margins. If activated, as e.g. during low and retracted vowels, its contraction might provoke some velum lowering due to its muscular connections between the velum and the tongue. If this pattern applied systematically, velum lowering differences would also be expected for /a/ in the oral context, which, however, was not evident in our data. On the other hand, the palatoglossus is a direct antagonist to the levator palatini, which basically functions to elevate the soft palate (Dixit et al., 1987; Lubker et al., 1970). Thus, it is conceivable that in German oral contexts, the levator strength might dominate that of the palatoglossus, because the latter is much smaller than the former, such that the soft palate is raised in spite of palatoglossus activity (cf. Kuehn and Azzam, 1978, p. 356). In contrast, little or even no levator activity is present in nasal contexts, such that the pull-down effect of the palatoglossus may become visible.

The other question is why back vowels in nasal contexts are not generally produced with a lowered velum, because the palatoglossus is active during tongue retraction not only for low vowels. Given the approach above, stronger lowering effects for pre-nasal /u/ and /o/ would be expected. On the other hand, in coarticulatorily nasalized vowels, a passively pulled down soft palate by palatoglossus constriction might be in conflict with a high retracted tongue, causing a higher restriction on velar movement.

Alternatively, the perceptual approach suggests that in high vowels, excessive velum lowering is avoided to prevent the formant structures from too large perturbations due to nasal coupling. Previous studies have provided evidence that listeners perceive high vowels as nasalized with only small changes in the formant frequencies, while larger distortions are tolerated for low vowels (House and Stevens, 1956; Maeda, 1993). Possibly, because heavily nasalized vowels are not part of the Standard German language, speakers provide sufficient closure of the velopharyngeal port to circumvent the undesired effect of a strongly nasalized sound. For /a/ in nasal contexts then, this concern can be neglected to some extent on the part of the speaker because the low vowel tolerates a good amount of velum lowering before it is perceived as nasalized in an unfamiliar way. The data in this chapter do not clearly argue in favour of one specific account. In fact, the physiological and perceptual approach do not exclude each other; they may even support each other: if the speaker lowers the velum on purpose, the palatoglossus may facilitate the lowering gesture; on the other hand, if the palatoglossus is primarily responsible for the low velum position during /a/, the acoustic result would still be acceptable for the speaker without the need to go against the lowering gesture. Thus, both factors may contribute to the phenomenon that tense /a/ is so often produced with a lowered velum and that it is also often affected first by contrastive nasalization in the process of nasal vowel evolution. This issue will be discussed again in the following chapter, which considers the temporal extent of vowel nasalization.

Chapter III

Vowel duration, vowel nasalization and nasal duration

Abstract

The relation between the temporal extent of vowel nasalization and vowel duration is tested for tense and lax pre-nasal vowels in German CVNV and CVNCV sequences. The goal is to investigate whether a clear relationship is apparent between vowel duration and anticipatory nasalization and if so, whether vowels of different tensity and height are affected to the same extent. Prior research provides evidence for vowel length as one factor in the development of contrastive vowel nasalization, especially in perceptual terms. In this sense, the vowel length parameter suggests long vowels being affected first before short vowels by contrastive nasalization. Our research, however, produced rather mixed results with respect to the relation between vowel length and nasalization: in absolute terms, tense vowels tend to exhibit a higher extent of vowel nasalization than lax vowels, while proportionally to the respective vowel length, nasalization differences are less evident.

Furthermore, the question is tested whether the extent of the overall lowering gesture of the soft palate is impacted by variations in vowel tensity and the post-nasal context in CVNCV and CVNV sequences. The goal is to investigate whether the overall velum lowering gesture remains roughly constant but, depending on the specific context, is shifted across the sound segments. Prior research reported evidence for the idea that more extensively nasalized vowels are followed by shorter nasal consonants, while less nasalized vowels precede longer nasals. Our findings suggest a more variable extent of the overall lowering gesture: it is decreased when the nasal is followed by a post-nasal stop rather than by a vowel. In addition, considering tense versus lax vowels in CVNCV sequences, the overall lowering gesture is increased when tense vowels are involved.

Thus, this chapter consists of two parts that consider different parameters affected by vowel tensity and the nature of the post-vocalic context: chapter 3.1 investigates the effect of tensity and the context on the temporal extent of vowel nasalization. Chapter 3.2 focuses on the overall temporal extent of the velum lowering gesture, i.e. the time span from velum lowering to raising in the target sound sequences under investigation.

3.1 Vowel duration and vowel nasalization

3.1.1 Introduction

As elaborated in chapter II, low vowels tend to be produced with a lower velum than high vowels, which may be one reason for low vowels to be preferentially affected by contrastive vowel nasalization. However, it has been suggested that the evolution of nasal vowels does not solely depend on the tongue position and the accompanying velum position but rather is related to the duration of the vowel. According to this scenario, low vowels are prone to becoming distinctly nasalized not because they are low but because they often are intrinsically longer than mid and high vowels (Hajek and Maeda, 2000). Thus, on the one hand, a long vowel potentially allows for longer nasalization during articulation and on the other hand, and perhaps more importantly, it is prone to being perceived as more nasalized than a short vowel (Hajek and Watson, 1998; Whalen and Beddor, 1989). In fact, duration differences between low and high vowels constitute a widespread phenomenon across languages (Busà, 2003; Hajek and Maeda, 2000; Laver, 1994; Toivonen et al., 2015).

Evidence for vowel length as a possible factor for establishing contrastive vowel nasalization comes from a number of studies of Hajek (1992, 1997) and Hajek and Maeda (2000), who investigated this phenomenon in Northern Italian dialects. Their data indicate that in these dialects distinctive vowel nasalization in stressed syllables is widely used in the context of historically long vowels, regardless of the vowel height, while short nasal vowels are not evident. Moreover, lengthening of short vowels mainly affects low vowels and only in some cases also mid vowels, as illustrated by the examples below:

Nasal vowel evolution in Italian dialects (from Hajek and Maeda, 2000, p. 11)

Latin	/pa:ne/ 'bread'		/annu/ 'year'
Proto-Northern Italian (PNI)	*/pa:n/		*/an/
Tavetsch	[pawn]		[ɔn]
Milan	[pɑ̃ː]		[an]
Cairo	[pãŋ]	(< [pãᶭ̃])	[an]
Bergamo	[pa(:)]	(< [pãː])	[a(:)n]
Bologna	[pæ̃ŋ]	(< [pæ̃ᶭ̃])	[a:n]
Rimini 1918	[pɛːn]/[pɛ̃ːn]		[a:n]
Rimini 1991	[pɛ̃ːn]		[ã:n]
Lugo	[pɛ̃ː]		[a:n] (no nasalization)
Imola	[pɛ̃ː]		[ɛ̃ːn]
Ravenna	[pʌ̃ː]		[ʌ̃ːn]

These findings, among others, lead Hajek and Maeda (2000) to the assumption that "the spread of nasalization phenomena is entirely dependent on a gradual process of height conditioned vowel lengthening." (Hajek and Maeda, 2000, p. 11). In fact, it seems that if vowel nasalization becomes contrastive in a language, it always affects long vowels first before short vowels. This relation is formulated by the *Vowel Length Parameter* (VLP):

Vowel Length Parameter (Hajek and Maeda, 2000, p. 10)

V:N >> VN

The VLP states that no language has been attested so far in which short nasal vowels have become contrastive without the long vowels having been affected first. Consequently, several languages have long nasal vowels or both long and short nasal vowels, but, according to the authors, no language is reported which solely exhibits short nasal vowels. If vowel length has some impact on the evolution of contrastive vowel nasalization, and if the initiation is grounded in phonetic principles, differences in articulation and perception between long and short vowels are expected to occur, including articulatory variation in the temporal and spatial extent of velum lowering as well as perceptual differences of vowels with the same amount of nasalization but varying lengths.

Articulatory evidence With respect to the velum lowering timing patterns in short versus long vowels, three scenarios are conceivable. First, the lowering gesture may start at a roughly fixed point before the nasal, resulting in a temporally constant nasalized part of the preceding vowel independent of its length. In this case, the relation of the oral to nasalized portion would depend on the vowel length, with the shorter vowel being nasalized to a larger proportion. In the second scenario, the anticipatory lowering gesture may start at a fixed point after vowel onset, leading to a greater proportional and absolute amount of nasalization in the long vowel. A third option includes a similar proportion of nasalization in long and short vowels, leading to a more temporally extended nasalization of the long vowel in absolute terms, but not relative to vowel length compared to short vowels.

The precise velum lowering patterns during pre-nasal vowels of different lengths have been considered in a handful of studies. Considering the timing of velar movements during pre-nasal vowels in particular, Moll and Daniloff (1971) used cineflourographic films to investigate coarticulatory velar behavior for three American English speakers and one speaker of Canadian English. The authors tested both anticipatory and carryover coarticulation of velopharyngeal movements in various sequences of vowels, nasals and oral consonants combined within and across word boundaries in natural sentences. Their major finding was that anticipatory velum lowering in VN sequences was initiated very early in the vowel, near the beginning of the primary articulatory gesture. Although no concrete distinction was made between lax and tense vowels, a crucial finding was that this early velum lowering gesture also extended to two vowels preceding a nasal, starting at the beginning of the first vowel. This pattern was found both within (non-)words (as in ‘freon’) and even across word boundaries (as in ‘free Ontario’), indicating that anticipatory nasalization started as soon as possible after the pre-nasal vowel onset.

The temporal extent of vowel nasalization in production was also considered by Clumeck (1976), who examined the nasal portion within vowels of different heights in pre-nasal position. In general, the timing of velum lowering turned out to be highly language-specific across the six languages tested, with a late onset of the lowering gesture in Amoy but nearly full nasalization in American English. Clumeck’s data also revealed that for five of the six languages under examination, the low vowel /a/ tended to be longer in overall duration than the mid and high vowels (although the difference was not always significant), and for four languages it also showed the greatest percentage of nasalization. Moreover, in absolute terms, the longest vowels in all but one case also exhibited the longest duration of nasalization. With respect to contrastive vowel length, only Swedish was tested for vowel nasalization differences in long versus short vowels. Results revealed that the duration of nasalization in absolute terms was quite similar for both long and short vowels, indicating a constant timing of gesture lowering before the nasal consonant in this language.

Considering vowel nasalization relative to vowel length in two different languages, Solé (1992) tested the hypothesis that vowel nasalization was phonological for American English but not for Spanish. Accordingly, the extent of nasality was predicted to vary relative to the speaking rate in English (and hence to the varying length of the vowel), whereas for Spanish, no extent differences were expected, suggesting that anticipatory velum lowering started at a similar point due to phonetic-mechanical principles. American and Spanish participants were measured by means of the nasograph[1] (Ohala, 1971) while reading CVVC target words with five different speaking rates. Findings suggested multi-level stages of velum lowering during pre-nasal vowels for American English (which was in accordance with Bell-Berti and Krakow, 1991), but not for Spanish. Moreover, the Spanish data revealed that velum lowering started at a constant point before the nasal stop, independent of the speaking rate (i.e. independent of vowel duration), while for American English, the extent of vowel nasalization was found to be proportionally adjusted, which corresponded to full nasalization for all vowel lengths. Based on these findings, Solé suggested that vowel nasalization in Spanish was "the result of a physiological time constraint" (Solé, 1992, p. 38), but it was "present in the phonological specification of the segment" (Solé, 1992, p. 39) for American English.

[1] Cf. section 1.5 which provides an overview of measurement techniques for tracking velum movements.

For Italian, Busà (2003) investigated the duration and nasalization of vowels of different heights in Italian VNC sequences, with C containing stops, fricatives or trills. By means of nasal airflow measurements from two speakers, results revealed differences in vowel duration: while the low vowel /a/ was found to be longer than high /i/, the percentage of nasalization turned out to be nearly identical in both vowels.

Perceptual evidence While the articulatory data provide evidence that long vowels are often nasalized to a temporally greater extent than short vowels, the question arises whether this pattern is also reflected in perception, and whether length itself is sufficient to create the perceptual impression of nasalization. For example, Delattre and Monnot (1968) edited synthesized items of French *l'aide* [lɛdᵊ] 'help' and *l'Inde* [lɛ̃dᵊ] 'India', such that the spectral modifications of vowel nasality were intermediate between oral and nasal. Moreover, the stimuli exhibited vowels of nine different lengths. Stimuli were presented to nine French and ten American English listeners who had to opt for either *l'aide* or *l'Inde* after each token. Responses clearly revealed a correlation between vowel duration and perceived nasality: the longer the vowel, the more likely the word was rated as *l'Inde*.

In a series of experiments, Whalen and Beddor (1989) focused on Eastern Algonquian languages in which a distinctive nasal vowel had evolved from a long oral vowel without any nasal context. This change only affected the long low vowel /a:/ but not the other vowels of the vocalic system, which lead the authors to investigate whether a tendency was evident for long vowels to be preferentially perceived as nasalized on the one hand and produced with more extended velum lowering on the other hand. In a first experiment, a synthesized low vowel /a/ was rated more nasalized by American English listeners when velopharyngeal port (VP) opening was increased, as expected. In addition, ratings were also higher when the duration of the vowel was enlarged, even when no nasal coupling was involved. This pattern was also found for the synthesized high vowels /i/ and /u/, with /u/ showing only duration effects but no impact of VP opening. In contrast, subsequent experiments with natural speech tokens lacked the duration effect. Only when these tokens were edited to create iterated pitch periods of different lengths was duration found to have an effect on the nasality ratings, but solely on those stimuli with nasal vowel iterations.

According to the authors, the discrepancy between the participants' judgements of nasality of the synthetic and natural oral vowels might have been due to the fact that in naturally produced vowels, the oral low vowel /a/ was inherently produced with a slightly lowered velum. To be perceived as nasalized, this vowel required more VP opening than the other vowels and might be otherwise perceived as oral, despite some nasal coupling. In contrast, the synthetic oral vowels were perceived as more nasalized when duration was increased without any VP opening, which was supposed to be related to the composition of the formant bandwidths. From the perceptual results, the authors concluded that "longer vowels with some appropriate degree of nasalization sound more nasalized than shorter vowels of the same spectral shape." (Whalen and Beddor, 1989, p. 473). However, they doubted that the small intrinsic length differences between vowels of various heights were sufficient to contribute to the emergence of contrastive vowel nasalization. Instead, they suggested that this process was more likely for vowels of contrastive length with larger duration differences (Whalen and Beddor, 1989, p. 482).

However, the finding that nasalization was perceived by increasing the duration independent of the vowel height did not explain why it was often /a:/ being affected by contrastive nasalization first. Therefore, the authors assumed that the intrinsic velum position for specific vowel heights could not be ignored. They considered data from Henderson (1984) who showed that, at least in Hindi, the velum lowering differences between oral and nasal /a:/ were quite small and overall less than velum height differences for the other oral versus nasal vowels. This finding was in contrast to studies on perceived vowel nasalization in which only little VP coupling was required for high vowels to change the spectral pattern from oral to nasal, while for the low vowel, much more coupling was necessary (House and Stevens, 1956; Maeda, 1993). The authors suggested that this discrepancy might be resolved by considering the stimuli used for these experiments: while Hendersen provided data on natural speech, synthesized stimuli were utilized in the perception studies. During a natural low oral vowel, velum lowering was already present and, as Hendersen outlined, only a little bit more lowering was needed to create nasality. In contrast, a synthesized low vowel with a closed VP obviously needed a much greater lowering distance until it had the same VP size as the natural vowel.

Considering coarticulatory nasalization in production, the velum height in long versus short vowels was examined indirectly by further perception tests. Naturally produced vowels from Western Abenaki, surrounded either by nasal or oral contexts, were edited such that all stimuli had the same length but consisted of pitch period iterations from either long or short vowels in their specific contexts. By doing this, the authors tested whether long vowels showed a higher degree of velum lowering than short vowels in production. If so, the participating listeners were supposed to notice some nasality differences between the nasal stimuli. Results revealed no duration effect for the oral vowels but a slight effect for vowels in nasal contexts: stimuli with iterations of the long vowels were rated as more nasalized than stimuli with short vowel iterations. However, for the same experiment with Cherokee, no such correlation was found, for which reason the authors concluded that "there is not a strong universal tendency to introduce nasalization into phonemically long low vowels." (Whalen and Beddor, 1989, p. 478).

Comparable perception results were reported by Hajek and Watson (1998), who presented synthetic disyllabic stimuli to ten British listeners. The stimuli comprised [asa]-like sequences, in which the first vowel exhibited three different stages of VP opening, two duration levels and two stress conditions. The second vowel always omitted VP coupling. Listeners were asked to rate the heard items with respect to nasality. Similar to Whalen and Beddor, results revealed a direct relation between vowel duration and nasality rating. However, significant rating differences were found only between the closed VP level and the two VP opening stages, but not between the medium and large VP opening stages. The authors attributed this finding to the fact that listeners were confronted with disyllabic stimuli that involved two near-adjacent vowels with differing nasalization degrees. This might have lead listeners to rate the two open VP stages similarly because in comparison to the second vowel, both levels sounded equally nasalized.

The findings from the production and perception data on vowel length and vowel nasalization can be summarized as follows:

- the temporal extent of vowel nasalization is highly language-specific: vowel nasalization may start at a fixed point before the nasal stop independent of the vowel length or may variably extend to the whole vowel (Clumeck, 1976; Moll and Daniloff, 1971; Solé, 1992)
- longer vowels are often accompanied by longer vowel nasalization in absolute terms, but not necessarily in proportion to the vowel length (Clumeck, 1976; Busà, 2003)
- vowel duration is related to perceived vowel nasalization: the longer the vowel, the more likely it is perceived as nasalized (Delattre and Monnot, 1968; Whalen and Beddor, 1989; Hajek and Watson, 1998)

As these findings suggest strong variations of the nasalization patterns across the languages, our investigation of German productions contributes to further insights about the interplay of vowel length and nasality.

3.1.2 Experiment II: Vowel nasalization and vowel tensity

Predictions

Based on the results from past studies, differences in the duration of vowel nasalization are expected for different vowel lengths in German vowels. The perceptual findings indicate that listeners tend to perceive longer vowels as more nasalized. The articulatory data also suggest a tendency for more nasalization in longer vowels, although various scenarios are conceivable, as outlined in the beginning (and as demonstrated by Solé, 1992): a) the velum lowering gesture may start before the nasal at a roughly fixed time point in both tense and lax vowels, b) the lowering gesture may be initiated at a roughly fixed time point after the vowel onset or c) the lowering gesture is timed relatively to the vowel length such that tense and lax vowels are nasalized to a similar proportion. Since it is not clear which of the three scenarios is to be expected for the German data, the following more general prediction is formulated:

Hypothesis: Nasalization and tensity in German vowels

H1 German pre-nasal tense vowels in CVNV and CVNCV sequences exhibit a larger extent of temporal vowel nasalization than lax vowels.

Speech materials

The data analyzed for this specific question are based on a subset of the original corpus. Separate analyses of items with CVNV and CVNCV sequences are presented in which the first V either was tense or lax. The post-nasal C in the CVNCV sequences refers to the oral voiceless stop /t/. It has been frequently reported that vowels become preferentially nasalized and the following nasal stops become reduced when these are followed by an oral voiceless obstruent (Beddor, 2009; Busà, 2007; Ohala and Ohala, 1991; Sefton and Beddor, 2005), presumably due to the distinct articulatory and acoustical incompatibility of a nasal and an oral voiceless consonant (Beddor, 2009; Busà, 2007; Ohala and Ohala, 1991, 1993, see also chapter IV, section 4.1.2). Thus, to examine the effect of the post-nasal context on the temporal extent of vowel nasalization, /t/ was chosen to provide a context that is scarcely compatible with nasal stop production and which is in contrast to a post-nasal weak vowel that does not require complete velar closure. A total of 1,182 items from 35 words was analyzed: 725 CVNV items and 457 CVNCV items. All target words were embedded into a carrier phrase and read with broad prosodic focus, as described in section 1.6.1. The subset data allowed for two main investigations with respect to vowel nasalization: the first one focused on the impact of vowel length (here: vowel tensity) affecting vowel nasalization, which was examined separately for CVNV and CVNCV sequences. The respective target items for the analysis of this question are listed in tables A.4 and A.5 in the appendix. The second aspect considered whether the post-nasal context (i.e. a post-nasal vowel versus oral stop) affected the extent of vowel nasalization, for which the target words in tables A.6 and A.7 were used.

Besides differences in vowel tensity, some of the word pairs additionally show variations in their word onsets (e.g. *Diener* versus *Finne*). It has been previously shown, though, that the velum is raised with comparable closure of the velopharyngeal port for both stops and fricatives (Amelot and Rossato, 2006, 2007), such that the achieved intra-oral pressure ensures sufficient energy for the required air turbulence. Therefore, the consonantal variation preceding the vowel in this subset is not expected to affect the velum position during the vowel. Furthermore, German tense and lax vowels do not only differ in length but also systematically vary in quality: tense vowels are generally articulated more at the periphery of the vocal space, while the shorter vowels are produced more centrally (Hoole and Mooshammer, 2002). Thus, the statements in the following about the effect on vowel nasalization refer to tensity rather than exclusively to length, although length is considered an important factor.

Participants and procedure

Detailed information about the participants and measurement procedure are outlined in section 1.6. The MR images were processed with the method sketched in section 1.6.4. The subset analysis data refer to measurements of vowel duration and vowel nasalization, in which vowel duration was determined manually from the acoustic vowel onset and offset as described in section 1.6.4. Vowel nasalization was defined as the difference between the point of maximum velocity during velum opening, i.e. the point where the velum was fastest in its lowering gesture and the acoustic vowel offset.

As for experiment I, statistical analyses were carried out in the programming environment *RStudio* (version 1.2.5033) by applying linear mixed models with the *lmer* function from the *lmerTest* package. To test the hypothesis just outlined, vowel length and the duration of vowel nasalization were selected as the dependent variables, respectively. Vowel tensity and the vowel category were defined as the fixed effects and the speaker and word onset as the random effects.

Results: Vowel duration and vowel nasalization

CVNV: tense versus lax vowels

Figure 3.1 shows the mean differences of vowel duration and vowel nasalization between tense and lax vowels in CVNV sequences, scaled in seconds. For each vowel per participant, the duration means of lax vowels were subtracted from the duration means of tense vowels. Likewise, for each vowel and each participant, the mean differences of vowel nasalization were calculated for tense versus lax vowels, which allows for investigating whether a tendency is apparent across individual speakers' duration differences. With respect to vowel length, all tense vowels generally exhibit larger duration values than their lax counterparts, which particularly is the case for tense /a/, /o/ and /ø/. Similarly, vowel nasalization tends to be increased in tense vowels, although here, clear nasalization differences seem to be present only for /a/. Table 3.2 lists the mean values of vowel length and vowel nasalization for each vowel category as well as the percentage of the nasalized portion. Since the values were

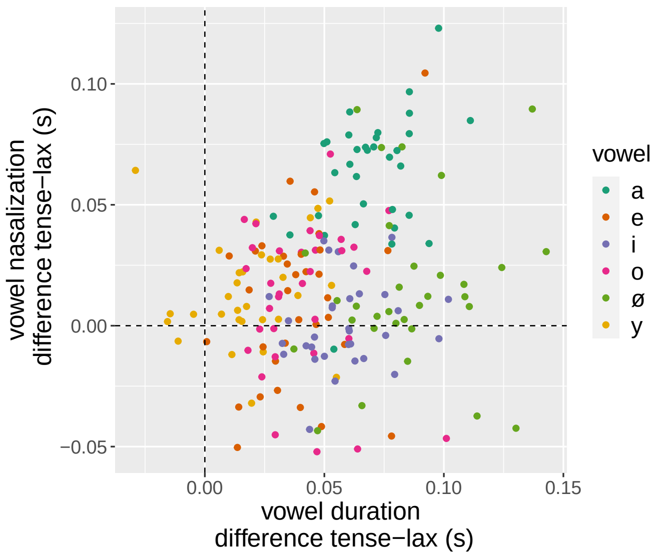


Figure 3.1: Differences of vowel length and duration of nasalization in tense–lax vowels in CVNV sequences per vowel per participant. X-axis: mean difference of vowel duration (s); y-axis: mean difference of vowel nasalization (s).

averaged across all speakers, speaker variability is not captured[2]. For the statistics below, however, speaker variability was considered as a random effect.

vowel	tensity	vowel duration (mean ms)	vowel duration (s.d. ms)	vowel nasalization (mean ms)	vowel nasalization (s.d. ms)	nasalized portion (in %)
/a/	tense	174.44	27.72	100.22	41.85	57,45
/a/	lax	104.68	24.66	35.33	25.71	33,75
/e/	tense	131.08	19.57	43.11	27.51	32,89
/e/	lax	92.77	14.37	34.06	22.05	36,71
/i/	tense	145.32	41.61	15.76	25.05	10,85
/i/	lax	88.38	39.27	14.42	21.94	16,31
/o/	tense	160.48	38.23	49.72	37.29	30,98
/o/	lax	121.01	34.95	37.90	25.77	31,32
/ø/	tense	191.42	31.15	34.26	28.74	17,89
/ø/	lax	106.23	13.86	17.65	19.08	16,61
/y/	tense	122.40	28.81	17.39	25.15	14,21
/y/	lax	102.14	12.83	5.47	19.78	5,36

Table 3.2: Means and standard deviations (s.d.) of vowel duration and vowel nasalization in pre-nasal vowels (CVNV).

The data in fig. 3.1 and table 3.2 suggest that on the one hand, tense and lax vowels differ in length as expected and on the other hand, nasalization in absolute terms appears to be overall slightly longer in tense vowels than in lax vowels. Indeed, statistical analysis showed that the length differences between tense and lax vowels varied significantly. Results suggested two main effects for the vowel category ($F[5, 38] = 40.67$, $p<0.001$) and tensity ($F[1, 13] = 61.18$, $p<0.001$) as well as a significant interaction between these effects ($F[5, 36] = 30.38$, $p<0.001$), which revealed that tense vowels generally had a significantly longer duration than lax vowels in these data except for /ø/. Regarding vowel nasalization, fig. 3.1 shows large differences within the respective vowel categories, suggesting that the nasalization difference between tense and lax /a/ was much greater than for the other vowel pairs. This impression was also confirmed by the statistical analysis with vowel nasalization as the dependent variable. Again, the vowel tensity and vowel category were reported as main effects (tensity: $F[1, 17]= 54.92$, $p<0.001$, vowel: $F[5, 23]= 23.27$, $p<0.001$). In addition, a significant interaction was indicated ($F[5, 41]= 31.54$, $p<0.001$). Post-hoc corrected bonferroni tests confirmed the tendency apparent in the data: tense /a/ was significantly more nasalized than lax /a/ ($p<0.001$). Similarly, tense versus lax /o/ exhibited a marginally significant difference in nasalization ($p<0.05$). Apart from that, no other pair showed notable differences in vowel nasalization.

[2]Note that the standard deviations in table 3.2 are less informative because variability across and within speakers is not captured here, but it is factored out by the linear mixed model applied below.

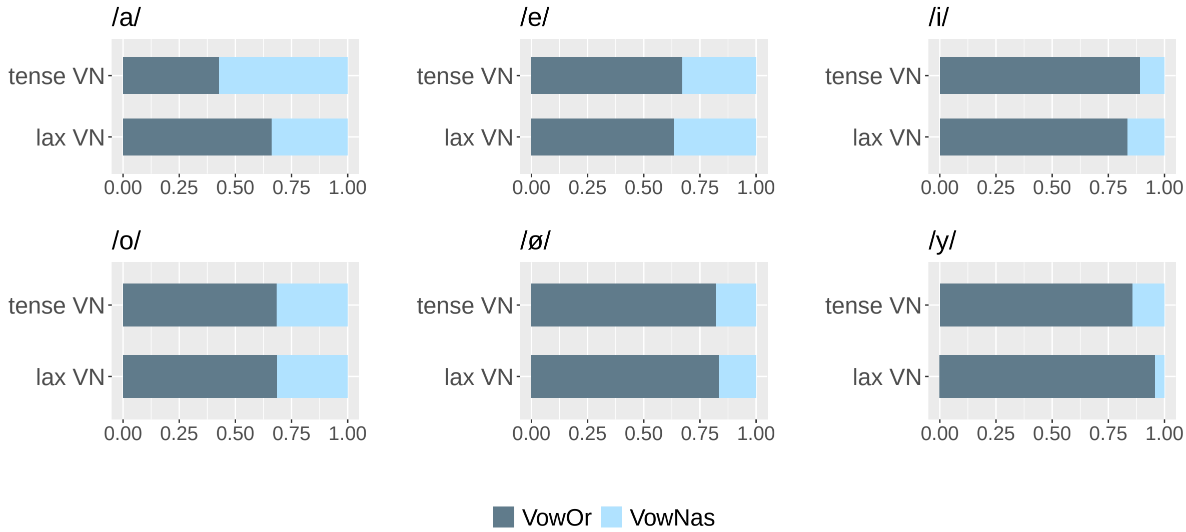


Figure 3.2: Relation of oral (VowOr) and nasalized (VowNas) portion in tense and lax vowels in CVNV sequences.

Considering the proportion of the oral and nasalized parts within the vowels (fig. 3.2), the only clear difference is noticeable for tense versus lax /a/ ($\approx$ 24%), whereas all other vowel pairs were nasalized to a similar extent within their own category (maximum difference of $\approx$ 9%). Evidently, this pattern argues against scenario b) outlined in the beginning: if the soft palate were lowered at a fixed point after the vowel onset, this would be reflected by the proportions, where the tense vowels should clearly exhibit a longer portion of nasalization due to their overall longer duration. Apart from /a/, however, this is not the case. Similarly, scenario a) seems unlikely because if velum lowering occurred at a fixed time point before the nasal onset, lax vowels should be nasalized to a higher percentage than tense vowels. Instead, the data are more compatible with scenario c), with /a/ as an exception: pre-nasal tense and lax vowels in CVNV contexts are nasalized to a similar proportion, suggesting that the point of maximum velocity during velum lowering is different for tense and lax vowels relative to the nasal onset.

CVNCV: tense versus lax vowels Similar to fig. 3.1, fig. 3.3 depicts differences between the duration and nasalization of tense and lax vowels for the CVNCV sequences under investigation. In this condition, values are largely positive in both dimensions, indicating that tense vowels in the CVNCV context appear to be generally longer and exhibit overall more nasalization than lax vowels. Accordingly, the averages of vowel duration and nasalization across participants show more discrete distances between the tense and lax counterparts of each vowel category (table 3.3). Equivalent to the CVNV context above, dependencies between vowel duration, tensity and the vowel category were tested for the CVNCV context, which revealed significant length differences between tense and lax vowels: tensity ($F[1, 59] = 255.2$, p<0.001) and the vowel category ($F[4, 10] = 17.43$, p<0.001) were rated as main effects with no interaction reported. Considering vowel nasalization, table 3.3 indicates that pre-nasal tense vowels of CVNCV sequences were more nasalized than lax vowels in absolute terms. In fact, with vowel nasalization as the dependent variable, statistical analysis suggested tensity ($F(1, 56) = 65.32$, $p<0.001$) and vowel category ($F(4, 16) = 12.92$, $p<0.001$) as main effects with no interaction, indicating that tense vowels exhibited more vowel nasalization. Proportionally to vowel length, however, larger differences were only apparent for /a/ and /i/, with relatively more nasalization in tense /a/ (≈16%) and lax /i/ (≈14%), respectively. Figure 3.4 illustrates the relation between the oral and nasalized parts of the respective vowels in CVNCV sequences. Since the oral and nasal proportions are comparable for tense and lax vowels (with slightly differing patterns for /a/ and /i/), scenario a) and b) proved to be unlikely for the CVNCV contexts as well: with a constant timing gesture after the vowel onset or before the nasal onset, tense and lax vowels would have systematically exhibited different proportions of nasalization. Obviously, this is not the case: except from /a/ and /i/, the vowels were nasalized to a similar proportion in both tensity categories, indicating that velum lowering occurred at different points of time relative to the nasal onset.

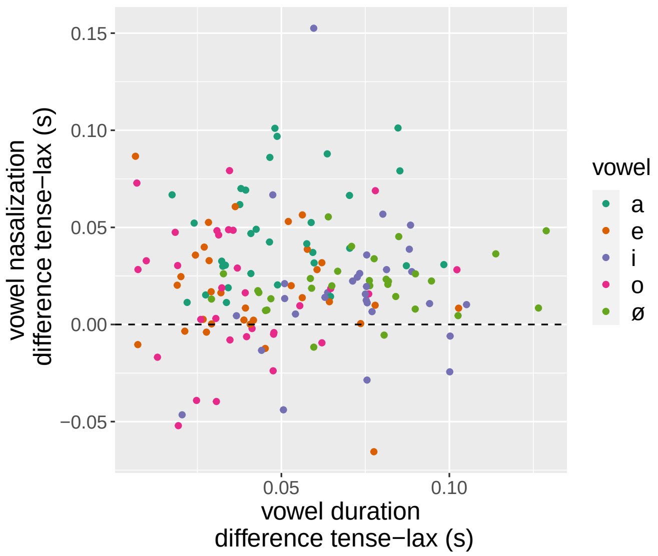


Figure 3.3: Differences of vowel length and duration of nasalization in tense–lax vowels in CVNCV sequences per vowel per participant. X-axis: mean differences of vowel duration (s); y-axis: mean differences of vowel nasalization (s).

vowel	tensity	vowel duration (mean ms)	vowel duration (s.d. ms)	vowel nasalization (mean ms)	vowel nasalization (s.d. ms)	nasalized portion (in %)
/a/	tense	154.45	23.47	97.92	30.98	63,39
/a/	lax	104.35	21.74	49.84	22.95	47,76
/e/	tense	127.14	25.59	62.72	30.40	49,33
/e/	lax	84.37	15.032	44.70	18.62	52,98
/i/	tense	124.24	18.58	49.09	36.23	39,51
/i/	lax	54.95	12.99	29.16	22.39	53,07
/o/	tense	132.14	23.73	72.57	31.15	54,92
/o/	lax	94.25	15.44	53.57	22.37	56,84
/ø/	tense	189.09	26.05	54.92	22.13	29,04
/ø/	lax	113.85	17.89	35.30	23.09	31,01

Table 3.3: Means and standard deviations (s.d.) of vowel duration and vowel nasalization in pre-nasal vowels (CVNCV).

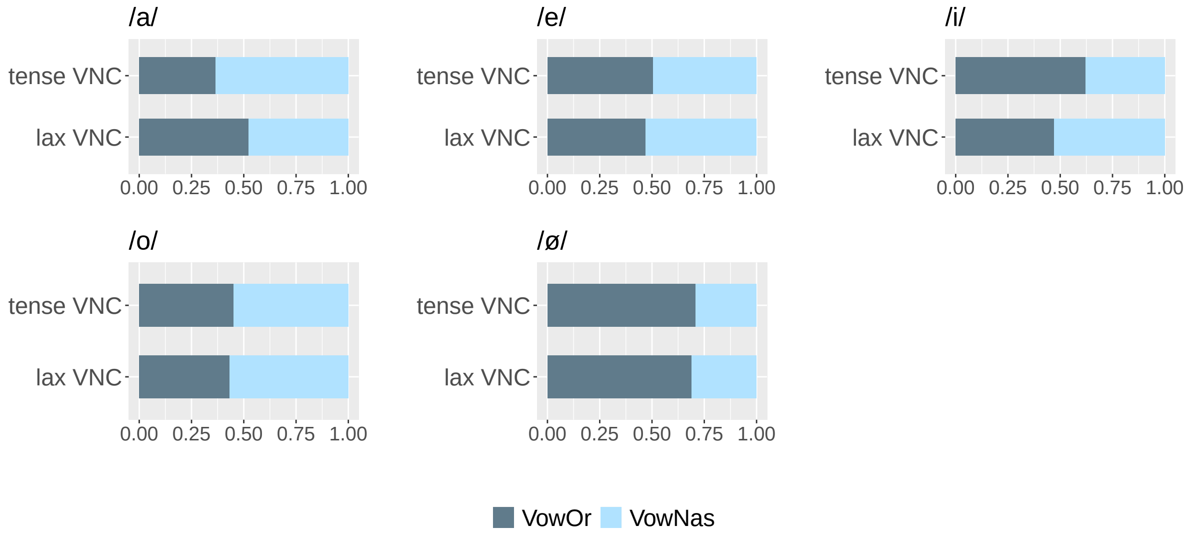


Figure 3.4: Relation of the oral and nasalized portion in tense and lax vowels in CVNCV sequences.

Interim summary

The tensity effect on vowel nasalization was found to vary across the two context conditions analyzed: while for vowels in CVNCV sequences, tensity had some general impact, this was only observed for /a/ and /o/ in the CVNV context. Considering the proportional relation of vowel nasalization to vowel length, findings are more compatible with the timing scenario c): tense and lax vowels showed a similar extent of nasalization relative to the vowel length, except for /i/ in CVNCV contexts and /a/ in both conditions. As with the data on the spatial amount of velum lowering, the temporal extent of nasalization was conspicuous for tense /a/, with more than half of the vowel being nasalized in both CVNV and CVNCV sequences.

Apart from the differences depending on vowel tensity, a comparison of the absolute and relative amounts of vowel nasalization between the two conditions leads to the impression that vowels in CVNCV contexts generally exhibit more vowel nasalization than in CVNV contexts. At the same time, vowel duration appears to be a little shorter in CVNCV sequences. Therefore, the following section explores the effect of the post-vocalic context on vowel nasalization and vowel duration.

Results: Contextual effects on vowel duration and vowel nasalization

To investigate the impact of the post-vocalic context on vowel duration and nasalization, the speech material subsets from tables A.4 and A.5 were combined to create groups of CVNV versus CVNCV items separated by tense and lax vowels. These are listed in tables A.6 and A.7 in the appendix.

Post-nasal context: tense vowels

Figure 3.5 shows differences for each participant per vowel for tense vowels only. The differences refer to mean values for vowel nasalization and vowel duration in CVNCV versus CVNV contexts, where values of CVNV were subtracted from CVNCV. A tendency is apparent for tense vowels showing a shorter duration when followed by NCV compared to NV. Vowel nasalization, in turn, appears to be slightly increased, although for /a/ no clear pattern is evident. The proportional relation between the oral and nasalized part of the vowels in their respective context is illustrated in fig. 3.6, with proportions referring to the means for the oral and nasal part of each vowel in the respective context, averaged across the participants. The impression given by fig. 3.5 is confirmed by fig. 3.6: vowels in CVNCV contexts appear to be shorter in duration but exhibit a higher proportion of vowel nasalization relative to vowels in CVNV contexts.

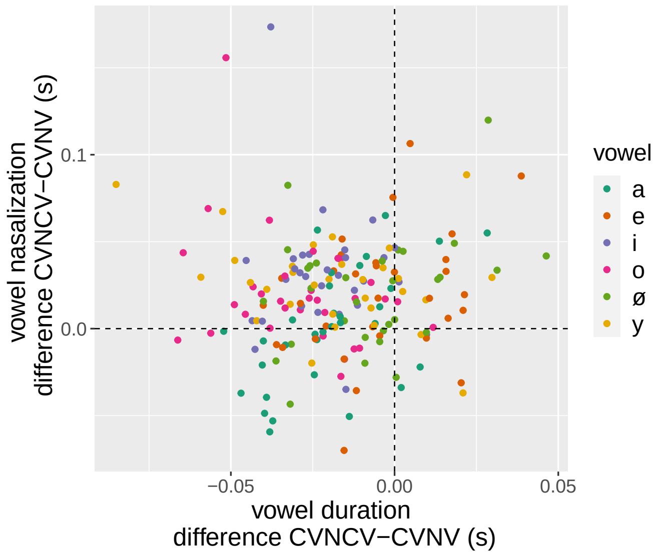


Figure 3.5: Differences of vowel length and duration of nasalization in tense vowels of CVNCV–CVNV contexts per vowel per participant. X-axis: mean difference of vowel duration (s); y-axis: difference of mean vowel nasalization (s).

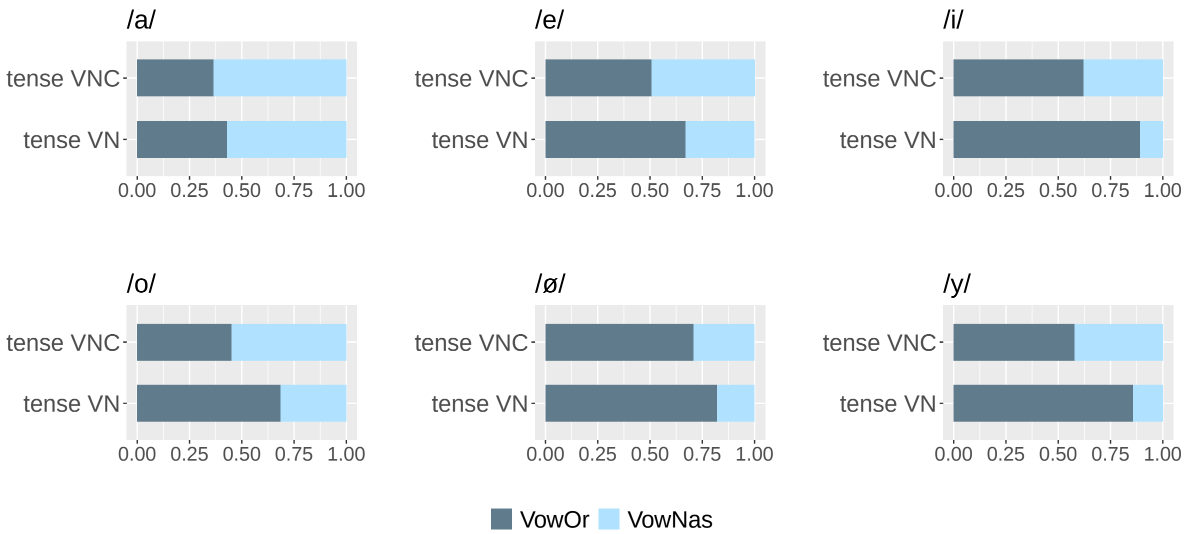


Figure 3.6: Oral and nasalized parts in tense vowels of CVNCV versus CVNV sequences.

For statistical analysis, two linear mixed models were applied, one with vowel duration and the other one with vowel nasalization as the dependent variable. For both models, the context and vowel category were defined as the fixed effects and the speaker and word onset as the random effects. Considering vowel duration, both the context and vowel category were rated as main effects ($F[1, 423]=6.39$, $p<0.05$ and $F[5, 15]=80.99$, $p<0.001$, respectively). In addition, an interaction was found between these factors ($F[5, 423]=6$, $p<0.001$). Post-hoc corrected bonferroni tests revealed length differences for /a/ ($p<0.001$) as well as for /i/ (interestingly with shorter /i/ in CVNV[3], though with only marginal significance ($p<0.05$)).

In terms of vowel nasalization, the context ($F[1, 427]=77$, $p<0.001$) and the vowel ($F[5, 16]=28.75$, $p<0.001$) were reported as main effects. An interaction was also found ($F[5, 414]=7.36$, $p<0.001$), which revealed that all vowels except for /a/ showed increased vowel nasalization when followed by NCV rather than NV, although with varying significance levels (/e/: $p<0.05$, /ø/: $p<0.01$, /i, o, y/: $p<0.001$). Thus, given an equal vowel category and tensity, the post-vocalic context appears to have a stronger impact on the nasalization of a vowel than on its duration.

Post-nasal context: lax vowels The effect of the post-nasal context on the extent of nasalization and the duration of lax vowels is illustrated in fig. 3.7. At first glance, no clear tendency is apparent for a duration difference of lax vowels in CVNV versus CVNCV sequences, with different vowels showing varying patterns: for /a/ or /e/, hardly any differences are evident, while /i/ and /o/ seem to be shorter when followed by NCV. In contrast, vowel nasalization seems to be generally more extensive in CVNCV contexts. Considering the proportional relation between vowel duration and vowel nasalization, fig. 3.8 suggests a clear tendency for lax vowels in CVNCV contexts to be more nasalized on average than in CVNV sequences, as was the case for tense vowels. Given that vowel duration is roughly comparable in both contexts, fig. 3.8 indicates that the point of maximum velocity during velum opening occurs earlier when the nasal is followed by an obstruent rather than by a vowel. In terms of statistical analysis, results revealed that the vowel category had a marginal impact on vowel duration ($F[4, 3] = 7.76$, $p=0.04669$], while

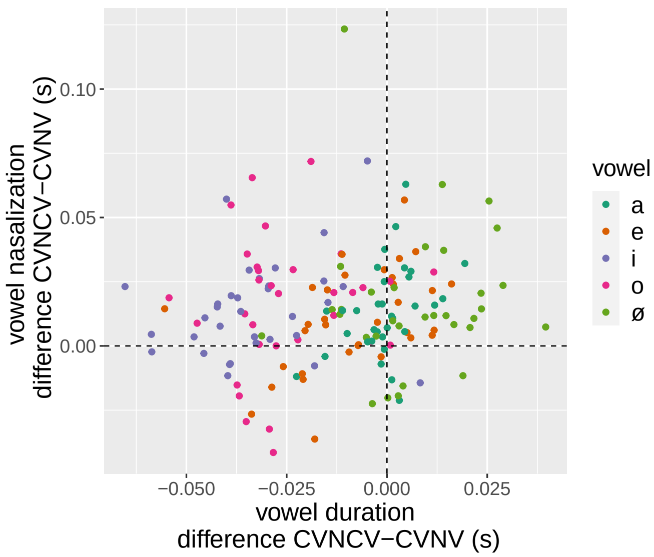


Figure 3.7: Differences of vowel length and duration of nasalization in lax vowels of CVNCV–CVNV contexts per vowel per participant. X-axis: mean differences of vowel duration (s); y-axis: mean differences of vowel nasalization (s).

[3]This might be surprising, as fig. 3.5 and tables 3.2 and 3.3 do not indicate a decreased vowel duration of tense /i/ in CVNV. However, closer inspection of the specific target words revealed inconsistencies with respect to the mean duration of tense /i/ in the CVNV words. As the mixed model applied here explicitly included the word onset as a random factor, this might have caused the reported outcome which is different from the impression given by the data above.

the context condition was not deemed an influencing effect. Focusing on the nasalization patterns, in fact both the context ($F[1, 455]$ = p<0.001) and the vowel category ($F[4, 8]$ = 6.57, p<0.05) were rated as significant main effects, with no interaction reported.

In summary, lax vowels were generally more nasalized when followed by NCV rather than NV contexts, while vowel duration was not affected in a significant way.

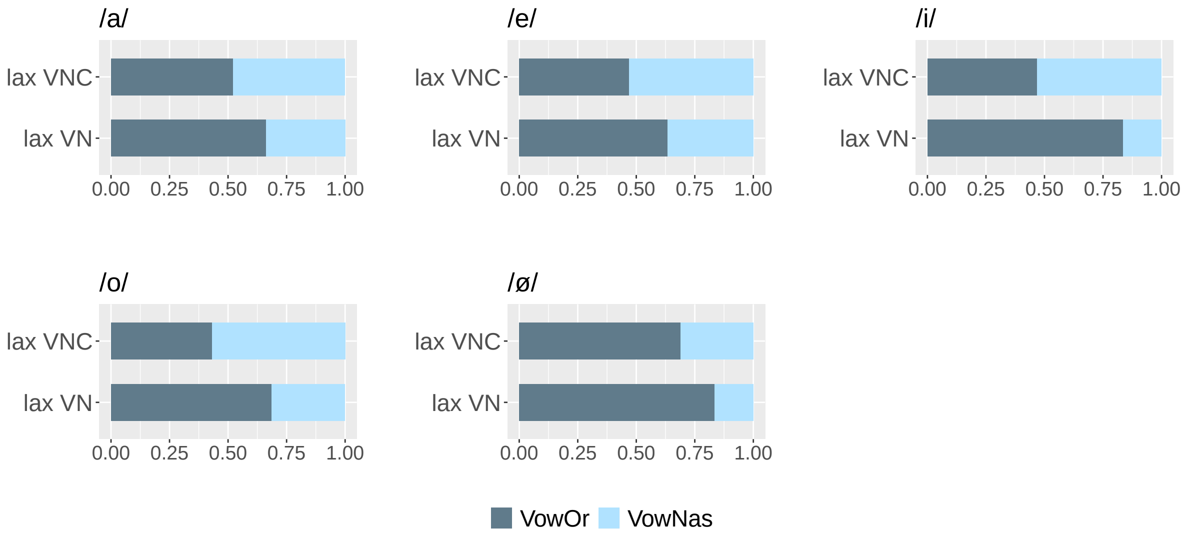


Figure 3.8: Oral and nasalized parts in lax vowels of CVNCV versus CVNV sequences.

Interim discussion

The data on vowel tensity and vowel nasalization primarily revealed two findings. First, in absolute terms, tense vowels were nasalized to a larger extent than lax vowels, i.e. the velum lowering gesture started sooner during tense vowels relative to the following nasal consonant. While this pattern was generally observed for tense versus lax vowels in CVNCV contexts, only /a/ and /o/ showed significant differences in the CVNV context. That tense vowels tend to be more nasalized in absolute terms might be expected, since longer vowels may allow for a longer time span to elicit gestural adjustments. Relative to the vowel length, however, the inverse case could rather be expected, because a shorter vowel is 'closer to' the following nasal, suggesting that velum lowering might occur quickly after the vowel onset to provide sufficient velum lowering for the following nasal stop. In this scenario, shorter vowels should be nasalized to a higher percentage than tense vowels. As it turned out, however, the proportion of nasalization in tense and lax vowels in our data was quite similar, which is in agreement with scenario c). A clear exception to this was the low vowel /a/: from all vowels tested, tense /a/ was nasalized to the largest extent, which is in line with the finding that tense /a/ also exhibited the lowest velum position (at least in CVNV contexts, which served as basis for the velum height survey in chapter II). If velum lowering starts decidedly earlier than the vowel midpoint, the soft palate is expected to exhibit a lower position at the midpoint, which was indeed found for tense /a/. Otherwise, if the point of maximum velocity during velum lowering occurs later than the vowel midpoint, the lowest velum position will accordingly also be achieved after the vowel midpoint. In fact, as can be seen from table 3.2, vowel nasalization in CVNV sequences generally started after the vowel midpoint, with the exception of tense /a/. This might explain the similar degrees of velum height for the other vowels explored in chapter II, where some general velum lowering was present in the nasal context but probably not to the degree it occurred later in the vowel.

Second, the post-vocalic context affected the extent of vowel nasalization in the preceding vowel: both tense and lax vowels were nasalized to a larger extent when followed by a NCV sequence rather than NV, with the exception of tense /a/. In contrast, within the same tensity and vowel category, the duration of the vowel was hardly affected by the context: significant differences were only found for tense /a/ and /i/ but not for the other tense or lax CVNV versus CVNCV pairings. If vowel duration is not systematically influenced by the post-nasal context, but vowel nasalization is, this suggests that the context is a strong factor affecting the extent of vowel nasalization.

Considering the two conditions of CVNV and CVNCV in more detail, one might argue that these pairings (as well as tense versus lax CVNV) differ in syllable affiliation, such that the following nasal either belongs to the second syllable (after tense vowels as in /za: .nə/) or appears as a coda segment if followed by an oral stop plus vowel, as in /za:n .tə/, and that the varying nasalization patterns emerge from the different syllable affiliation. Indeed, prior research has indicated that nasal stops show systematically varying temporal and spatial velum lowering patterns in production depending on their position within the word: syllable-final nasal stops were found to exhibit a lower velum and longer low plateaus than nasals in initial syllable position (Krakow, 1989, 1993, 1999; Schourup, 1973). It seems likely that a longer lower velum position is accompanied by a higher assimilatory effect on the preceding vowel, for which evidence has been provided (Krakow, 1993, 1999). This aligns well with the data presented by Clumeck (1976) who reported that in Portuguese, the word /famba/ revealed slightly more nasalization on the first vowel than it was observed for the word /fama/. However, it is striking that in our data nasalization differences in CVNCV versus CVNV contexts were also apparent even in lax vowels. According to the syllable prediction, in these pairings little or no differences should occur, as in both contexts the lax vowel is followed by a tautosyllabic nasal (or ‘ambisyllabic’ in CVNV). However, in our data, the vowel was systematically more nasalized when the following context involved an oral stop. These ‘ambisyllabic’ cases were also briefly discussed by Krakow (1999), who suggested that such medial segments are affiliated with those syllables that carry primary stress (see also chapter IV). Further evidence for the role of stress comes from Byrd et al. (2009) who compared the timing patterns of the tongue tip and velum gesture in nasal stops in different syllable positions but also in words with similar syllable structure but different stress patterns (*beknow* vs. *bono* vs. *bonafide*). Along the line of Krakow, when the pre-nasal syllable had primary stress, the timing patterns resembled those observed from the coda position (velum lowering before tongue tip raising for /n/), while in words with primary stress on the post-nasal syllable the inter-gestural timing was comparable to that observed for the onset position (tongue tip raising for /n/ slightly preceded velum lowering). Thus, the syllable position alone cannot be considered sufficient for explaining the temporal patterns of velum lowering that occur during speech.

Furthermore, a CVNV context involving a final 'weak' unstressed schwa vowel may not be perfectly comparable to sound sequences with longer or secondary stressed post-nasal vowels: the schwa vowel alone may not elicit any relevant gestural changes of the soft palate, such that velum lowering in CVNV (with the second vowel being /ə/) might behave as if there were no final vowel at all. As an alternative to the syllable-based explanation, the varying nasalization patterns can be better explained by a pure phonetic approach, in which the surrounding context has some systematic impact on the vowel based on physiological and acoustic factors (Ohala and Ohala, 1991; Ohala, 1993; Ohala and Ohala, 1993; Steriade, 2000). In our data, the nasal consonant preceded an oral stop in one context and an unstressed schwa vowel in the other context. Hence, the difference in vowel nasalization between CVNV and CVNCV can be accounted for by the articulatory interplay during the production of these sequences: a post-nasal oral stop requires a closed velopharyngeal port to facilitate sufficient intra-oral pressure. The change from an open to a closed stage may occur within a larger temporal space in CVNV but needs to be quickly performed in CVNCV sequences. This may lead to a greater extent of anticipatory nasalization to ensure a sound that is sufficient in nasality. Moreover, it is not unlikely that due to the aerodynamic conflicts the nasal is additionally weakened (i.e. shortened or at least acoustically less salient) to some extent before the oral stop. In contrast, when the nasal is followed by a weak vowel, there is no need to provide for full velar closure, which generally allows for a later lowering gesture in anticipation of the nasal consonant in a CVNV sequence.

The finding that the extent of vowel nasalization is related to the post-vocalic context leads to the question of whether the extent of nasal duration - or more generally, the overall extent of velum lowering - is also affected in these sequences. This issue will be explored in the next section.

3.2 Overall temporal extent of velum lowering

3.2.1 Introduction

As suggested by the evidence from prior research and the results in section 3.1, nasalization in production is observed for long vowels to a greater extent than for shorter vowels (although the nasalized portion may comprise a similar proportion relative to vowel length). At the same time, perceptual tests have shown that vowels are perceived as more nasalized when length is increased (Delattre and Monnot, 1968; Hajek and Watson, 1998; Whalen and Beddor, 1989). However, in the interim summary of the previous section it was also indicated that the post-vocalic context had an effect, such that the extent of vowel nasalization was affected by the issue of whether the nasal was followed by an oral stop or a vowel. In fact, similar observations in prior research considering the relation between vowel nasalization and the following context have suggested that vowel nasality and nasal consonant duration show an inverse temporal relationship: a vowel exhibiting increased nasalization is often followed by a shorter nasal consonant and vice versa, longer nasals are preceded by vowels that are less nasalized. Much research on this has been carried out by Beddor and colleagues (Beddor, 2007, 2009, 2015; Beddor et al., 2018; Onsuwan, 2005; Sefton and Beddor, 2005). According to Beddor's approach, the temporal overall nasality across the vowel and the nasal segment is suggested to be roughly constant, independent of its specific assignment to the segments. Thus, a trade-off relationship between vowel nasalization and nasal duration is postulated: if the overall velum lowering gesture is shifted more into the vowel, the vowel becomes extensively nasalized, while at the same time the nasal is temporally reduced. In the extreme form, nasality is associated solely with the vowel and the nasal consonant becomes lost. Consequently, if this pattern becomes systematic and finally phonologized in a language, the emergence of contrastive vowel nasalization is a possible scenario.

In addition, the post-nasal context is assumed to affect the shifting of the velum lowering interval, which is preferentially elicited if the nasal is followed by a voiceless rather than voiced obstruent (Beddor, 2009; Busà, 2007; Ohala and Ohala, 1991; Sefton and Beddor, 2005). This effect is attributed to the temporal alignment of the velar and oral constriction gestures: in the voiceless condition, a lowered velum during the nasal is incompatible with the high air pressure required for a voiceless stop, such that an early onset of velum lowering in the vowel may help resolve this conflict (Beddor, 2009; Ohala and Ohala, 1991, 1993). The following section summarizes findings from studies providing articulatory and perceptual evidence for a relationship between vowel duration, vowel nasalization and nasal duration with and without considering the post-nasal consonantal context. Subsequently, the German data are tested for differences in duration of the overall velum lowering gesture within and across CVNCV and CVNV contexts.

Articulatory and perceptual evidence Most research on this topic has been done on American English, but occasionally other languages were investigated with respect to the timing and temporal extension of vowel nasalization as well. Sefton and Beddor (2005) analyzed CVN and NVC words produced by five speakers of American English and found the vowel in CVN more nasalized and the nasal more shortened compared to NVC. Although the distribution of nasality varied for the two sequences, overall nasality was similar for both contexts. This finding was confirmed by a subsequent experiment with CVNC words, with the final consonant varying in voicing: nasal duration was reduced before voiceless obstruents compared to voiced obstruents and vowel nasalization inversely co-varied with nasal duration. This inverse relation was also reported for Thai (Onsuwan, 2005).

In fact, the role of a post-nasal obstruent for the articulatory gestures during production had been described earlier by Clumeck (1976), who reported that both vowel duration and vowel nasalization in Portuguese tended to be more extended if the nasal was followed by an oral stop. The difference, though, was not rated significant. Similarly, Kawasaki (1986) mentioned that in Tunica, a nasal followed by a voiceless stop became devoiced while the pre-nasal vowel was decidedly more nasalized than in other pre-nasal positions. Malécot (1960) provided data from a perception task in which American English listeners perceived a nasal stop in words that contained a nasalized vowel followed by a voiceless stop.

Focusing on the trading relation hypothesis, Beddor (2007) presented data from American English, Thai and Botswanan Ikalanga. In addition to testing the trade-off between vowel nasalization and nasal duration, she also considered the relation between nasal duration and the temporal extent of oral stop constriction. Acoustic analyses for the American English data indeed showed a clear inverse relation both for the temporal nasalization extent across the vowel and nasal and for the constriction of the nasal and oral stop: the shorter the nasal, the more extended was the oral stop. In contrast, the Ikalanga data showed no influence of the post-nasal context on nasal duration. Moreover, the temporal extent of nasalization across the vowel and the nasal was not inverse in Ikalanga but roughly held constant in both voiced and voiceless contexts.

In addition, the effect of the vowel length on vowel nasalization and nasal duration was investigated. Following Onsuwan (2005), who had run the experiment for Thai speakers, Beddor (2007) provided comparable results for American English: long vowels (tense in English) were more nasalized and followed by a shorter nasal stop, while short vowels (lax in English) were less nasalized but followed by a longer nasal consonant.

To test the trading relation idea from a perceptional perspective, 23 English and 24 Ikalanga listeners were presented with stimuli pairs of three types: the first type involved a VN sequence with a constant portion of vowel nasalization and an altering length of the nasal consonant. The second type contained pair members with short overall nasalization in contrast to long overall nasalization and the third type comprised pair members with a similar overall nasalization that was differently distributed across the segments: an item with long vowel nasalization and short nasal duration was paired with an item of short vowel nasalization and long nasal duration. The prediction was that listeners would have the most difficulties in discriminating the pair members of type three, relying more on nasality itself rather than on its precise distribution. Results confirmed this prediction: listeners from both languages had most problems in identifying differences between pair members of type three, when overall nasalization was similar but differently distributed on the vowel and nasal segment. However, as expected due to less co-variation of nasalization in their own language, the Ikalanga listeners had overall better performances in discriminating the pair members of all types.

In her following work, Beddor (2009) amplified the trade-off approach in more detail. The production and perception experiments were repeated with new participants, all native speakers of American English. Data were presented for six participants who took part in the production study and for 27 listeners in the perception test. Again, the production results showed a clear negative correlation between the extent of vowel nasalization and the duration of the nasal stop in English CVNC words, where the final consonant was either voiced or voiceless. Likewise, the voicing condition of the post-nasal stop affected the length of the nasal, which showed a more extended duration when followed by a voiced rather than voiceless stop. However, voicing showed also some impact on the remaining gestures: when followed by a voiced stop, the preceding gestures for vowel length, nasalization across ṼN and alveolar constriction across NC were slightly longer compared to the voiceless condition. As suggested in her previous work, Beddor attributed this difference to the temporal alignment of the velum and oral constriction gestures, suggesting that an early onset of velum lowering may occur to ensure sufficient nasality before the achievement of a spread glottis and a closed velopharyngeal port for the production of the voiceless stop. In contrast, voicing is not in conflict with a nasal murmur until the velopharyngeal port is completely closed, which allows for a more extended nasal stop.

Similarly, the perception experiment with manipulated Ikalanga CVNCV words confirmed the results obtained earlier: listeners' performance was best at discriminating those pair members that differed in the extent of the overall temporal nasalization. The most difficulty occurred with similar-nasality pairs, that is, pair members with a similar overall nasalization that was distributed variably across the vowel and the nasal segment.

To sum up the findings on the contextual effect on vowel nasalization and nasal duration, the studies outlined suggest that

- vowels of VNC sequences are nasalized to a higher extent if the nasal stop is followed by a voiceless obstruent (Beddor, 2007, 2009; Kawasaki, 1986; Sefton and Beddor, 2005; Onsuwan, 2005)

- nasal stops are shortened if they are followed by a voiceless obstruent (Sefton and Beddor, 2005; Onsuwan, 2005)

- vowel nasalization and nasal duration are negatively correlated in some languages (Beddor, 2007, 2009; Onsuwan, 2005)

Further findings from subsequent perceptual studies on the trade-off hypothesis (Beddor et al., 2013; Beddor, 2015; Beddor et al., 2018) are delineated in chapter V. The results presented in the current section primarily focus on production data, exploring the temporal interaction of the articulatory gestures during German VN sequences.

3.2.2 Experiment III: Velum lowering extent in CVNCV versus CVNV sequences

Predictions

Most research on the trade-off hypothesis has involved data from speakers and listeners of American English, a language which has often been described as exhibiting extensive and partly even distinctive vowel nasalization (Chen et al., 2007; Clumeck, 1976; Malécot, 1960; Solé, 1992). Although the other languages tested in the mentioned studies showed similar results (Ikalanga and Thai), it would not be appropriate to simply transfer these findings to further languages. For example, as was outlined in chapter II, velum behaviour in German vowels is decidedly different from the lowering patterns in American English, especially with respect to oral vowels. Nonetheless, the second part of section 3.1.2 suggests that the post-nasal context affects the extent of vowel nasalization, which leads to the question whether nasal duration itself is also affected and whether there is a relation between nasal duration and vowel nasalization in the sense of Beddor's work. In the thesis at hand, however, this latter aspect will be explored by analyzing the overall extent of velum lowering instead of correlating the interval of vowel nasalization with the duration of the adjacent nasal consonant. As indicated in the introduction (section 1.6.4), this approach is motivated by the fact that in our data, these two parameters share one acoustic boundary: vowel nasalization was defined as the difference between the acoustic vowel offset and the point of maximum velocity during velum opening. Correspondingly, nasal duration would be defined as the interval between the acoustic vowel offset and the point of maximum velocity during velar closing. If these two adjacent intervals, sharing the acoustic vowel offset boundary, were to be correlated, this would necessarily result in a negative correlation. As has been pointed out in previous research (Ohala and Lyberg, 1976), the effect of such an erroneous negative correlation is related to inevitable measurement errors in manual analysis: if one interval is slightly increased, the adjacent interval is necessarily slightly decreased. Since the correlation is calculated based on single pairs that all include such small measurement errors, it will automatically be overall negative. To avoid this issue in our analysis, a different parameter was used to receive an impression of the relation between the extent of overall nasalization and the contextual influence of CVNV and CVNCV sequences. This parameter refers to the interval between the points of maximum velocity during velum opening and during velum closure and thus indicates the time span of clear velum lowering during the target word. Therefore, the question will be addressed how vowel tensity as well as the post-nasal context affect the overall velum lowering interval (instead of the nasal consonant duration) and how this can be interpreted in terms of the finding that vowel nasalization is increased in vowels preceding NCV rather than NV.

Considering the trading relation prediction in terms of the overall velum lowering extent, no significant differences should be evident for the lowering gesture in CVNV versus CVNCV words or in same-conditioned contexts with varying vowel tensity. The respective hypotheses are formulated as follows:

Hypotheses: Overall velum lowering extent in German CVNV and CVNCV words

H1 The time span of the overall velum lowering, defined as the interval between the points of maximum velocity during velum opening and closing, is not affected by the pre-nasal vowel tensity.

H2 The time span of the overall velum lowering is not affected by the post-nasal context.

Speech materials

The speech material for this analysis is the same used in experiment II (see appendix tables A.4, A.5, A.6, A.7), i.e. one data set with tense versus lax vowel comparisons separated by the post-vocalic context and the other one with context comparisons separated by vowel tensity. These subsets allow for exploring the data with respect to the predictions just outlined: to survey the impact of vowel tensity on the overall velum lowering extent (H1), CVNCV and CVNV contexts are separately explored. The influence of the post-nasal context (H2) is investigated by comparing the overall velum lowering in CVNV versus CVNCV sequences separated by the pre-nasal vowel tensity. As with the previous experiments, all items were embedded into a carrier phrase and read with broad prosodic focus, as described in section 1.6.1.

Participants and procedure

Detailed information about the participants, the MRI measurement procedure and image processing are given in section 1.6. For the experiment at hand, measurement values are given for vowel nasalization and the overall time span of clear velum lowering, which are defined as follows: vowel nasalization refers to the difference between the acoustic vowel offset and the point of maximum velocity during velum opening, i.e. the point where the velum was fastest in its lowering gesture (section 1.6.4). Overall velum lowering (henceforth OVL) was defined as the difference between the points of maximum velocity during velum opening and closing. Thus, the definition of OVL solely considers velum activity instead of acoustic parameters. For the present data, velum movement patterns were expected to differ from the acoustic correlation visible in a spectrogram: the point of closure was expected to highly depend on the post-nasal context[4]. This was based on the consideration that velum raising in the CVNV contexts was likely to occur later than would have been acoustically evident, since a post-nasal weak vowel generally does not require full closure of the velopharyngeal port; full closure is probably not required until the consonantal onset of a subsequent word. Accordingly, OVL in our CVNV data presumably extended to the post-nasal schwa vowel rather than to the offset of the nasal stop. This point will be also considered in the discussion below.

As for experiments I and II, statistical analyses were carried out in the programming environment *RStudio* (version 1.2.5033) by applying linear mixed models with the *lmer* function from the *lmerTest* package. To test the hypotheses outlined, a linear mixed model was run for each condition. Details on the model parameters are given in the respective sections.

Results

CVNV: tense versus lax vowels Figure 3.9 (left) shows the mean differences for OVL in CVNV contexts for tense versus lax vowels, indicating values per vowel per participant. Apparently, no overall tendency for OVL duration is evident for any specific direction. A linear mixed model was applied to test the effect of tensity on the overall velum lowering extent, with OVL as the dependent variable, tensity and vowel category as the fixed effects and speaker and word onset as the random effects. A main effect was found for the vowel category ($F[5, 43] = 11.36$, $p<0.001$). In addition, an interaction between the tensity and vowel category was reported ($F[5, 62] = 3.34$, $p<0.01$): post-hoc corrected bonferroni tests showed that OVL was significantly longer for tense /a/ compared to lax /a/, but not for the other vowels. Figure 3.9 (right) adds the mean difference values for vowel nasalization in tense versus lax vowels in CVNV contexts. As reported in section 3.1.2, significant differences were evident for /a/ and marginally for /o/, with longer nasalized portions in the tense vowels.

Figure 3.10 illustrates the portion of vowel nasalization relative to the overall velum lowering extent. On the left, means of OVL and vowel nasalization are given for each vowel, averaged across speakers. The figures on the right refer to the vowel nasalization proportion relative to OVL. The tendency found in fig. 3.9 is reflected in fig. 3.10: except for /a/, only small differences in OVL duration and vowel nasalization are evident for tense versus lax vowels, which holds both in absolute and relative terms.

[4]Note that this expectation does not refer to OVL in general but only to the point of closure, such that H2 is still to be tested.

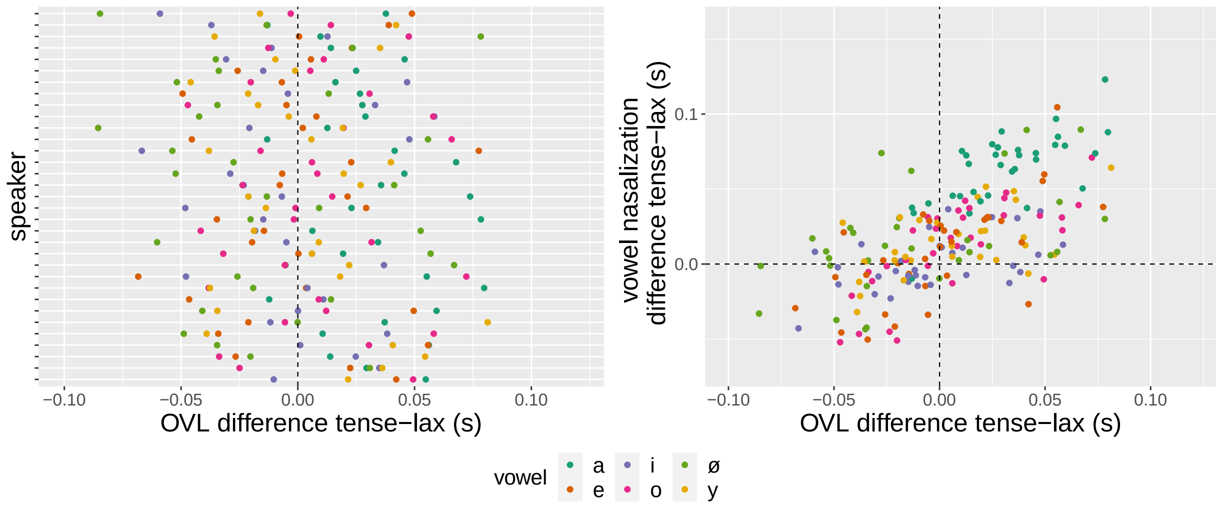


Figure 3.9: Left: differences of the overall extent of velum lowering (OVL) for tense–lax vowels (CVNV context). The single values refer to mean differences of OVL per participant per vowel. The ticks on the y-axis refer to speakers S03–S38 from bottom to top, omitting S11, S25 and S30 due to registration issues (see section 1.6.1). Right: vowel nasalization differences are added.

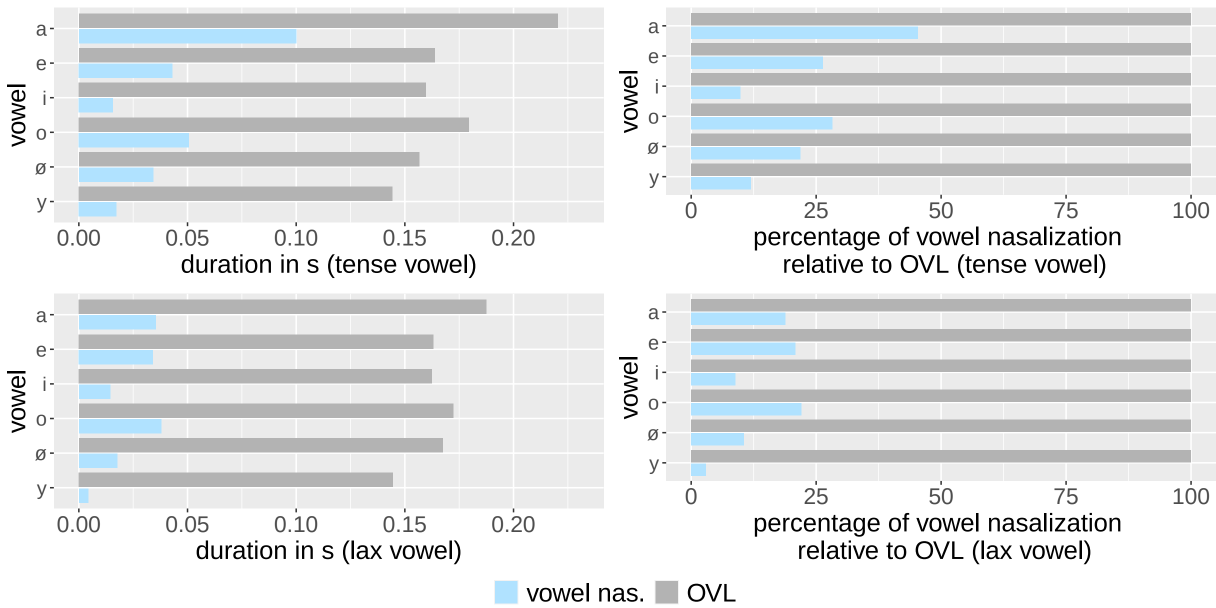


Figure 3.10: OVL duration and vowel nasalization in tense versus lax vowels (CVNV context). Plots are given for means across speakers for both absolute and proportional relations.

CVNCV: tense versus lax vowels Figure 3.11 (left) shows differences of OVL in CVNCV sequences with pre-nasal tense versus lax vowels. At first glance, a small tendency is apparent for longer OVL duration when tense vowels are involved, which is indicated by a bias towards the positive values along the x-axis. In fact, statistical testing revealed a main effect both for the vowel category and the tensity condition, with no interaction reported (tensity: $F[1, 67] = 9.73$, p<0.01; vowel: $F[4, 15] = 7.01$, p<0.01). Thus, OVL duration was found to be increased when tense vowels were involved.

By adding the mean differences for vowel nasalization to the plot, fig. 3.11 (right) shows

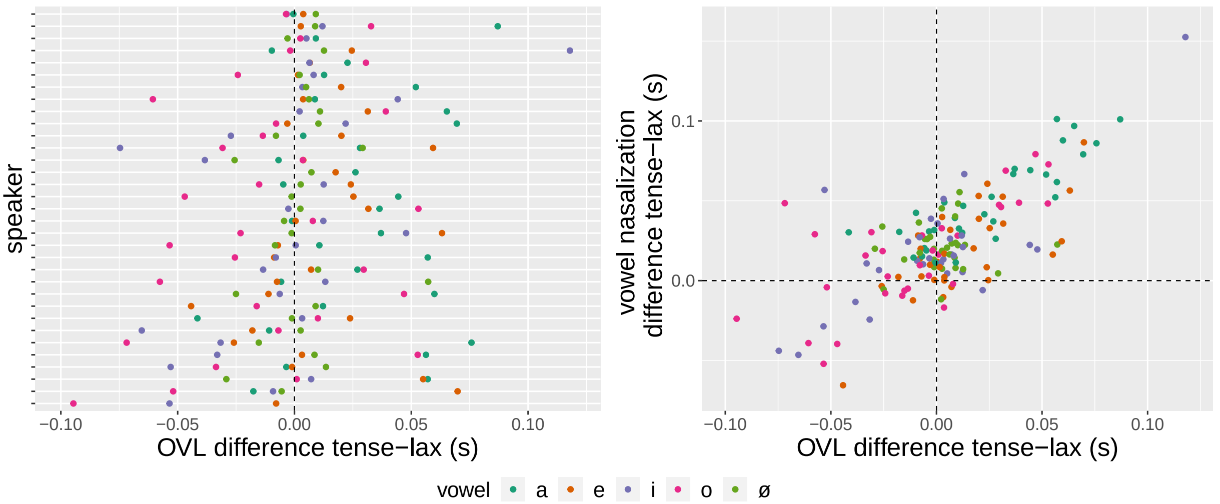


Figure 3.11: Left: differences of the overall extent of velum lowering (OVL) for tense–lax vowels (CVNCV context). The single values refer to mean differences of OVL per participant per vowel. The ticks on the y-axis refer to speakers S03–S38 from bottom to top, omitting S11, S25 and S30. Right: vowel nasalization differences are added.

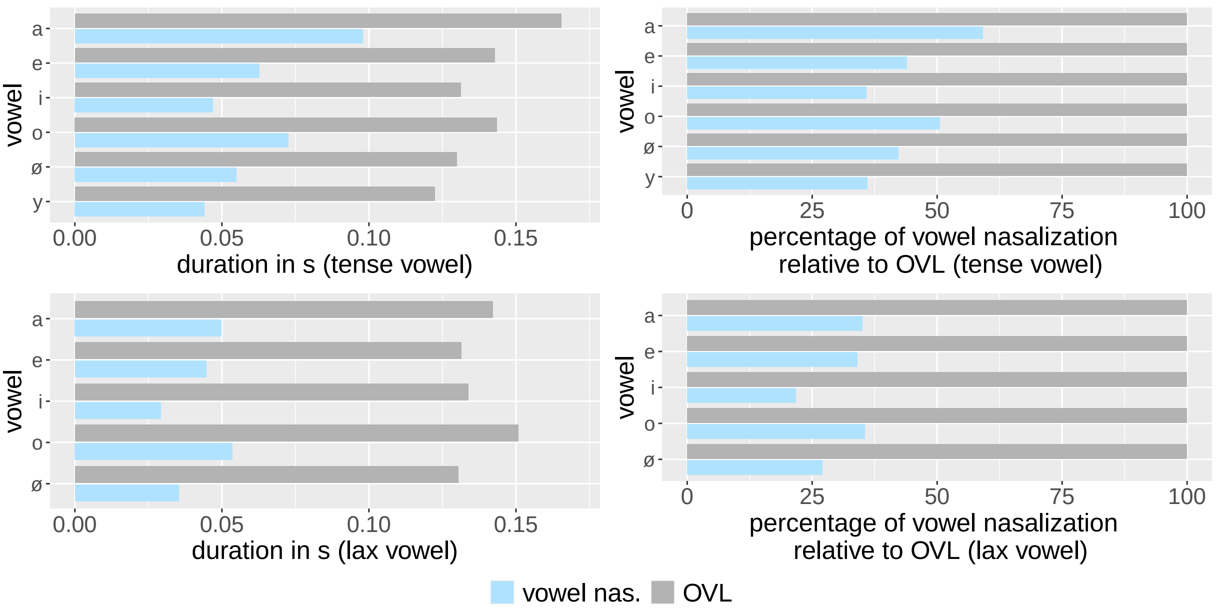


Figure 3.12: OVL duration and vowel nasalization in tense versus lax vowels (CVNCV context). Plots are given for means per OVL and vowel across speakers.

largely positive values along the y-axis, indicating more extensive vowel nasalization in tense vowels (as reported in section 3.1.2). That OVL duration was slightly longer when tense vowels were involved is only partly depicted by figure 3.12, which shows mean values for OVL and vowel nasalization in CVNCV contexts for tense and lax vowels, respectively. However, the left side suggests that vowel nasalization seems to cover a greater portion of OVL in tense than in lax vowels. This is also reflected by the percentage plots on the right: within the interval of the overall lowering gesture, vowel nasalization was proportionally more extended when pre-nasal tense vowels were involved.

Post-nasal context: tense vowels Figure 3.13 (left) gives an impression of the impact of the post-nasal context on the velum lowering extent. The data include sequences with tense vowels only. Mean differences are depicted for OVL between CVNCV–CVNV contexts for each vowel per participant. A clear tendency is apparent for OVL to be shorter if the nasal is followed by an oral stop rather than a vowel. Statistical results are compatible with this impression (dependent variable: OVL; fixed effects: vowel category, context; random effects: speaker, word onset): two main effects were found for the vowel category ($F[5, 28] = 18.18$, $p<0.001$) and the context ($F[1, 40] = 51.27$, $p<0.001$) as well as a significant interaction ($F[5, 383] = 2.39$, $p>0.05$). Post-hoc corrected bonferroni tests showed overall significant OVL differences between the two contexts for all vowels except for /y/ (/a, i/: $p<0.001$, /o, ø/: $p<0.01$, /e/: $p<0.05$, /y/: $p=0.087$). Overall, results suggested an effect of the post-nasal context: OVL was decreased if the nasal was followed by an oral stop compared to a schwa vowel.

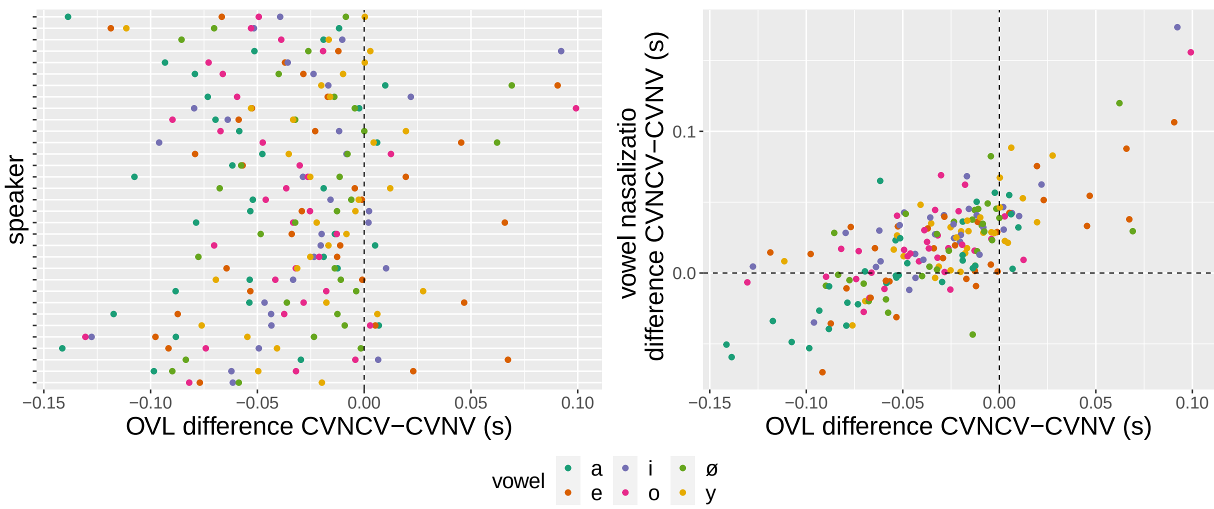


Figure 3.13: Left: differences of the overall extent of velum lowering (OVL) for CVNCV–CVNV contexts (tense vowels). The single values refer to mean differences of OVL per participant per vowel. The ticks on the y-axis refer to speakers S03–S38 from bottom to top, omitting S11, S25 and S30. Right: vowel nasalization differences are added.

Figure 3.13 (right) illustrates the mean values for OVL differences as a function of vowel nasalization differences in CVNCV – CVNV contexts for each vowel per participant. As was described in section 3.1.2, vowel nasalization was significantly extended in CVNCV contexts except for /a/. At the same time, OVL was decreased. This is also illustrated by fig. 3.14, which depicts the means of OVL and vowel nasalization for each vowel in CVNV versus CVNCV contexts, averaged across speakers. Longer OVL is evident in the CVNV context, while vowel nasalization is slightly decreased. Correspondingly, vowel nasalization relative to OVL is proportionally enlarged when the nasal is followed by an oral stop.

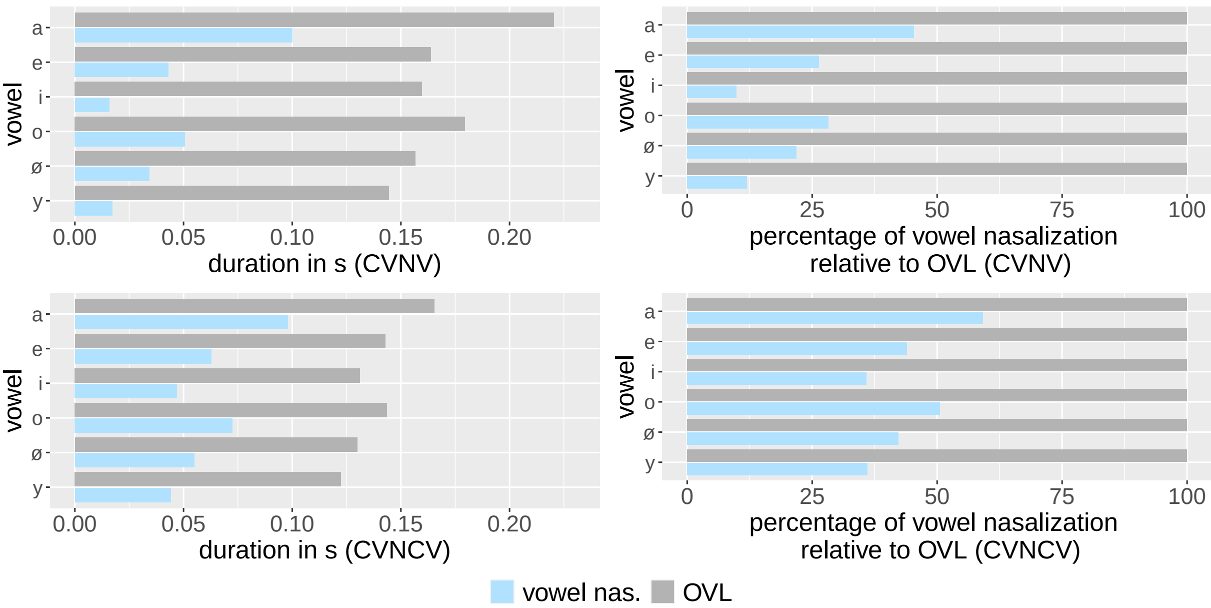


Figure 3.14: Means of vowel nasalization and OVL in absolute and relative terms separated by CVNV and CVNCV contexts (tense vowels).

Post-nasal context: lax vowels Figure 3.15 illustrates differences between CVNCV and CVNV contexts for both OVL and vowel nasalization for target words with pre-nasal lax vowels. Similar to the tense vowels, a tendency is apparent for negative values along the x-axis in 3.15 (left), indicating that OVL was systematically longer in CVNV compared to CVNCV contexts. Applying a linear mixed model with OVL as the dependent variable, results suggested both the context $F[1, 36] = 52.69$, $p<0.001$) and the vowel category ($F[4, 10] = 4.28$, $p<0.05$) as main effects, with no interaction reported. Figure 3.15 (right) includes differences for vowel nasalization as well, for which statistical analyses revealed that lax vowels in CVNCV contexts were significantly more nasalized than in CVNV sequences (see section 3.1.2).

The relation of vowel nasalization to OVL is depicted in fig. 3.16. As with the data for tense vowels, CVNV sequences with pre-nasal lax vowels involved overall longer OVL duration than the CVNCV contexts. Proportionally to OVL, vowel nasalization was slightly extended in CVNCV contexts.

Interim discussion

Results provided evidence that the extent of OVL was affected both by the post-nasal context and partly by vowel tensity: for CVNCV sequences, the time span of overall velum lowering was increased when tense vowels were involved rather than lax vowels. This pattern could not be confirmed for the CVNV contexts, in which OVL had a similar extent for both tense and lax vowels except for those cases involving /a/. On the other hand, OVL was affected by the post-nasal context in both tensity groups: irrespective of whether the pre-nasal vowel was tense or lax, overall velum lowering was shorter in CVNCV contexts. With respect to the effect of the post-nasal context, the point of velar closure was expected

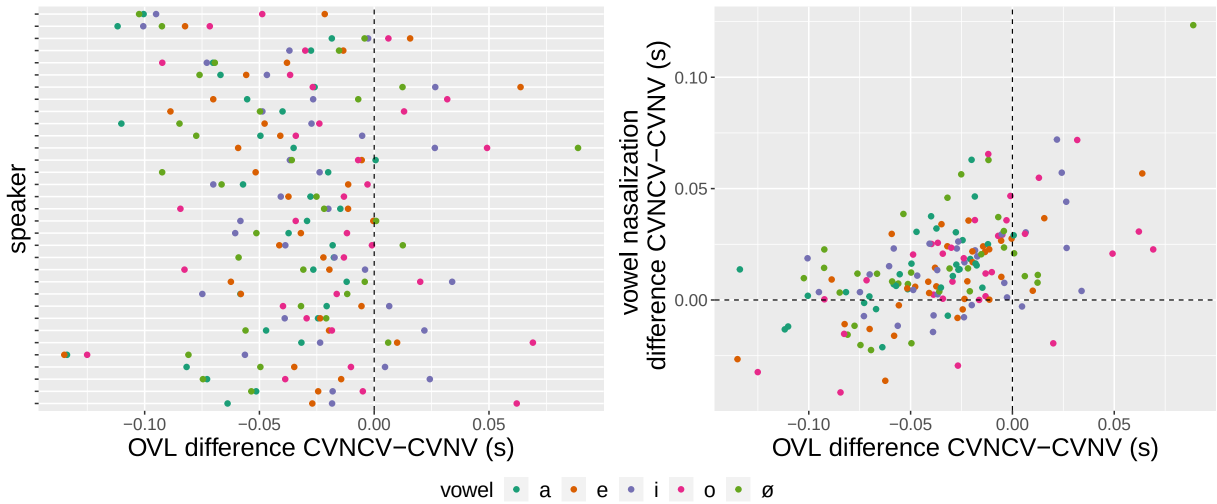


Figure 3.15: Left: differences of the overall extent of velum lowering (OVL) for CVNCV–CVNV contexts (lax vowels). The single values refer to mean differences of OVL per participant per vowel. The ticks on the y-axis refer to speakers S03–S38 from bottom to top, omitting S11, S25 and S30. Right: vowel nasalization differences are added.

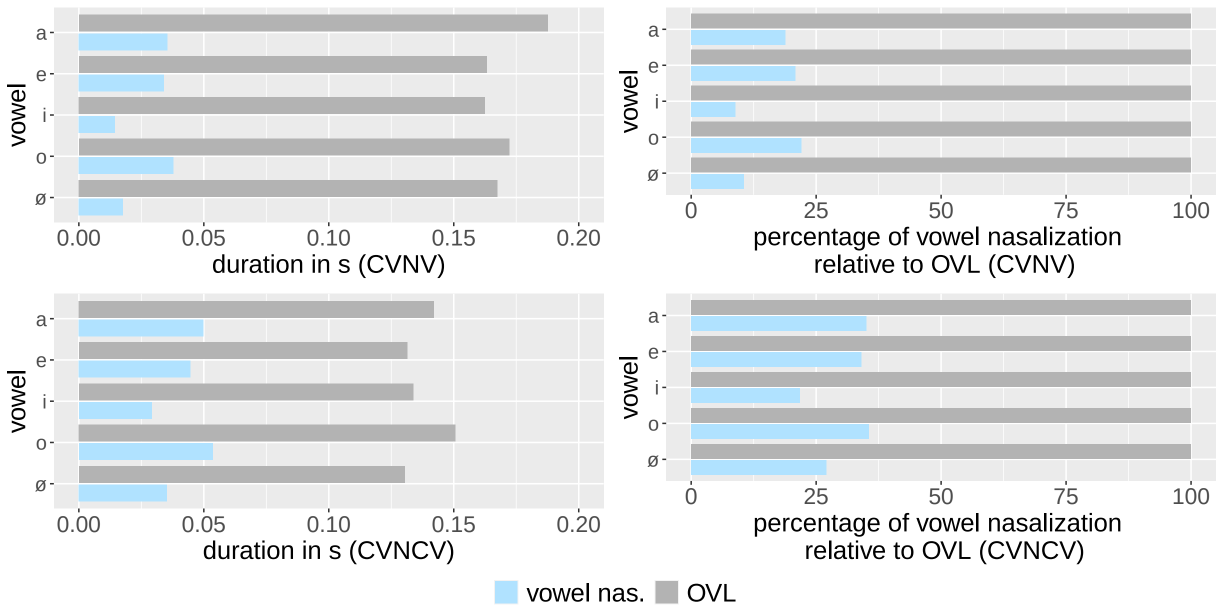


Figure 3.16: Vowel nasalization and OVL in absolute and relative terms separated by CVNV and CVNCV contexts (lax vowels).

to occur later in CVNV contexts because a post-nasal weak vowel was not supposed to exhibit full velar closure for production. This scenario might have indeed contributed to the patterns delineated before: OVL was found longer in CVNV contexts, while the vowels were less nasalized compared to CVNCV sequences. This suggests a velar gesture that started later in the vowel and extended throughout the nasal stop to the schwa vowel. In other words, both the velum opening and closing gesture occurred at an earlier time point in the CVNCV contexts. Although this seems to be compatible with the concept of a stable gesture being shifted along the segments, prediction H2 still is challenged: on the one hand, vowels were more nasalized in CVNCV sequences, but on the other hand, OVL was generally shorter in CVNCV. This suggests that a) the velum lowering gesture began earlier in the vowel, but b) not to the extent predicted by H2. In case of a shifted constant lowering gesture, vowels in CVNCV should have exhibited even more nasalization to compensate for the earlier closing gesture. However, this was not the case: CVNC and CVNCV sequences differed with respect to the time span of overall velum lowering. Nonetheless, the general differences in the initiation of velum lowering between the two contexts are to be considered in the context of a trading relationship between the nasalized part of the vowel and the duration of the nasal stop. The data above indicate that vowel nasalization is in fact increased when velum raising is required immediately after the nasal stop, but it is less pronounced when there is no need for rapid raising. Moreover, most of the research dealing with the trading relationship hypothesis is engaged with nasalization differences in CVNC sequences that primarily differ in the voicing of the post-nasal stop. In contrast, this section considered the temporal extent of velum lowering in two different sequence patterns, which is different from comparing the effects of post-nasal voicing. More appropriate are the data provided by Carignan et al. (2021), who, based on the same overall corpus used for this thesis, investigated velum movement patterns in words containing /Vnd/ vs. /Vnt/ sequences. They found only marginal differences in the time point of the velum lowering onset during the vowel but a decreased duration and magnitude for the nasal stop in /Vnt/ relative to /Vnd/ sequences. This alteration of proportionality was suggested to play a key role for the perceptual re-weighting of the acoustic cues on the part of the listener, who is assumed to pay closer attention to vowel nasality if it is increased proportionally to the overall nasal gesture. The data presented in this chapter contribute to this assumption: velum lowering started earlier in the vowel when the post-nasal stop was followed by an oral stop, while at the same time, the overall lowering gesture was reduced compared to sequences with a post-nasal weak vowel, resulting in an increase of the relative proportion of vowel nasality.

Moreover, prediction H1 seems to be in conflict with the findings. In CVNCV context, tense vowels were found to be more nasalized and OVL was increased compared to lax vowels. H1 assumes that, according to the trade-off hypothesis, no large velum lowering difference should occur, such that velum closure is supposed to shortly appear after a heavily nasalized vowel and in turn that the velum raises at a later point if the preceding vowel is only slightly nasalized. In our data, however, OVL in CVNCV was longer when tense vowels were involved, although these vowels were also nasalized to a larger extent than the lax vowels. Considering the CVNV context, no large effects were reported with respect to the impact of vowel tensity on OVL, but also not on vowel nasalization. The only difference in both parameters was found for /a/: tense /a/ was more nasalized than lax /a/, but OVL was also increased when tense /a/ was involved, which argues against H1. These results may be specifically considered with respect to data provided by Onsuwan (2005) and Beddor (2007), who presented results from a series of production and perception experiments with American English and Thai speakers. One of the experiments investigated the duration of vowel nasalization and the nasal consonant in V:N and VN sequences (tense and lax vowels in English) based on acoustic analyses. In fact, short vowels were less nasalized and occurred with longer nasals and long vowels exhibited a larger part of nasalization and occurred with shorter nasal stops. The authors pointed out that "the total duration of acoustic nasalization (Ṽ plus N) [was] not precisely constant across long (tense) and short (lax) vowel contexts", but that "VN sequences in both languages exhibit[ed] the predicted trade-off in the relative durations of vowel nasalization and N" (Beddor, 2007, p. 252). The data discussed in this section suggest that in CVNCV sequences, OVL is different in contexts with tense vs. lax vowels[5]: as vowel nasalization is increased in tense vowels, OVL increases as well (fig. 3.11). Thus, if the increase of OVL can be explained by the increase of vowel nasalization, this allows for the consideration that the duration of the post-vocalic nasal segment in CVNCV may be unaffected by the preceding vowel nasality. Consistent with this consideration, in CVNV sequences, both vowel nasalization and OVL do not differ for tense and lax vowels; the only exception is /a/, for which both vowel nasalization and OVL are increased in the tense vowel.

3.3 Summary and discussion

As outlined in this chapter, vowel tensity, vowel nasalization and the extent of the overall velum lowering are related to each other in a complex way. Differences for these parameters were elaborated for pre-nasal tense versus lax vowels in CVNV versus CVNCV contexts.

Considering the effect of vowel tensity in CVNV contexts, vowel duration was found to be significantly longer in tense vowels except for /ø/. At the same time, vowel nasalization was different only for /a/ and /o/, with longer portions in the tense vowels. With respect to the proportional nasalization of the vowels relative to vowel length, differences were observed for the vowel categories rather than for vowel tensity: apart from /a/, tense and lax vowels were nasalized at similar proportions and nasalization usually started later than the vowel midpoint for both lax and tense vowels. This is compatible with scenario c) introduced in the beginning: the German data at hand suggest that velum lowering starts at variable points before the nasal consonant. Moreover, OVL duration in the CVNV context was not affected by the vowel tensity except for those items involving /a/, with longer OVL for tense /a/. Findings differed for the sequences that included a post-nasal oral obstruent (CVNCV): all tense vowels were longer than lax vowels, and all tense vowels were nasalized to a greater extent in absolute terms. However, relative to vowel length, tense and lax vowels showed similar proportions of nasalization, with /a/ and /i/ exhibiting the most differing values within their vowel category. In addition, OVL duration was generally

[5] It should be recalled that OVL is based on kinematic rather than on acoustic data. However, at least in the CVNCV data, OVL can be assumed to be comparable to the acoustic Ṽ N extent in the study, because OVL comprises the nasalized part of the vowel plus the nasal stop without any post-nasal material, as the velum must raise for the post-nasal oral stop.

increased when tense vowels were involved.

With respect to the impact of the post-nasal context on tense vowels, vowel duration was hardly affected, with /a/ showing increased length and /i/ exhibiting decreased length in CVNV compared to CVNCV contexts. In contrast, vowel nasalization was significantly increased in CVNCV sequences except for /a/, while OVL duration was overall longer in CVNV, with the exception of /y/. Correspondingly, within the interval of OVL, vowel nasalization was proportionately more extended in CVNCV than in CVNV sequences.

For lax vowels, vowel duration was not affected by the following context, while vowel nasalization was generally increased in CVNCV contexts. As with the tense vowels, OVL duration was found longer in the CVNV than in the CVNCV context. Vowel nasalization made up the larger proportion of OVL when the nasal was followed by an oral stop.

Thus, these data can be summarized as follows: within the same segmental context, OVL is increased as vowel nasalization is increased. This does not hold for the comparison of CVNCV and CVNV sequences, given the same vowel tensity: when the nasal is followed by a voiceless stop, vowel nasalization is increased, whereas OVL is reduced.

The initial purpose of this chapter was to investigate which intrinsic linguistic effects other than tongue height may show some impact on vowel nasalization patterns. As it turned out, both the vowel tensity and the post-nasal context were found to be relevant factors. In light of this, the question must be considered of how exactly these parameters interact with each other: does vowel nasalization depend on the intrinsic length characteristics for German lax versus tense vowels, or is it rather the post-nasal segmental context which affects both the extent of overall velum lowering and vowel nasalization?

Considering the data above, there was a clear pattern for vowels being more nasalized in CVNCV compared to CVNV contexts. If tensity was the only factor for increased vowel nasality, vowel duration should have also differed for the two context conditions to explain the systematic nasalization patterns. However, in our data there was no systematic vowel duration difference apparent in CVNV versus CVNCV contexts (tense and lax vowels considered separately), whereas the vowels were generally more nasalized when followed by a NC rather than NV sequence. For this reason, it is appropriate to assume that the post-nasal context plays a major role for the extent of vowel nasalization. However, the impact of vowel length on vowel nasalization is not to be ignored. Consistent with the general observation that specifically /a/ is often affected first in a process of nasal vowel evolution (Chen, 1972; Hajek, 1992, 1997; Hombert et al., 1979; Whalen and Beddor, 1989, see Hajek and Maeda (2000) for a discussion), the findings in this chapter indeed suggest that in tense /a/, velum lowering is initiated earlier both in absolute time and relative to vowel length compared to the lax vowel and also compared to the other tense vowels. As discussed in chapter II, this behaviour is likely to be motivated by physiological (i.e. palatoglossus connection) and perceptual factors. The finding that in our data velum lowering occurs so early in tense /a/ suggests some timing control on the part of the speaker, which may be consistent with a scenario in which the speaker is aware of tense /a/ tolerating a relatively open velopharyngeal port without eliciting the percept of nasality. On the other hand, as indicated by previous studies (Delattre and Monnot, 1968; Hajek and Watson, 1998; Whalen and Beddor, 1989), low vowels are perceived more easily as nasalized when their length is increased. The combination of both – speakers lower the velum early in the low vowel, while listeners perceive this nasality especially in long vowels – may lead to the phenomenon that across languages, long or tense /a/ is often affected first by contrastive nasalization.

Considering the extent of overall velum lowering in our data, OVL was affected by the context, with shorter duration in the CVNCV context. It is striking that vowel nasalization was increased in the CVNCV context, although an overall longer lower plateau was found for CVNV: contrary to the consideration in Krakow (1993, 1999), this longer low plateau did not elicit a more nasalized vowel. Instead, the pre-nasal vowel was more nasalized when the nasal segment was followed by an oral stop, a constellation with only little temporal space being available for velar adjustments.

In summary, findings are not compatible with all hypotheses formulated in this chapter: in agreement with H1 from section 3.1.2, vowel nasalization in tense vowels was found to be more extensive than in lax vowels. Contradicting H1 and H2 from section 3.2.2, however, the extent of overall velum lowering was found to differ for tense versus lax vowels in the CVNCV contexts as well as for CVNV versus CVNCV contexts: in CVNCV sequences, vowel nasalization was more extended in tense than in lax vowels, while the time span up to the point of velum closing was not reduced, resulting in a longer overall velum lowering gesture. Similarly, velum lowering started earlier in vowels of CVNCV compared to CVNV sequences, but at the same time, the temporal extent of the overall gesture was decreased. Thus, vowel nasalization was proportionally more extended relative to the overall velum lowering gesture in CVNCV contexts, which may lead to some re-weighting of the perceptual cues on the part of the listener who may start to pay closer attention to the nasalized part in the vowel (Carignan et al., 2021).

Chapter IV

Stress, speaking rate and velum control

Abstract

The temporal and spatial amount of vowel nasalization is explored for pre-nasal vowels and consonantal sequences produced in two different prosodic conditions in which the target word was uttered with either contrastive focus or broad focal accent. These data serve as basis for investigating the impact of prosodic enhancement on velar behaviour during vowels followed by a nasal stop as well as during consonantal /nd/ and /nt/ sequences. For the consonantal contexts, additional data are provided for the tongue tip gestures. Results indicate enhancements of both the lingual and velar gesture during /nd/ and /nt/ when produced with contrastive focus. In addition, distinctions between the two voicing contexts are evaluated, revealing that the velum is consistently lower in /nd/ than in /nt/ independent of the focus condition. However, while the tongue tip shows more presence in /nt/ than /nd/ in the broad focus condition, this difference is neutralized with contrastive focus, unlike the velar behaviour. Considering vowels, findings are generally compatible with those from prior research suggesting that stress enhances the velum position intrinsic to vowels under regular conditions: the velum tends to be higher for high vowels and lower for low vowels in words produced with contrastive focus.

In addition to the effect of focus, this chapter is also concerned with the impact of an increased speaking rate on the velum and tongue tip position in post-vocalic oral and nasal coronal stops. Moreover, the velum velocity is explored for the nasal stop. The positional differences during fast versus moderate speech are examined for each stop separately. Moreover, it is investigated to what extent the deviation patterns observed for the rate conditions differ for the stops and conversely, how the gestural deviation between the nasal and oral stop is affected by the rate condition. Results suggest an overall lower tongue tip position for the nasal compared to the oral stop. Furthermore, while both the velar and lingual gestures exhibit a reduced amplitude in fast speech during the nasal stop, this is observed only for the tongue tip gesture in the oral stop. Less differences are found between the fast and moderate condition with respect to the gestural deviation between the oral versus nasal stops. Furthermore, the tongue tip differences between the fast and moderate rate are similar for both stops, whereas the velum shows a significantly larger positional distance in the nasal stop. Moreover, the velocity of the velum movement during the nasal is found to be similar in both speaking rates. Considerations are outlined about the relationship between the velocity and amplitude of the velum gesture as well as the relationship between tongue and velum movement patterns during fast speech.

4.1 The role of stress

4.1.1 Introduction

The effect of lexical and utterance-level stress on the spatial and temporal properties of articulatory gestures is well investigated for numerous languages. For example, the tongue position during vowels has been found to be more peripheral when produced with stress, such that high vowels are produced with a higher tongue position and low vowels with a lower position compared to their unstressed counterparts (De Jong, 1995; Farnetani and Vayra, 1996; Kent and Netsell, 1971; Mooshammer et al., 1999; Straka, 1963). Likewise, lingual contact with the palate is increased for consonants (Cho and Keating, 2009; Giot, 1977; Meynadier et al., 1998) and jaw movements show enlarged amplitudes (Fletcher and Vatikiotis-Bateson, 1994; Harrington et al., 2000; Kent and Netsell, 1971). In addition, both lip protrusion and tension is enhanced with stress (De Jong, 1995; Slis, 1971). However, with respect to velum movement patterns, findings diverge in their results regarding the impact of stress on the temporal and spatial amount of the lowering gesture. While some studies reported a generally higher velum position in stressed nasal stops (indicated by less nasal airflow, Fougeron, 2001), others found the velum in a lower position, but only in the coda (Krakow, 1993; Vaissière, 1988). Similarly, varying patterns were reported for the velum position within vowels: for some speakers, the velum was generally found in a lower position than in unstressed vowels, while for others, the intrinsic velum height of the specific vowel was enhanced with stress (Krakow, 1993). Moreover, studies on the interplay between the oral and velar gesture during nasal stops indicated that these two gestures work independently of each other: in onset position, oral closing and velar lowering occurred synchronously, while in coda position, velum lowering preceded the oral closure (Byrd et al., 2009; Krakow, 1989). Other findings suggested that in word-final positions, the tongue tip might be reduced in its gestural amplitude, while the velum remained unaffected by the alteration of the oral gesture (Jaeger and Hoole, 2011).

All these studies indicate that stress affects the temporal and spatial amount of the participating articulators to some extent. The current chapter is concerned with the question of whether the German data at hand provide further insights into the impact of stress on vowel nasality in both spatial and temporal terms. Moreover, the interplay between the velum and the tongue in nasal consonants is considered in more detail to explore whether the alterations of the tongue tip during stressed versus unstressed sound sequences affect the patterns of velum lowering. Under the assumption that during a nasal stop some cue trading is involved between the oral and the velar gesture, a lower velum position may be expected if the tongue tip gesture is reduced in its amplitude, such that nasality is maintained even with no complete oral closure being provided. In this chapter, stress effects on the consonantal sequences of /nd/ and /nt/ (experiment IV) and the vowel (experiment V) are presented and discussed at the end of the respective sections. In the following, some of the studies mentioned above are presented in more detail to provide a short overview of the main findings from prior research.

Prior research on stress effects on velum movement patterns

While numerous studies have focused on the effects of the preceding and following segmental context on vowel nasalization, less research has been done on the impact of lexical or utterance-level stress on velum movement patterns in vowels or consonants. Likewise, little is known about the interplay between velar and lingual gestures in stressed versus unstressed contexts with respect to the temporal and spatial gestural amount. Most experiments investigating these aspects involved a small number of participants, resulting in partly contradicting findings.

In his cross-language analysis, Schourup (1973) provided evidence for stress as a relevant factor for vowel nasalization: when nasalization was present in a language, in many cases only stressed rather than unstressed vowels were nasalized. In other languages, nasalized vowels became denasalized when in unstressed position. Moreover, the author observed that if a language exhibited unstressed nasalized vowels, it also had stressed nasalized vowels; the reverse conclusion, however, was not true. In summary, Schourup characterized vowels as predestined for nasalization when they were "optimally low, back, and stressed." (Schourup, 1973, p. 213).

In her exploration of velum movement patterns in fluently spoken sentences, Vaissière (1988) presented data from two native speakers of American English. The data were acquired by means of a computer-controlled X-ray microbeam system and analyses were provided for 22 sentences comprising lexically stressed and unstressed CVC and CVN syllables in different positions. As a main finding, stress was reported to enhance the intrinsic velum position observed for consonants: in stressed CVN syllables, the velum was found higher for the oral consonant and lower for the nasal consonant compared to the unstressed counterparts. With respect to the velum movement velocity, the two speakers showed slightly different patterns in the explored contexts: for one speaker, the maximum velocity slope from the oral to the nasal consonant was higher in stressed syllables, while the other speaker showed no differences between the two contexts.

Considering the temporal interaction between the velum and the lips for bilabial nasals in varying word positions, Krakow (1993, 1999) recorded two speakers of American English by means of the velotrace[1] and an optoelectronic tracking system. Analyses were

[1] The velotrace (Horiguchi and Bell-Berti, 1987) is a mechanical device for tracking movements of the soft palate. It consists of an external and an internal lever, which are connected to each other. The internal lever is inserted into the nasal cavity, such that contact to the soft palate is induced and movements are directly transmitted to the external lever. Recordings can be accomplished by monitoring signals of infrared

provided for the timing patterns in target words with /m/ in word-initial, word-medial and word-final position. As a result, lip movement and the onset of velum lowering were found to occur roughly simultaneously at the beginning of a word, whereas in word-final position, the velum gesture preceded lip closure. When the nasal was located within the word and surrounded by two vowels (e.g. *homey, pomade*), the timing patterns corresponded to those found for either the initial or final word position: in these words, synchronous lip and velum gestures occurred when the syllable that followed the nasal had primary stress (as in *pomade*), whereas the velum gesture preceded lip closure when the stressed syllable preceded the nasal (as in *homey*). Based on these observations, and contrary to the concept of some kind of 'ambisyllabicity' of such segments, Krakow suggested "that the nasal consonant was typically affiliated with the preceding or the following vowel, but not simultaneously with both", such that "syllabification was determined by the stress pattern, with primary stressed syllables attracting the nasal consonant." (Krakow, 1993, 94).

More concerned with the effect of word stress rather than the position within the word, Krakow (1993) investigated velum movement patterns in /i/ and /a/ in pre-nasal and post-nasal positions of bisyllabic words, with stress on either the first or the second syllable (e.g. *babám, bábam, mábab, mabáb*). Two participants were recorded by means of the velotrace system. Velar height was measured at the onset, mid and offset of the target vowel. With nasals in word-initial position, intrinsic differences in velum height were found, with /a/ showing a lower velum position than /i/ at all measurement points. Considering the effect of stress, however, speakers differed in their strategy: while for one speaker, the intrinsic velar height differences were enhanced with stress, the other speaker exhibited a more lowered velum for both vowels in the stressed condition. Similarly, for vowels preceding a word-final nasal, intrinsic differences in velum height were evident, but only at the mid and offset measurement points. As with the findings from the word-initial context, the velum position was enhanced with stress for one speaker, while a lower velum position was observed in both vowels for the other speaker. Krakow noted that the finding of different articulatory strategies for implementing stress was in accordance with the varying results on tongue position (Krakow, 1993, p. 105), for which prior studies had shown that the intrinsic tongue position for vowels was enhanced for some speakers, but generally lowered for others. Moreover, and contrary to Schourup (1973), Krakow pointed to the observation that in her data stress did not generally promote vowel nasalization, but – in accordance with Schourup – that pre-nasal low vowels indeed were prone to becoming nasalized.

In a nasal airflow study with accompanying acoustic recordings, Busà (2003) provided data gathered from two speakers of Northern Italian who read out VNC sequences in bilsyllablic words with primary stress on the first syllable and trisyllabic words with primary stress on the second syllable. Both speakers showed overall longer vowel durations in the

light emitting diodes attached to the velotrace.

stressed syllable compared to the unstressed counterpart and for one speaker, the stressed vowel was also temporally nasalized to a significantly greater extent. For both speakers, however, the percentage of the nasalized portions in the two conditions was not significantly distinct.

Considering tongue movements during speech, it has been widely noticed that the tongue position is clearly affected by lexical and utterance-level stress. For various languages, findings suggest that vowels in accentuated position are articulated more in the periphery than unaccented vowels (De Jong, 1995; Farnetani and Vayra, 1996; Kent and Netsell, 1971; Straka, 1963) and that accentuated consonants show a higher amount of closing duration and contact to the palate (Fletcher and Vatikiotis-Bateson, 1994; Meynadier et al., 1998; Straka, 1963, see Fougeron, 1999 for a detailed review). Focusing on the interplay between lingual and velum movements, Fougeron (2001) investigated lingual contact and nasal airflow for five consonants and two vowels in the domain-initial positions of the syllable, word, accentuated phrase and intonational phrase of a sentence. For a total of four French participants, electropalatographic (EPG) and nasal airflow data (two and four speakers, respectively) were presented. Speakers read out complete sentences with the target word being placed at the respective position. Results showed significantly more lingual contact for the nasal consonant /n/ corresponding to the higher prosodic domain, that is, the largest contact was observed in the intonational domain, followed by the accentuated phrase. Similar patterns were evident for /t/, although here, distinct lingual contact was observed for only three of the four domains. Moreover, the author noted that for both coronal stops, the contact surface tended to extend more toward the palatal center in a higher-level domain. Considering nasal airflow, the data showed the reverse pattern, with decreasing airflow when the nasal was produced in higher prosodic domains. In many cases however, the only significant distinction was found between the intonational phrase and the other three domains, suggesting that nasal airflow showed less variation between the four domains than the EPG results. The author assumed that this might have been due to a smaller degree of freedom of the velum position compared to the tongue and that airflow measurements could be insensitive to slight variations in velum position (Fougeron, 2001, p. 125). The finding that nasal airflow was decreased in the initial position of the accentuated and intonational phrase was further suggested to reflect articulatory strengthening in the sense of increased contraction of those muscles that were involved during a specific articulatory gesture. Hence, decontraction of the levator palatini might have been reduced, such that velum lowering was counteracted to a larger degree (Fougeron, 2001, p. 132).

Following Krakow (1993), Byrd et al. (2009) investigated timing interactions between the tongue and the soft palate in three different syllable conditions and three different stress conditions. Unlike Krakow, however, the authors explored movement patterns during the nasal alveolar stop instead of bilabial stop. Data were presented for four native speakers of

American English, who were measured via real-time MRI with a temporal resolution of 89 ms (11 fps). The authors investigated the order of the gestural adjustments by identifying the time lag between the tongue tip plateau region and the velum gesture plateau region. In addition, the velum displacement, i.e. the spatial magnitude of the velum gesture was examined. With respect to the syllable position, findings confirmed the patterns described by Krakow: near-synchronicity between the tongue and velum movements was found in the onset, with the tongue gesture slightly preceding the velum gesture. Coda or geminate nasals, in contrast, showed the opposite pattern, with the velum lowering gesture clearly preceding the tongue tip movement. Considering stress patterns, the stimuli set comprised single target (non-)words with a word-medial alveolar nasal in three stress conditions: primary stress on the post-nasal syllable (e.g. *beknow*), secondary stress on the post-nasal syllable (e.g. *bono*) and main stress on the pre-nasal syllable (e.g. *bonafide*). Indeed, the temporal adjustment between the tongue and the velum differed for these conditions: when the pre-nasal syllable had main stress, the velum gesture preceded the tongue tip gesture, reflecting the pattern found for the coda position. In words with primary stress on the post-nasal syllable, the tongue movement preceded the velum lowering gesture, corresponding to the pattern found for the onset. These findings were compatible with Krakow's suggestion that the nasal consonant was attracted by stress. With respect to the spatial displacement of the velum, participants generally showed differences in velum position for the varying conditions, but these differences were highly individual and not consistent across speakers.

Focusing on regressive place assimilation across word boundaries in German, Jaeger and Hoole (2011) investigated tongue tip movements and velum gestures in word-final /n/ and /t/ when these were followed by a word-initial velar or bilabial oral stop. In addition, the effect of word frequency was considered. For data acquisition, four native speakers of German were measured by means of electromagnetic articulography (EMA), in which the movements of the tongue tip, mid and back were captured as well as the movements of the upper and lower lips and the velum. With respect to the tongue tip, findings clearly pointed to lower positions in the nasal than in the oral stop as well as in high frequency compared to low frequency words. Furthermore, it was considered whether the reductions of the tongue tip movements were also reflected by the velum position. However, data did not suggest a close coupling link, indicating that the amount of velum lowering occurred independently of the magnitude of tongue tip closure, at least in the condition under investigation. The authors suggested that "[s]peakers only receive the freedom to reduce the [tongue tip] gesture if they simultaneously ensure that nasality remains robustly present acoustically, by this means ensuring that the weakening of place of articulation information does not become too salient for the listener" (Jaeger and Hoole, 2011, p. 422). Although this study did not focus on stress effects in particular, its findings nonetheless are related to the current exploration: if velum lowering is largely independent of the magnitude of tongue tip reduction, it may be interesting to see what patterns are evident in the reverse case, i.e. when the tongue tip has pronounced contact with the palate in stressed contexts.

Summing up these findings, stress appears to affect both the temporal adjustment patterns as well as the spatial amount of velum lowering in the following ways:

- speakers exhibit individual strategies of velum lowering during stressed vowels in nasal contexts: the intrinsic velum position associated with specific vowels is enhanced with stress or the velum is generally lowered in stressed positions (Krakow, 1993)
- speakers exhibit individual strategies of velum lowering during stressed nasal stops, with some showing lower velum positions in the stressed nasal and others showing less nasal airflow, indicating a higher velum position (Vaissière, 1988; Fougeron, 2001)
- velum lowering in nasal consonants is associated with the adjacent stressed rather than unstressed speech material, such that the lowering gesture precedes oral closure when the nasal follows a stressed syllable, whereas it is synchronized with the oral gesture when the nasal precedes the stressed syllable (Byrd et al., 2009; Krakow, 1993, 1999)
- considering the spatial interplay of the tongue and velum movements, a reduced raising gesture of the tongue tip may be accompanied by a certain degree of velum lowering sufficient to generate the concept of nasality (Jaeger and Hoole, 2011)

In the following sections, data are presented for Standard German, which contribute to further insights on the impact of stress (in this case: focus) on the spatial and temporal extent of velum lowering both during vowels and nasal consonants. With respect to the consonantal sequences, additional findings are outlined for the tongue tip position in the respective stress conditions.

4.1.2 Experiment IV: Focus effects on oral and velar gestures in coronal consonants

Most of the experiments outlined involve comparisons between target words with different lexical stress patterns to explore the impact of stress on velar behaviour. In contrast, the data in the following comprise velum lowering patterns for target words that were originally read out in three different prosodic conditions. Thus, the specific target word was the same in all conditions and had lexical stress on the primary syllable, but it was uttered with different focal accents. In condition 1, the target word was nuclear accented, such that the pitch accent occurred on the syllable with primary lexical stress. This condition refers to a stress pattern that is often used in a broad focus context (henceforth BF). Condition 2 involved target words with contrastive focus (CF) in the sentence. Condition 3 applied contrastive focus on the final verb, such that the target word was pre-focal unaccented (PF). Examples of the three conditions are given below (repeated from section 1.6.1):

Conditions for target words (example: Künste)

1. broad focus (BF): *Wieder Künste gesagt.*
2. contrastive focus (CF): *Bis er Künste sagt.*
3. pre-focal unaccented (PF): *Bis er Künste sagt.*

As indicated in the introduction, numerous speakers had difficulties with consistently producing the target word in the PF condition with pre-focal accent and the final verb with appropriate contrastive focus, i.e. with a higher amplitude and pitch than the target word. Although speakers were instructed in a training session and corrected in their stress patterns if necessary, many of them still had problems in providing the verb with sufficient contrastive focus, such that the target word was often not as focally unaccented as required. This impression was confirmed by considering the vowel duration based on the acoustic analyses: vowels in the target words were longer when produced with focal accent compared to pronunciations with broad focus. However, no significant differences of vowel duration were apparent between the broad focus condition and the pre-focal accent condition. Figures 4.1 and 4.2 give a general impression on vowel duration differences between the CF and BF conditions as well as between the PF and BF conditions for tense and lax vowels.

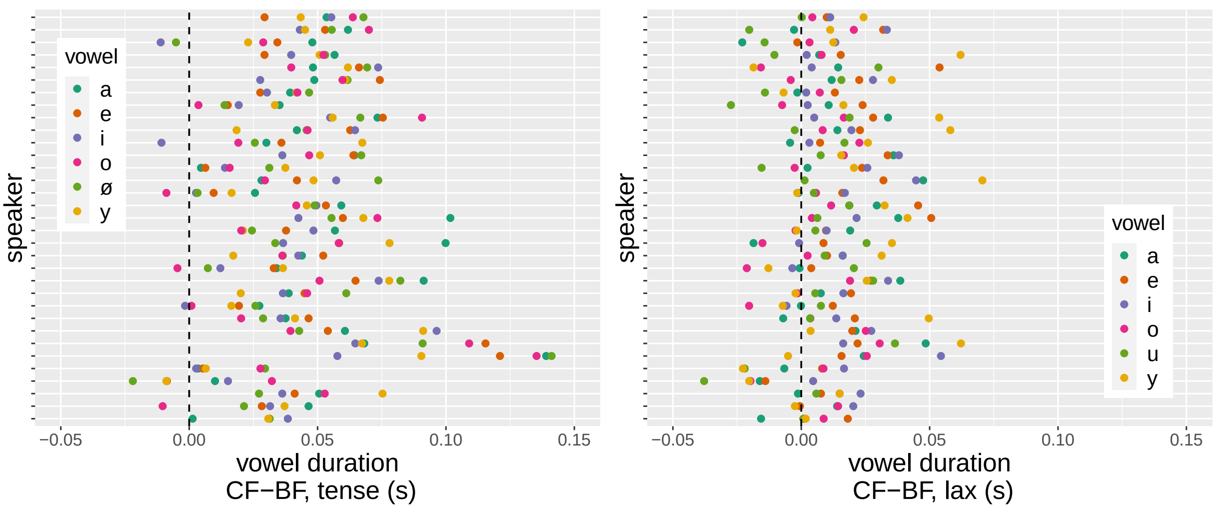


Figure 4.1: Vowel duration differences of tense and lax vowels with contrastive focus (CF) versus broad focus (BF). Mean values are given for each vowel per speaker. The ticks on the y-axes refer to speakers S03–S38 from bottom to top, omitting S11, S25 and S30 due to registration issues (see section 1.6.1).

A clear tendency is evident for vowels with focal accent being longer than their BF counterparts (fig. 4.1), which is reflected by the overall positive values. In contrast, no clear difference in vowel duration is apparent between BF vowels compared to the their

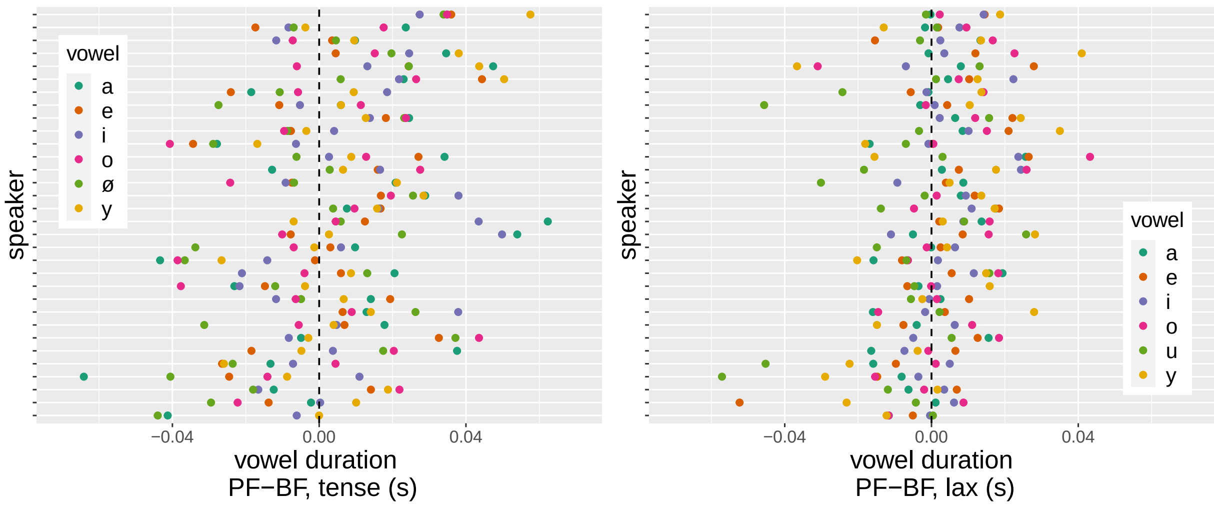


Figure 4.2: Vowel duration differences of tense and lax vowels with pre-focal accent (PF) and broad focus (BF). Mean values are given for each vowel per speaker. The ticks on the y-axis refer to speakers S03–S38 from bottom to top, omitting S11, S25 and S30.

counterparts in the PF condition (fig. 4.2). Statistical analyses (provided in the appendix B.1) confirmed the overall tendencies: differences between CF and BF were found to be significant, while this was not the case for the BF versus PF conditions.

Given the problems with the PF condition experienced by the speakers, and given that BF and PF did not differ in vowel length (and also probably not in other acoustical aspects), the data presented in the following include differences between the contrastive focus and broad focus conditions rather than between contrastive focus and pre-focal accent.

Predictions

Velum and tongue positions are examined in /nd/ and /nt/ sequences in CVNCV target words uttered with contrastive focus (CF) and broad focus (BF). To explore the effect of focus and voicing on the velar and alveolar region, the maximum signal values during these sequences estimated by the PCA (for the velum signal) or the vocal tract aperture (VTA) function (for the tongue signal) are considered. The effect of the voicing of a post-nasal stop on the preceding sound segments has been considered in prior research, as delineated in chapter III. Findings generally suggest that when the nasal is followed by a voiceless rather than voiced obstruent, vowels are often nasalized to a greater extent and the nasal consonant is shorter in duration (Beddor, 2007, 2009; Onsuwan, 2005; Sefton and Beddor, 2005). In a recent study, Carignan et al. (2021) provided evidence that the overall velum lowering gesture in German (extending to the pre-nasal vowel) was slightly reduced both in spatial and temporal terms in words with /nt/ contexts compared to /nd/. This reduction primarily affected the nasal stop interval, suggesting that the articulatory conflict between a nasal consonant and a voiceless stop was resolved by shortening the nasal duration and diminishing the amplitude of the velum lowering gesture. While this study focused on target words with no specific focus contrast, the current experiment explores whether the voicing effect is also apparent when the target word is produced with contrastive focus. As a prediction, the velum in our data is expected to show less spatial lowering during /nt/ compared to /nd/ in both focus conditions, because independently of the stress pattern, the articulatory requirements of the two consecutive sounds are more incompatible if the nasal is followed by a voiceless rather than voiced stop. This point will be further considered in the interim discussion of this section. With respect to the differences in the gestural amount of lingual closure between /nd/ and /nt/ sequences, the tongue tip is expected to show a higher degree of closure in /nt/ due to a typically longer temporal closure phase during the voiceless stop.

Taking the two focus conditions into account, lingual closure is expected to be increased for both /nd/ and /nt/ contexts in the CF condition, as suggested by the findings on stress effects in prior research (e.g. Fletcher and Vatikiotis-Bateson, 1994; Meynadier et al., 1998; Straka, 1963). However, it is not clear what to expect with respect to velum behaviour. On the one hand, velum lowering might reflect the gestural enhancement of the tongue position by exhibiting larger degrees of lowering in the CF condition. This might be true especially for the /nd/ context, for which the velum is not required to strictly close the velopharyngeal port for the upcoming voiced stop, at least not until stop release (Ohala and Ohala, 1991). On the other hand, considering the findings from Jaeger and Hoole (2011), some kind of cue trading could be involved: when the velum in /nd/ is lowered to a certain extent, allowing for sufficient nasality, the tongue tip may not induce strong alveolar contact. This may be the case, however, only for target words with BF because if strong lingual contact is initiated in the CF condition, the velum might not be lowered to the extent observed for the BF condition due to a possibly longer alveolar closure for the oral stop, requiring a more pronounced velar closure similar to that for /nt/. Thus, because the velum may be either lowered (showing enhanced amplitude) or raised (ensuring velar closure) with contrastive focus, no directional prediction is given for /nd/ with respect to velum lowering alterations in the CF vs. BF condition. For those contexts involving /nt/, however, the velum is expected to show a higher position in the CF condition to ensure sufficient closure for a voiceless stop that is likely to be produced with an enlarged closure phase. Based on these considerations, the following predictions are to be tested:

Hypotheses: The effect of focus and voicing on velum lowering and tongue tip participation in /nd/ and /nt/

H1 In words with post-vocalic /nd/, the spatial extent of velum lowering is greater than in /nt/.

H2 In words with post-vocalic /nd/, the tongue tip position is lower that in /nt/.

H3 In post-vocalic /nd/ and /nt/ sequences with CF, the tongue tip position is generally higher compared to BF.

H4 In post-vocalic /nd/, the spatial amount of velum lowering is affected by the focus condition.

H5 In post-vocalic /nt/, the spatial amount of velum lowering is decreased in CF compared to BF.

Speech materials

The speech material consists of a subset of the original overall corpus and comprises target items with CVNCV sequences in which the pre-nasal vowel is lax. The post-vocalic consonants consist of /nd/ or /nt/ sequences (e.g. *Sonde* [zɔndə] ‘probe’ vs. *sonnte* [zɔntə] ‘sunned’). All target words were embedded into carrier phrases that were read with broad and contrastive focus on the target word. A total of 1,160 items from 17 words were analyzed: 279 /nd/ and 313 /nt/ items with contrastive focus plus 272 /nd/ and 296 /nt/ items in the broad focus condition. A detailed list of the target words is provided in the appendix (table A.9).

Participants and procedure

Data are presented from 33 native speakers of Standard German, who were measured via real-time MRI as outlined in the introduction. The spatial amount of velum lowering was quantified by PCA analyses (see section 1.6.4), such that higher values indicate a more open velopharyngeal port. The tongue position was captured by means of the VTA function (section 1.6.4). Thus, higher values of the tongue signal indicate a physically higher tongue position. Results are illustrated by plots showing differences between the two consonantal contexts or the two focus conditions for velum lowering and the tongue position. These differences were estimated as follows: for each target word, the interval of the post-vocalic consonantal context was determined manually from the acoustical data. The context comprised both the nasal and the stop without an intermediate boundary between these two segments. The left boundary referred to the point where changes in the F2 and F3 amplitudes were clearly visible during the transitions from the vowel into the nasal murmur. The right boundary referred to the point of the stop release, i.e. omitting the aspiration portion of the voiceless stops. Next, the maximum signal value within these consonant sections was determined for the alveolar and velar region by means of the VTA function or PCA, respectively. Afterwards, for each vowel per speaker, the means of the maximum values were calculated, such that subtractions were possible for either the two contexts (/nd/–/nt/) separated by the condition or the two conditions (CF–BF) separated by the context. Statistical analyses were carried out in the programming environment *RStudio* (version 1.2.5033) by applying linear mixed models with the *lmer* function from the *lmerTest* package. To test the hypotheses presented in this section, the degree of velum lowering (velar signal) and tongue tip presence (alveolar signal) was selected as the dependent variable, respectively. To determine the effect of the post-nasal voicing, data were separated by the focus condition (CF, BF), such that the context (/nd, nt/) and the category of the pre-nasal vowel (/a, e, i, o, u/) were defined as the fixed effects and speaker and word onset as the random effects. To figure out the impact of the focus condition, data were separated by the consonantal context. For these subsets, the condition (CF, BF) and the category of the pre-nasal vowel (/a, e, i, o, u/) were defined as the fixed effects and speaker and word onset as the random effects.

Results

Broad focus: /nt/ vs. /nd/ Figure 4.3 (left) shows the differences of the means per vowel per speaker for /nt/ versus /nd/ sequences, with the target words being produced in the broad focus condition. Differences of velum lowering are reflected along the x-axis, while the y-axis depicts differences captured from the alveolar region. A slight tendency is apparent for higher alveolar signals in /nt/ (indicated by the positive values), whereas a more intense velum lowering signal is evident for /nd/. Statistical results suggested the post-nasal stop voicing ($F[1,34]=45.83$, $p<0.001$) and the pre-nasal vowel ($F[4,48]=18.05$, $p<0.001$) as main effects on velum lowering. An interaction ($F[4,434]=4.22$, $p<0.01$) showed significant velum lowering differences between the consonantal sequences independently of the preceding vowel category with the exception of /u/ ($p=0.0568$). With respect to the lingual movement patterns, a significant main effect was reported for both the voicing context and the pre-nasal vowel (context: $F=[1,33]=8.99$, $p<0.01$; vowel: $[4,20]=5.01$, $p<0.01$), suggesting overall increased lingual presence in the alveolar region for /nt/.

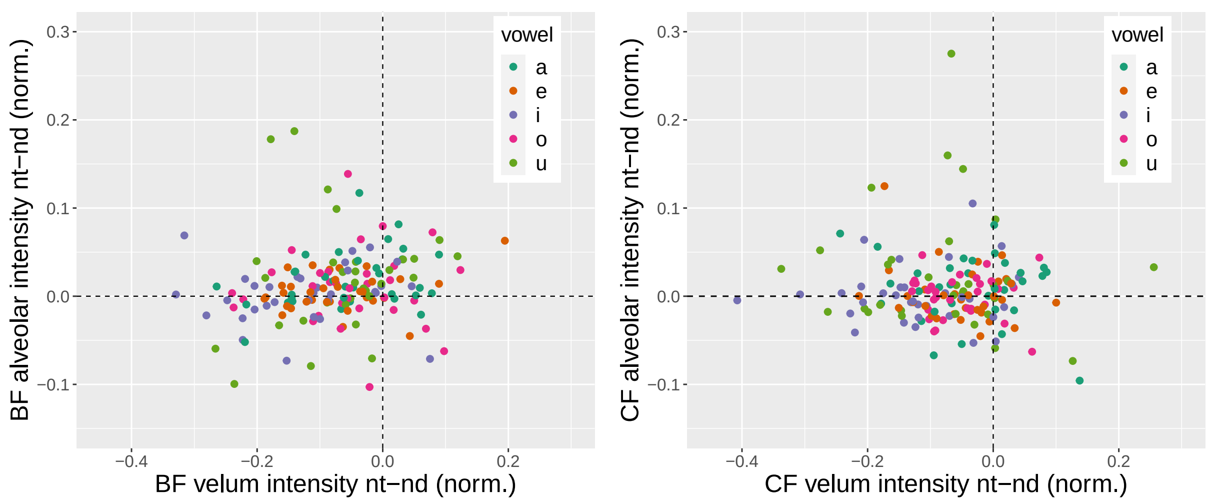


Figure 4.3: Mean differences of the signal intensity in the consonantal contexts /nt/ versus /nd/ for the velar region (x-axis) and alveolar region (y-axis). Left: broad focus condition (BF); right: contrastive focus condition (CF). Normalized mean values are given for the /nt/–/nd/ differences, separated by vowel per speaker.

Contrastive focus: /nt/ vs. /nd/ Figure 4.3 (right) shows the differences of velum lowering and tongue position between /nt/ and /nd/ sequences in the CF condition. Similar to the BF data, velum lowering appears to be increased in the /nd/ context. However, alveolar intensity does not show any tendency towards a specific context direction, indicating a similar tongue position during /nd/ and /nt/ when the word was produced with focal accent. Appropriately, statistical results indicated no significant effects of the post-nasal voicing or the pre-nasal vowel on the tongue position, but found velum lowering to be affected by both the context and the preceding vowel (F[1,37]=50.49, p<0.001 and F[4,22]=8.45, p<0.001, respectively). A significant interaction (F[4,417]=5.63, p<0.001) suggested overall higher velum signal values in the /nd/ contexts for all vowels except for cases in which the consonants were preceded by /a/ (p=0.1153).

Voicing context /nd/: CF vs. BF Next, differences between the CF and BF conditions are considered for the /nd/ context with respect to the velum and tongue movement patterns. Figure 4.4 (left) suggests a more lowered velum and a higher tongue position in the CF condition, indicated by more positive values for both articulatory regions. With respect to the velum intensity signal, statistical findings reported two main effects for the condition and the pre-nasal vowel (F[1,32]=11.26, p<0.01 and F[4,46]=7.08, p<0.001). In terms of the tongue position, a main effect was found for the condition (F[1,33]=37.45, p<0.001) as well as an interaction between the condition and the vowel (F[4,384]=2.64, p<0.05). Post-hoc corrected bonferroni tests showed significant differences between the two conditions for all vowels (/i/: p<0.05, other vowels: p<0.001), with overall higher values in the CF condition.

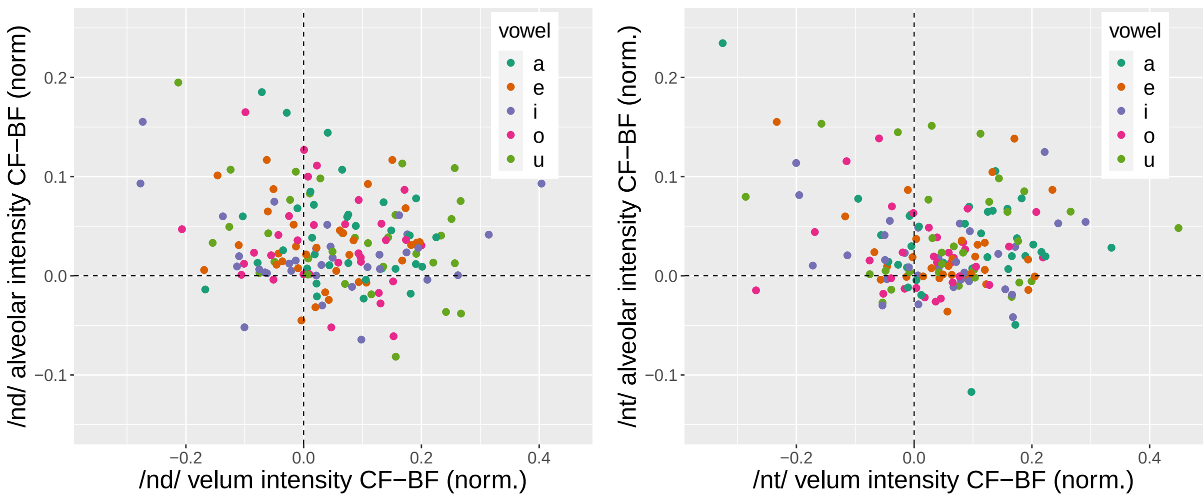


Figure 4.4: Mean differences in the signal intensity between CF vs. BF conditions for the velar region (x-axis) and alveolar region (y-axis). Left: /nd/ context; right: /nt/ context. Normalized mean values are given for the CF–BF differences, separated by vowel per speaker.

Voicing context /nt/: CF vs. BF The same tendency is also apparent for the /nt/ data (fig. 4.4, right) with slightly minor differences for the alveolar signal. Again, velum lowering appears to be more extensive in the CF condition. Statistical results suggested the condition and the pre-nasal vowel as main effects, with no interaction reported (F[1,32]=12.98, $p<0.01$ and F[4,31]=12.40, $p<0.001$). With respect to the tongue position, a main effect was found for the condition (F[1,34]=20.06, $p<0.001$) as well as an interaction between the condition and the pre-nasal vowel (F[4,453]=2.43, $p<0.05$). Significant differences in the alveolar signal were evident independently of the preceding vowel except for /o/ (p=0.0509).

Interim discussion

The exploration of the post-nasal voicing affecting the velum and tongue positions in /nd/ and /nt/ sequences revealed that on the one hand, the velum exhibited a lower position in /nd/ than in /nt/, which was evident in both focus conditions. The position of the tongue, on the other hand, was lower for /nd/ when produced with broad focus, but with contrastive focus, this difference was neutralized. With respect to the effect of focus, differences in the velum and tongue position were found in both voicing contexts: the velum exhibited a lower position when the sequence was produced with contrastive focus, while at the same time, the tongue tip was in a higher position.

Not all of these findings are compatible with the predictions suggested in this section. Velum lowering was indeed found to be increased for /nd/ compared to /nt/ (H1), and a higher tongue position was evident for /nt/ when pronounced with broad focus. However, when uttered with contrastive focus, /nd/ and /nt/ showed a comparable tongue position, which was not predicted by H2. The finding that in the CF condition, the tongue position in /nd/ became more similar to that of /nt/, but the velum did not, is in line with prior reports suggesting that velar and lingual movement patterns work independently of each other. Considering the concept of cue trading, one might have expected that due to the enhanced alveolar contact of the tongue in /nd/ in the CF condition the velum might also have taken a similarly high position as observed for /nt/. However, this was not the case: in the CF data presented in fig. 4.3 (right), the tongue tip had pronounced alveolar contact in /nd/ comparable to /nt/, but the velum was unaffected by the lingual adjustment, still showing a lower position in /nd/ than in /nt/. Moreover, the values referring to the velum position reflect the maximum signal achieved from the velum area during the interval of both consonantal segments in /nd/ and /nt/, respectively, indicating the maximum lowering position (which is strongly expected to occur during the nasal). As the velum is lower in the nasal stop of /nd/ compared to that of /nt/, it can be assumed that it does also not raise to the same extent for the post-nasal stop /d/ as for /t/, not even if produced with contrastive focus. The reason for these differing patterns may be related to the different requirements of the gestures to meet the articulatory adjustments: while strict velar closure is necessary throughout the pronounced voiceless stop, some velar leakage may be admissible during the initial part of the voiced stop, at least until the velum needs to be closed for the final release (cf. Ohala and Ohala, 1991). Our data suggest that this difference is still preserved even when the gestural amplitude is enhanced with contrastive focus.

Furthermore, prior research showed that velum raising is gradual and may continue even if the velopharyngeal port is already closed (Amelot and Rossato, 2006; Kuehn, 1976; Rossato et al., 2003). With respect to voicing effects on the velum position during oral stops, some studies found no impact of the voicing (Rossato et al., 2003; Ushijima and Hirose, 1974), while others reported that voiced stops were related to higher velum positions (or higher muscle activity in the levator palatini) compared to voiceless consonants (Bell-Berti and Hirose, 1973; Bell-Berti, 1975). This was explained by the speaker's effort to enlarge the pharyngeal cavity to maintain voicing, such that the supraglottal pressure is decreased and the increased pharyngeal cavity "would aid the maintenance of transglottal pressure differential necessary for the continuation of glottal pulsing through the period of vocal tract occlusion" (Bell-Berti, 1975, p. 457). In contrast, the data in the current experiment point to the opposite finding: although the velum lowering signal in this analysis is likely to refer to the nasal consonant and not to the oral stop (such that no final insight is available about its height during the oral stop), the lower position in /nd/ suggests that the velum does not arrive at the same height in the oral stop as for /t/. Instead, if the oral voiced stop tolerates some velar leakage at least to some extent, raising the velum higher than needed for cutting off the nasal leak is not required, apparently not even in a target word that is produced with contrastive focus.

4.1.3 Experiment V: Focus effects on velum lowering patterns in the vowel

The previous section investigated whether velar and lingual movements were affected by contrastive focus during post-vocalic /nt/ and /nd/ sequences and whether an interplay was apparent between the velar height and tongue position in the respective focus conditions. In addition, the effect of the post-nasal stop voicing was explored for both focus conditions separately. The current section focuses on both the temporal and spatial extent of velum lowering in tense and lax vowels in CVNC, CVNCV and CVNCC sequences that were produced with contrastive and broad focus. Although this corpus is different from the one used in experiment IV, it can be assumed that the velar adjustments during /nd/ and /nt/ in the different focus conditions seen in experiment IV can be transmitted to the sequences involved in experiment V, as the NC in these contexts refers to /nt/ (in CVNC and CVNCV) and /nst/ (CVNCC). Thus, the following data serve to examine how the temporal and spatial extent of velum lowering during vowels is generally affected by focus and whether the findings from experiment IV can be linked to the patterns of vowel nasality that occur with different focus conditions.

Predictions

Based on the findings from prior studies on lexical stress effects, comparable patterns of velum lowering are expected for our data considering the impact of prosodic focus. The predictions for experiment V are summarized in the following.

Hypotheses: The impact of focal accent on vowel nasality

H1 In words with focal accent, the spatial amount of velum lowering is generally enhanced compared to words with broad focus.

H2 In words with focal accent, the spatial amount of velum lowering intrinsic to vowels is enhanced compared to words with broad focus.

H3 In words with focal accent, the temporal amount of velum lowering is generally enhanced compared to words with broad focus.

Speech materials

Analyses of items with CVNCV, CVNC and CVNCC sequences in which the first V is either tense or lax are presented. The post-nasal consonants consist of either /t/ or /st/. As a prior survey to the main analysis revealed no differences on the temporal or spatial extent of vowel nasalization between words with /nt/ versus /nst/ (see appendix B.1.1 for statistical details), these groups were combined to constitute the basic subset. A total of 1,615 items from 24 words were analyzed: 827 items produced with contrastive focus and 788 items produced in the broad focus condition. Since during the measurement procedure, some stimuli blocks had to be repeated due to technical issues or mistakes in pronunciation, the numbers of the items are not equal for the two conditions. The detailed list and numbers of the target words can be looked up in the appendix (table A.8).

Participants and procedure

As with the previous chapters, data are presented for 33 native speakers of Standard German who were measured via real-time MRI as outlined in the introduction. The acoustic and kinematic analyses were accomplished in the same way as described previously (section 1.6.4). The measurement parameters for experiment V refer to a) differences in the spatial amount of velum lowering at the vowel midpoint estimated by the PCA and b) the duration of vowel nasalization, which was defined as the time span between the point of maximum velocity during the velum lowering gesture and the acoustic vowel offset (see section 1.6.4).

Statistical analyses were carried out in the programming environment *RStudio* (version 1.2.5033) by applying linear mixed models with the *lmer* function from the *lmerTest* package. To test the hypotheses given above, the duration of vowel nasalization and the degree of velum lowering were selected as the dependent variable, respectively. The focus condition

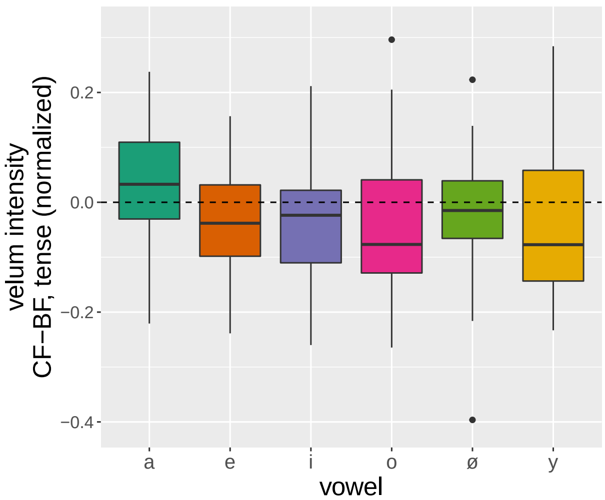
(CF, BF) and vowel category (tense: /a, e, i, o, ø, y/; lax: /a, e, i, o, u, y/) were defined as the fixed effects and speaker and word onset as the random effects. Data sets were tested separately with respect to tense and lax vowels.

Results

Velum lowering: tense vowels Figure 4.5 illustrates differences in the degree of velum lowering between the two conditions under investigation. The values refer to differences at the vowel midpoint. The means of the velum lowering signal were determined for each vowel per speaker for both the CF and BF condition, such that differences could be calculated between the two conditions. Thus, high values in fig. 4.5 correspond to a more intense signal (i.e. a lower velum) in the CF condition, whereas negative values indicate higher values in the BF condition. For tense vowels, fig. 4.5 (left) suggests a marginal preference for a higher degree of velum lowering in the BF condition, with the exception of /a/.

Statistical analysis was performed with the parameters mentioned above. A main effect was found for the vowel category (F[5,48]=34.33, p<0.001). In addition, a significant interaction was reported for the condition and the vowel category (F[5,705]=3.12, p<0.01). Post-hoc corrected bonferroni tests showed a slight tendency for vowels in the BF condition to be produced with a more lowered velum, except for /a/, where the reverse pattern was reported. However, none of these comparisons was rated significant.

Velum lowering: lax vowels Figure 4.5 (right) illustrates the mean differences in velum lowering at the vowel midpoint when lax vowels are involved. As with the tense vowels, a slight tendency is apparent for a lower velum position in the BF condition. Especially the high vowels seem to exhibit larger values when produced with broad focal accent.



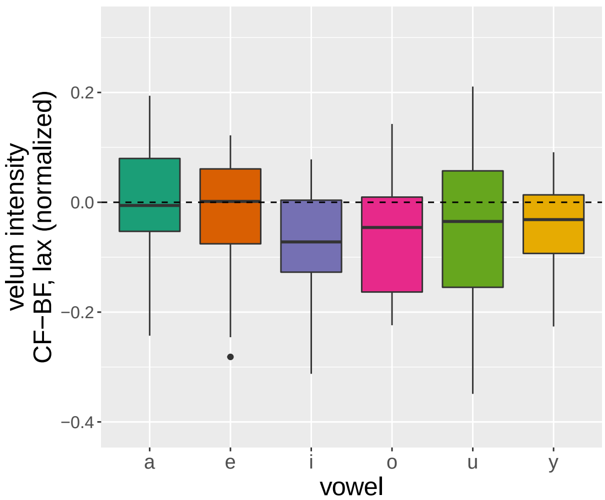


Figure 4.5: Differences of the spatial amount of velum lowering in tense (left) and lax (right) vowels at the vowel midpoint in two different focus conditions. Mean values are given for the differences of contrastive focus (CF) – broad focus (BF) for each vowel per speaker.

Statistical results indicated two main effects, one for the condition (F[1,35]=7.99, p<0.01) and one for the vowel category (F[5,6.10]=6.23, p<0.05). In addition, a significant interaction was reported (F[5,966]=4.06, p<0.01), suggesting significant differences in velum lowering only for /i/ and /u/.

Vowel nasalization: tense vowels Considering the temporal differences of vowel nasalization dependent on focal accent, fig. 4.6 (left) suggests that tense vowels produced in the CF condition do not generally exhibit a higher extent of vowel nasalisation compared to vowels with broad focus. A clear exception to this is the low vowel /a/, showing a longer portion of vowel nasalization in the CF condition. In fact, statistical analysis revealed two main effects for the condition and the vowel category (F[1,32]=14.11, p<0.001 and F[5,2]=52.44, p<0.05, respectively) as well as an interaction between these effects (F[5,702]=8.89, p<0.001). As suggested by the plot, only tense /a/ showed a significantly higher extent of vowel nasalization in the CF condition (p<0.001).

Vowel nasalization: lax vowels In fig. 4.6 (right), no clear tendency is visible for any of the lax vowels with respect to vowel nasalization in the BF versus CF condition. A main effect was found for the condition (F[1,40]=5.17, p<0.05) as well as for the vowel (F[5,6]=8.69, p<0.05). However, a significant interaction (F[5,576]=3.38, p<0.01) indicated no general tendency with respect to increased vowel nasalization in one specific condition (higher CF values for /a/ and /e/, higher BF values for the other vowels), and none of these comparisons was rated significant.

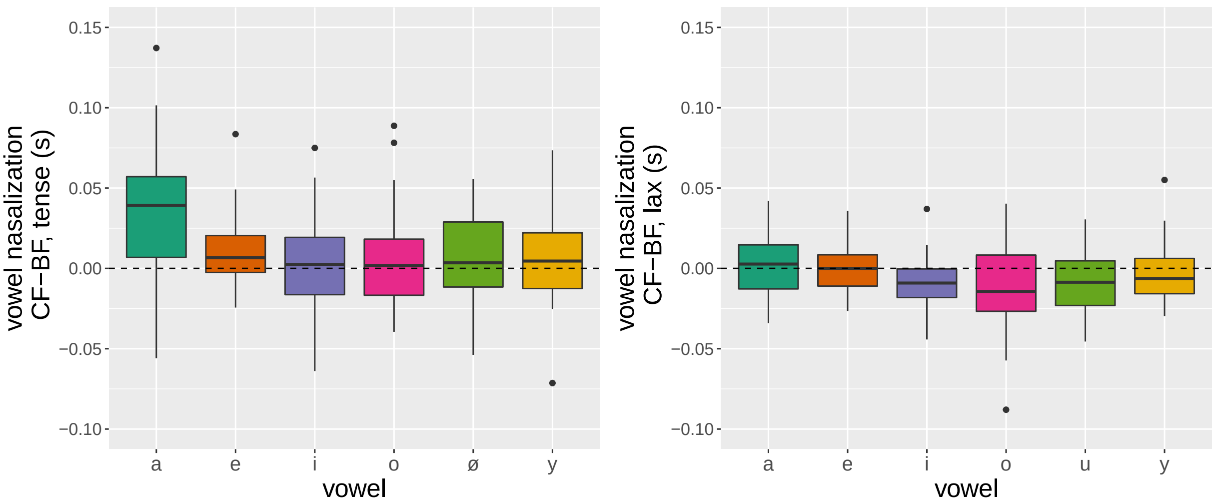


Figure 4.6: Differences in the extent of vowel nasalization in tense and lax vowels when produced in two different focus conditions. Mean values are given for the differences of contrastive focus (CF) – broad focus (BF) for each vowel per speaker.

Interim discussion

To investigate the impact of focus on vowel nasalization and the degree of velum lowering in tense and lax vowels, target words produced in two different focus conditions were analyzed. For tense vowels, the velum tended to be in a lower position in the broad focus condition except for /a/, for which it showed a slightly lower position when produced with contrastive focus. However, these tendencies were not found to be significant. In contrast, the temporal extent of vowel nasalization was significantly different for /a/, with a longer extent in the CF condition. Although a similar tendency was evident for the other vowels, no further significance was reported. Considering the nasality patterns for the lax vowels, results revealed a general tendency for vowels to be produced with a lower velum position in the BF condition, which, however, was significant only for /i/ and /u/. In terms of temporal vowel nasalization, no consistent tendency was evident for one specific focus condition, and for no vowel was the condition effect rated significant.

Several conclusions can be drawn from the overall findings. First, the results are not compatible with prediction H1: the velum was not generally found in a lower position in the CF condition. However, non-significant tendencies were apparent that are in line with H2: all high tense vowels showed a slightly higher velum position in the CF condition than in the BF condition, whereas for the low vowel /a/, the reverse pattern was found. This is in accordance with Krakow (1993, 1999), who suggests that stress enhances the intrinsic velum position in vowels, at least for some speakers. With respect to the lax vowels, /i/ and /u/ showed a significantly higher position when produced with contrastive focus, again suggesting that the velum position is enhanced with focus. However, as the overall data indicate only marginal alterations of the velum position dependent on the focus condition, statements about their compatibility with H2 must be handled with care.

Second, similar to the data on the spatial amount of velum lowering, the temporal extent of vowel nasalization was not significantly different for the lax vowels and was distinct only for the tense low vowel /a/. The finding that temporal nasalization in tense /a/ was affected by focus, but velum lowering was not, is somewhat unexpected if one considers the results of the previous chapters concerned with velum height and vowel length. In chapter II, the velum was clearly found in a lower position for pre-nasal tense /a/ compared to the other vowels. In chapter III, vowel nasalization was found to be more extended in tense /a/ than in lax /a/. As illustrated in fig. 4.1, vowel duration was in fact increased with contrastive focus, possibly leading to the same effect seen for tense versus lax /a/: a longer low vowel may be nasalized to a longer extent. It was further outlined in chapter III that the lower position found in tense /a/ is likely to be related to the starting point of the lowering gesture before the vowel midpoint, such that when the midpoint is reached, /a/ already exhibits a low velum, while the other vowels do not. Considering the current data, both vowel duration and nasalization of /a/ were increased with contrastive focus, while the spatial amount of velum lowering at the vowel midpoint remained largely unaffected.

This suggests that on the one hand, velum lowering starts earlier with contrastive focus in /a/ and on the other hand, more generally, data indicate that the degree of velum lowering may be independent of the temporal extent of vowel nasalization as long as a sufficient time span is provided for the velum to take a low position. In other words, in both focus conditions, velum lowering in /a/ is assumed to start before the vowel midpoint (at which the degree of lowering is captured). Although the lowering gesture generally starts even earlier with contrastive focus, this has no further effect on the velum position. Overall, findings are not compatible with H1 and H2 and are only in partial agreement with H3, namely for the tense low vowel /a/.

Summary and conclusion

The two sections explored the impact of contrastive focus on velar behaviour in German vowels and on tongue tip and velum movements in sound sequences consisting of a coronal nasal followed by a coronal voiced or voiceless stop. While during vowels, velum lowering was hardly affected by focus (with only /a/ showing a higher extent of temporal nasalization in the contrastive condition), a clear effect was evident for the consonantal contexts, in which the velum was found lower when the target sequences were produced with contrastive focus. At the same time, the tongue tip showed a higher position with contrastive focus compared to the broad focus condition. Interestingly, within the focus conditions, voicing had some impact on velum lowering as well, with a lower velum position being found when the nasal was followed by a voiced rather than voiceless stop. This pattern was consistent even with contrastive focus, different from the tongue position, for which voicing differences were apparent in the broad focus but not in the contrastive focus condition.

The finding that the velum lowering position during vowels was not affected by the focus condition suggests that there is no direct link between increased vowel nasalization and stress. In contrast, velum lowering during the /nd, nt/ clusters was more extensive with contrastive focus in both voicing contexts. Thus, our data suggest that in stressed environments, the nasal consonant is more enhanced rather than weakened: on the one hand, the tongue tip shows an overall higher position, suggesting that more contact is induced with the palate, on the other hand, the velum is in a lower position, allowing for a higher velar leakage that is characteristic for a nasal stop. In addition, and contributing to the findings from chapter III, results indicate that the voicing of the post-nasal stop affects the position of the velum, and it does so even with contrastive focus. This is compatible with assumptions about specific contexts that facilitate vowel nasalization, namely when vowels precede a nasal that is followed by a voiceless rather than voiced consonant (Beddor, 2009; Busà, 2007; Hajek, 1997; Malécot, 1960; Ohala and Ohala, 1991; Sampson, 1999). Thus, even if our data suggest that a) stress itself may play a minor role in affecting vowel nasality and b) that the differences in velum lowering during the post-vocalic consonants may not be accompanied by corresponding nasality variation in the preceding vowels (at least not at the vowel midpoint), they support nonetheless the idea that the voicing of the post-nasal stop is probably related to nasal weakening.

There are of course further options that generally come into question for investigating stress effects on articulatory enhancement and which may bring out differences more explicitly. For example, the stress effect may become more visible if a sound sequence with nuclear accent is compared to a matched sequence that is post-focally deaccented (e.g. *Bis er Sande sagt.* vs. *Bis ER Sande sagt.* 'Until he says sand.') (cf. Cho and Keating, 2009; De Jong, 1995; Kent and Netsell, 1971). This is because speech material which is not a point of information focus is likely to become more affected by temporal and spatial gestural reduction than speech material that carries main information (Eefting, 1991, 1992; Fowler and Housum, 1987). As indicated in the beginning of section 4.1, our database originally included target words that were supposed to be uttered as pre-focally deaccented, which, however, were finally excluded due to the insufficient contrast participants applied to the nuclear and pre-focal conditions. Although the data analyzed in this section exclusively include target words in nuclear accent position, findings indicate that simply providing the target word with narrow focus can have some impact on the gestural amplitude, at least in the consonantal sequences.

4.2 The role of speaking rate

4.2.1 Introduction

The previous chapter investigated focus effects on gestural behaviour. Another non-segmental parameter affecting velar and tongue movement patterns is the speaking rate. The effects of the speaking rate on the jaw, lip and tongue movement patterns have been well investigated by means of various measurement techniques such as EMA, X-ray microbeam or ultrasound recordings. When speakers produce syllables, words or sentences with increased speaking rate, they generally reduce the movement duration required for the sound sequences. However, such duration adjustments may affect different gestures and segments to varying extents. For example, in the phrase *tap a tad above* produced with fast speech, Adams et al. (1993) observed less reduction of the movement duration for the lower lip than for the tongue tip; in addition, they reported a higher consistency of the speaking rate effect for the opening than for the closing gesture. Gay (1981) found that vowels were proportionally more reduced in duration than their surrounding consonants. A general shortening of vowel duration was mentioned by Flege (1988), while Hoole and Mooshammer (2002) found a strong impact of the speaking rate on the length of German tense vowels but only a small effect on lax vowels.

To achieve some reduction of the movement duration, two basic strategies may be applied, which involve adjustments of the spatial amplitude, the velocity of the gestures or both. That gestures are reduced in their amplitude with fast speech has been widely reported. For vowels, Lindblom (1963) suggested a relation between vowel shortening and a more centralized position (or ‘undershoot’) of the tongue based on frequency changes in sound spectrographs. Later experiments with direct movement tracking confirmed his observation: Gay et al. (1974) used X-ray films and EMG recordings that revealed gestural reduction of the tongue during vowels accompanied by less activation of the genioglossus muscle. Similarly, Flege (1988) observed a lower tongue position for /ɪ/ and a higher position for /ɑ/ in fast speech (though not for all speakers). For consonants, Hertrich and Ackermann (2000) reported reduction effects on the amplitude of the tongue tip and lips in /tV/ and /pV/ syllables, respectively, but no consistent effects for jaw movements. Similar to the varying amplitudes for different articulators, velocity patterns during rapid speech were reported to differ for the jaw, lips and the tongue. For example, during the sequence /bab/, Gay (1981) found increased velocity for the opening and closing movements of the jaw but not of the lips. For the opening gestures of the tongue tip, Adams et al. (1993) observed a general effect on the number of velocity peaks, with one clear peak in fast speech targets compared to multiple peaks in slower-than-normal speech. Accordingly, velocity profiles were typically symmetrical in fast speech contrary to the asymmetrical profiles obtained from slower speech. Furthermore, especially for consonants, velocity alterations were also found to be strongly related to amplitude adjustments. Kuehn (1976) reported different strategies for different speakers, who either exhibited an increased velocity of the gestures while the amplitude remained unaffected, or, alternatively, showed similar velocity patterns independently of the speaking rate but exhibited a reduced amplitude with fast speech. The same finding was noted by Ostry and Munhall (1985), who considered speaking rate effects in terms of articulator stiffness, which they defined as the ratio of the maximum velocity to maximum displacement. This ratio was found to vary systematically with the duration of a gesture: a decrease of the gestural duration was accompanied by an increase of stiffness.

While the movements of the tongue, lips and jaw have been investigated from different perspectives, less is known about the impact of an increased speaking rate on velum movement patterns in spatial and temporal terms. Some of the basic findings with respect to velar behaviour and the interplay between the velum and tongue movements during rapid speech are outlined in the following.

Prior research on speaking rate effects on velum movement patterns

Using cinefluorographic films, Kent et al. (1974) presented tracing data of velum movements from two speakers of American English who produced sentences at a moderate and rapid speaking rate, with the fast sentences being uttered in about half the time of the moderate sentences. When the tracking curves for the fast sentences were expanded and superimposed, data showed that on the one hand, speakers preserved the relative timing patterns of the velum gestures in the fast rate condition, but on the other hand, they also showed different strategies for achieving the articulatory events in the rapid utterances: while one participant exhibited a reduced amplitude of the velum lowering and raising gesture, only small amplitude differences were found for the other speaker. The authors assumed that the latter participant, who showed an amplitude pattern in fast speech similar to that for moderate speech, might have exhibited an increased velocity of the velar gestures compared to the moderate speaking style. Overall, data suggested that speakers differed in their strategy of velum movement adjustments when producing sentences with rapid speech.

Similarly, Kuehn (1976) reported data from two speakers of American English, whose articulatory gestures were tracked by means of cineradiographic recordings. The speakers produced VNCV and VCNV non-words with a normal and fast speaking rate. For both speakers, the velar displacement and transitional duration were reduced with rapid speech, while the velocity was either increased or decreased. However, speakers were found to differ with respect to their strategy to reduce the velar amplitude: one speaker showed a similar velum position at its lowest point in both rates, but the maximum point during raising in rapid speech remained below that for normal speech. The converse pattern was found for the other speaker: here, the velum exhibited a similar height in its elevated position in both speaking styles, while its displacement during the lowering gesture was more distinct with normal speech compared to rapid speech. The author assumed that these different strategies were related to the overall velum position patterns reported for the two speakers: for the first one, prior trajectory analysis showed that the velum was still elevated even after contact had been initiated with the velopharyngeal wall. Accordingly, for rapid speech, "this movement during closure would appear to be aerodynamically dispensable and thus permissible for this subject to elevate the palate during rapid speech to a position just sufficient to achieve velopharyngeal closure " (Kuehn, 1976, p. 100). For the second speaker, in contrast, the velum did not exhibit an upward movement after velopharyngeal contact. Thus, an overall reduction of displacement in rapid speech was only possible by decreasing the lowering gesture. Overall findings suggested that the amplitude of velar movement was reduced and that speakers might have used different strategies to achieve gestural reduction.

Mixed results were also reported by Amelot and Rossato (2007), who considered velum lowering at the midpoint of nasal vowels when produced with different speaking rates. Two native speakers of French read out word sequences of /tṼ/ and /tṼt/ while being recorded via EMA. Speakers were supposed to read the carrier phrases with three different speaking rates indicated by a metronome. For data analysis, the amplitude of the velum lowering gesture was determined at the midpoint of the nasal vowel. While for one speaker, the degree of velum lowering was found similar in all three speaking rates, the other speaker showed a significantly higher velum in the fastest rate condition compared to the other rates.

More concerned with the general timing patterns of the velum, Bell-Berti and Krakow (1991) (reconsidered by Bell-Berti, 1993) argued for an increased temporal overlap of the velum gestures during adjacent segments produced with fast speech. Data were obtained by means of the velotrace device from three speakers of American English. Velum movements were investigated in CVN und CVC sequences in the carrier phrase *It's ... again*, with the number of the vowel varying from one single vowel to a sequence of several vowels plus /l/ (e.g. *It's ansal again* or *It's a lansal again*). The sentences were read out at a moderate and rapid speaking rate. Findings showed that during the transition from /s/ in *it's* to /n/ in e.g. *ansal*, velum lowering proceeded in terms of multi-stage levels in moderate speech, i.e. the velum showed a lowering stage for the vocalic material, which was lower than in /s/ but higher than in the nasal. With rapid speech, this stage was no longer observed, suggesting an increased temporal overlap between the oral and nasal segment. In addition, the velum peak at the release of /s/ was consistently higher in the moderate than in the fast rate condition, indicating less overlap between the gestures. The same pattern was found when the number of segments between /s/ and /n/ was increased, which was reported for both speaking rate conditions. Thus, results suggested that the separation between the extreme velum positions during /s/ and /n/ could be enlarged by either adding vocalic material or by slowing down the speaking rate (Bell-Berti, 1993, p. 107).

As noted previously in chapter III, Solé (1992) tested the hypothesis that vowel nasalization was phonological for American English in contrast to Spanish, expecting that the extent of vowel nasalization was adjusted relative to the speaking rate in English, but not in Spanish. American and Spanish participants were measured by means of the nasograph[2] (Ohala, 1971) while reading CVVC target words with five different speaking rates. Findings suggested multi-level stages of velum lowering during pre-nasal vowels for American English, but not for Spanish. In addition, velum lowering in Spanish started at a constant point before the nasal stop, while in American English, the extent of vowel nasalization was found to be proportionally adjusted to the speaking rate. Based on these findings, Solé suggested that vowel nasalization was phonologized to some degree in American English, while in Spanish, it was "the result of a physiological time constraint" (Solé, 1992, p. 38).

For European Portuguese, Oliveira et al. (2009) reported temporal reduction effects of velar gestures as well as a reduced overlap with the oral gestures in target sequences produced with rapid speech. Data were obtained from two speakers of European Portuguese by means of EMA; results were presented for one speaker who produced nasal vowels in the word-initial, word-medial and word-final position of non-existing words embedded in carrier phrases. The sentences were read out with two different speaking rates. Results indicated that both the overall velum lowering gesture and its single phases during the opening, plateau and closing phases of the nasal vowels were significantly reduced in duration in the

[2]See chapter III, section 3.1.1

fast productions. Next, the authors considered the parameter of stiffness, which referred to a measure of movement independent of its displacement and was achieved by dividing the peak velocity (cm/s) by the maximum displacement (cm) (Ostry and Munhall, 1985; Roon et al., 2007). The authors found an increase of stiffness during the opening and closing phase of the velum for word-medial nasal vowels when produced with fast speech. Finally, the inter-gestural timing was considered for the velum relative to the lip and tongue tip gestures. Data suggested that the oral target of the post-vocalic oral stop preceded the target of velum closing in both rate conditions, but this interval was found significantly reduced with fast speech, indicating an increased gestural overlap. Furthermore, based on additional acoustic analyses, the glottis was found to open before the velum closing target was achieved. Again, this interval was decreased with fast speech. In addition, the timing of the onset of velum lowering for the nasal vowel was considered relative to the release of the preceding oral stop. While in the moderate condition the velum onset started after stop release, the gestures occurred roughly simultaneously in fast speech.

Summing up the reports about speaking rate effects on velar behaviour, findings suggest that with fast speech,

- the lowering amplitude is generally reduced (Amelot and Rossato, 2007; Kent et al., 1974; Kuehn, 1976)
- the overall duration of velum lowering is decreased (Oliveira et al., 2009)
- movement velocity is increased (Kent et al., 1974; Kuehn, 1976)
- the temporal overlap of sequential velar gestures is increased (Bell-Berti and Krakow, 1991)
- the inter-gestural timing is affected, such that the time lag between the velar and oral gestures is decreased (Oliveira et al., 2009)
- multi-stage levels (if present in a language) of velum lowering during vocalic material are omitted (Bell-Berti and Krakow, 1991; Solé, 1992).

However, most of these studies also reported large variability across participants, with different speakers exhibiting different strategies to adjust their velum movements to rapid speech. The following experiment is concerned with speaking rate effects on the spatial amplitude of velum lowering and tongue tip movements as well as on velocity patterns of the velum during the lowering gesture in German. As the analyses comprise recordings from a larger group of participants, data may reveal some general tendencies of the specific gestural patterns as well as modifications of the interplay between the soft palate and the tongue tip during the coronal nasal and oral stop.

4.2.2 Experiment VI: Speaking rate effects on velum and tongue movements in post-vocalic /n/ and /t/

In this section, speaking rate effects on velar and tongue tip behaviour are investigated for post-vocalic nasal and oral stops in CVNV and CVCV words uttered in two different speaking rate conditions[3]. Although this survey explores movement patterns during consonants rather than vowels, insights from the inter-gestural arrangement in the post-vocalic context may nonetheless allow for some inferences on the temporal and spatial extent of the preceding vowel nasality. For example, the tongue tip may be found lower in /n/ in fast than in moderate speech, indicating less constriction, while at the same time, the velum exhibits a low position for the nasal. This gestural constellation may induce a nasalized sound that shows more vowel-like than consonantal properties. To illuminate such considerations, experiment VI was designed to explore the velar and lingual gesture patterns in the oral and nasal coronal stop in moderate and fast speech. More precisely, the questions are considered of a) how the gestures for the oral stop differ from those of the nasal stop in one specific speaking rate, b) how the gestures in the fast rate condition differ from those in the moderate style for one specific stop and c) how the speaking rate and the nature of the consonant interact with each other, i.e. whether the rate-dependent deviations are larger for one stop than for the other and conversely, whether the deviations between the nasal and the oral stop are greater in one specific rate condition than in the other. Data refer to movement patterns in spatial terms, for which the respective tongue tip and velum signal values during /n/ and /t/ were estimated by the VTA function and by PCA modelling, respectively (see section 1.6.4). In addition, the velocity of the velum gesture is considered for the nasal stop in CVNV sequences during fast and moderate speech.

Predictions

The predictions given below are based on assumptions that are reminiscent of those sketched for the effects of focus in section 4.1. With respect to the differences in the gestural amount of lingual closure between /n/ and /t/, the tongue tip is expected to show a higher position in /t/ (cf. Jaeger and Hoole, 2011) due to a generally longer temporal closure phase, which is required to generate sufficient intra-oral air pressure during the voiceless stop. This is expected for both speaking rate conditions. Moreover, in fast speech, the lingual gesture is expected to show a reduced participation both in /n/ and in /t/ compared to the moderate rate, but /n/ is assumed to be affected to a higher degree, i.e. the distance between the respective positions in fast versus moderate speech is assumed to be larger for /n/ than for

[3]The choice of the speech material may be somewhat unexpected in the light of the preceding chapter, from which one may expect a comparison between fast vs. moderate /nd/ and /nt/ sequences. However, as mentioned in the introduction of this thesis, the overall corpus was initially designed for covering a broad range of research questions on vowel nasalization, such that it does not contain all kinds of sound sequences in all prosodic conditions. For this reason, speech rate comparisons will be given for /n/ and /t/ rather than /nd/ and /nt/.

/t/. This assumption is based on the same consideration just noted: while for /t/, alveolar contact is required in both speaking rates (to slightly different degrees, though), the tongue tip gesture may be variably reduced during /n/ in fast speech by a considerable amount. Furthermore, the distance between the gestures for the oral and nasal stop should be found to be greater in the fast speaking rate, because the tongue tip still requires at least some constriction for /t/, whereas the gesture for /n/ may be highly reduced.

Investigating the differences of the velum position between the two stops with respect to fast and moderate speech is a bit more complex, because the direction of velar movement depends on the stop: a constricted velopharyngeal port, i.e. a raised soft palate is required during /t/ to build up intra-oral air pressure, while a low velum position is needed during the nasal stop to facilitate nasal airflow. A direct comparison between the velum position in /n/ with that in /t/ would thus be scarcely meaningful. Instead, however, predictions can be made about the specific behaviour for each individual stop with respect to the speaking rate and also with respect to the differences of the spatial distance between /n/ and /t/ relative to the speaking rate. Considering the velum gesture during /n/ in fast versus moderate speech, overall less lowering is expected in fast speech based on prior studies reporting a reduction of velar displacement. For /t/, only slight differences may be evident with respect to velar raising due to the reduced space when contact is initiated to the pharyngeal wall. Nonetheless, a somewhat higher velum position is expected in moderate speech compared to fast speech. With respect to the distance between the raised position in /t/ and the lowered position in /n/, it is predicted that this distance is decreased in fast speech compared to moderate speech: if articulatory gestures are generally reduced in fast speech, a less raised velum in /t/ and a less lowered velum in /n/ would be the consequence, such that the spatial distance between these extremes should be decreased.

Considering the velocity during the velum opening gesture in the nasal stop, prior studies reported mixed results, emphasizing the large variability found across speakers. Based on the reports by Kent et al. (1974) and Kuehn (1976), with speakers showing either increased velocity or reduced amplitude in fast speech, it is proposed that, given that the amplitude is affected, velocity will remain roughly stable across both speaking rates. These considerations are summarized below.

Hypotheses: The effect of the consonant and speaking rate on velum and tongue tip participation in /n/ and /t/

H1 In words with post-vocalic /n/, the tongue tip participation is decreased compared to /t/ both in fast and moderate speech.

H2 In post-vocalic /n/ and /t/ produced with fast speech, the tongue tip participation is generally reduced compared to moderate speech.

H3 In fast speech, the tongue tip shows a greater spatial distance between the oral and nasal stop than in moderate speech.

H4 In fast versus moderate speech, the tongue tip shows a greater positional distance in /n/ than in /t/.

H5 In post-vocalic /n/, the velum exhibits a higher position during fast speech compared to moderate speech.

H6 In post-vocalic /t/, the velum exhibits a higher position during moderate speech compared to fast speech.

H7 In fast versus moderate speech, the velum shows a greater positional distance in /n/ than in /t/.

H8 The distance of the velum position between the oral and nasal stop is more reduced in fast speech compared to moderate speech.

H9 The velocity during the velum opening gesture is stable across the speaking rates.

Speech materials

The speech material consists of CVNV and CVCV words produced in carrier phrases uttered at a moderate and fast speaking rate. Overall data include 2476 target words: with moderate speech, 599 CVNV and 646 CVCV words were recorded; with fast speech, 600 CVNV and 631 CVCV target words were obtained. For all CVNV words, the inter-vocalic nasal refers to the alveolar nasal stop /n/. The corresponding consonant in the CVCV words is the oral stop /t/. The pre-consonantal vowels are either tense or lax, for which separate analyses are provided. A detailed list of all target words and their number of occurrence is given in the appendix (table A.10).

Participants and procedure

As with the prior experiments, data are presented from 33 native speakers of Standard German who were recorded by means of real-time MRI. In addition, synchronized acoustic recordings were available (see section 1.6.3). As outlined in the introduction, the carrier phrases were presented on slides that switched automatically after a fixed number of seconds. To elicit differences in the speaking rate, the time interval of the slides was reduced, such that sentences with a moderate rate were presented for four seconds, whereas the fast speech phrases were displayed for only two seconds. Depending on the speaking rate condition, each recording block contained 14–19 sentences and speakers were informed about the upcoming condition before the rate switched. Data analyses were performed by capturing signal intensities from the MR images in pre-defined regions of the vocal tract by means of the VTA function as well as PCA modelling as described in the introduction (section 1.6.4). Acoustic analyses were performed by manually selecting the boundaries of the sentence, word, vowel and post-vocalic consonants. Prior to the main analyses provided in the following, the (acoustically determined) word and the vowel were tested for duration differences depending on the speaking rate to make sure that the two conditions exhibited significant rate differences (see appendix B.1.2 for details). Both the word and the vowel were found to be significantly reduced in duration when produced with fast speech.

Considering the quantification of the velum and tongue movement patterns in the post-vocalic consonants, data comprise the maximum alveolar signal captured during the post-vocalic stops, with higher values indicating increased participation of the tongue tip. With respect to the velum position, two different parameters were selected, dependent on the stop. As the velum signal was obtained from PCA and not from the VTA function, the polarity of the PCA signal was simply set such that higher positive values indicate more velopharyngeal port opening. Accordingly, for /n/, the maximum signal captured during the nasal stop was considered. For the oral stop /t/, however, this maximum signal parameter was assumed to be inappropriate because the velum inherently exhibits a high position during /t/ to constrict the velopharyngeal port. The maximum signal value would thus probably rather reflect the velum position right after the transition from the vowel to the stop (as the velum is supposed to be slightly lowered during the vowel), but not the maximum low position during the actual stop. For this reason, the minimum instead of maximum signal was considered for /t/, because this parameter was assumed to reflect the small differences in the raised position more reliably. The interpretation of the signal values, however, does not change: lower values indicate velopharyngeal closure, whereas higher values suggest more opening.

As with the data on focus, results on the speaking rate comparisons are illustrated by plots showing the differences in the tongue and velum position a) between the two rate conditions for each consonant and b) between the two consonants for each rate condition. These differences were estimated as follows. For each target word (occurring in both speaking rate conditions), the interval of the post-vocalic consonant was determined manually from the acoustical data. For CVNV words, the left boundary referred to the point where changes in the F2 and F3 amplitudes were clearly visible during the transitions from the vowel into the nasal murmur. The right boundary referred to the point of transition into the following vowel. In the CVCV words, the consonant was defined as the interval between the transition point of the vowel into the silence gap typical for oral stops and the point of stop release, omitting the aspirated portion. Next, for these intervals, the maximum signal values were determined from the alveolar region by means of the VTA function, while the maximum velum signal in /n/ and the minimum velum signal in /t/ were captured by PCA modelling. To generate the differences displayed in figs. 4.7 and 4.8, a new variable *maxmin* was created that included the maximum velum signals for the nasal stop and the minimum velum signals for the oral stop. Subsequently, the means of the maxmin velum signals and maximum alveolar signals values were calculated for each speaker, such that for both gestures, subtractions were possible for the two stops (/t/–/n/) separated by the rate condition or for the two conditions (fast–moderate) separated by the stop. Comments on the logic of comparing the maximum /n/ values with the minimum /t/ values of the velum signal are given further below.

Statistical analyses were carried out in the programming environment *RStudio* (version 1.2.5033) by applying linear mixed models with the *lmer* function from the *lmerTest* package. To test the hypotheses of this section, the models were separated by vowel tensity. This is because the data are not entirely balanced with respect to the consonantal context and vowel tensity (CVNV lacking lax /u/, CVCV lacking tense and lax /y/)[4]. Based on a preliminary mixed model suggesting an interaction between the consonant and vowel tensity, the data were considered separately for tense and lax vowels.

Two models were applied for each tensity condition, one considering the effect of the consonant and speaking rate on the tongue position and one focusing on the velum gesture affected by these parameters. As the variable *maxmin* comprises the maximum velum signals for /n/ and the minimum velum signals for /t/, a comparison between these values would by definition lead to the redundant finding that the velum exhibits higher values (i.e. a lower position) in /n/. However, the focus is not on the direct comparison of the velum position in /n/ versus /t/ but rather on the interaction effects between the speaking rate and the consonant: it is investigated whether the maximum distance between /n/ and /t/ is significantly greater in one rate condition than in the other and also whether the deviation between the fast and moderate speaking rate is greater for one specific stop. With respect to the tongue tip position, in contrast, a direct comparison between /t/ and /n/ is possible because both stops are produced with a similar tongue tip constellation, which allows for considering the same parameter, i.e. the maximum alveolar signal.

In summary, the two models consider the dependent variables of a) the maxmin velum signal to determine the interaction effects of the consonants and the speaking rate on the velum position and b) the maximum alveolar signal to examine the impact of the consonant and speaking rate (and their interactions) on the tongue position. For all models tested, the consonant (/n/, /t/) and the speaking rate (fast, moderate) were selected as the fixed effects; the random effects included the word stem (word onset plus vowel) and the speaker.

Finally, in addition to the spatial data on the velum and tongue position, alterations in the velocity of the velum are examined by comparing the maximum velocity during the velum onset gesture in moderate speech with that in fast speech, for which only CVNV items are considered. The velocity parameter was derived from the basic velum signal estimated by the PCA and refers to the alterations of the velum signal per second. The velocity value used in this analysis indicates the (normalized) signal alteration per second at the time point of the maximum velocity during the velum opening gesture determined by kinematic analyses (section 1.6.4). The closing gesture was not considered because less consistency was expected with respect to the timing and velocity during velar closure. This

[4]As mentioned in the introduction, the original overall corpus was designed to cover a broad spectrum of various questions on vowel nasalization in different contexts in natural words. For this reason, the data considered are not perfectly balanced.

is due to the fact that the nasal stop in the current data was generally followed by an unstressed vowel, which allowed for a relatively large space for the velum closing gesture and thus for less methodological control across speakers. For statistical analysis, the velocity was defined as the dependent variable, with the speaking rate as the fixed effect and the word stem and speaker as the random effects.

Results

The effects of the speaking rate and the post-vocalic stop on the tongue and velum position are delineated on the basis of figs. 4.7 and 4.8. For tense and lax vowels separately, the positional differences between the nasal and oral stop are considered and the gestural alterations between fast and moderate speech are explicated, followed by statistical analyses.

Difference oral-nasal Figure 4.7 depicts the maximum distance of the velum position between the oral and the nasal stop along the x-axis. By definition, the maximum velum signal in /n/ is higher than the minimum signal in /t/[5], which is indicated by the negative values. However, the focus is on the question of whether the maximum distances between the stops are greater in one speech rate condition than in the other, reconsidering the prediction that during fast speech, the distance is expected to be decreased for the velum. Figure 4.7, however, suggests no systematic differences in the distribution of the values for fast and moderate speech.

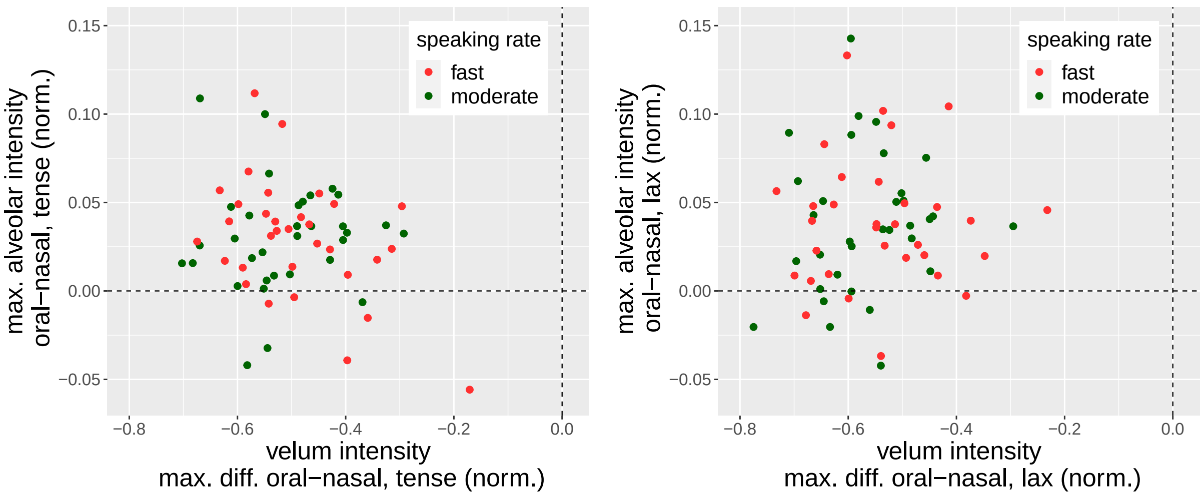


Figure 4.7: Maximum differences in signal intensity between /t/ and /n/ uttered with fast versus moderate speech. Results are given for the velar region (x-axis) and the alveolar region (y-axis). Left: stops preceded by tense vowels; right: stops preceded by lax vowels. Normalized mean values are indicated for the oral–nasal difference, separated by the speaker.

The differences of the tongue tip position between /t/ and /n/ are depicted along the y-axis. As the same parameter was considered for both stops (i.e. the maximum alveolar

[5]See the appendix B.1.2 for a comparison of the maximum velum signal for both /n/ and /t/, demonstrating velum position differences when the same parameter is considered.

signal), a direct comparison of the tongue position between the oral and the nasal stop is possible. A clear tendency is apparent for a higher tongue during the oral stop, indicated by the predominant amount of positive values. As with the velum distance patterns, a comparison of the tongue position differences with respect to the two rate conditions suggests that the speaking rate does not have any significant impact on the maximum alveolar signal. This pattern is reflected by both vowel tensity conditions.

Difference fast-moderate The gestural position differences between fast and moderate speech are displayed in fig. 4.8 for each stop individually; the data are again separated by vowel tensity. Clearly, the positional differences in fast versus moderate speech are much less distinct than the differences between the individual stops just outlined, which is expressed by the value distribution close to the center of the coordinate system. For both stops and both tensity conditions, alterations in velum height do not seem to systematically vary in favor of one rate condition. If the velum were regularly lower in moderate speech during /n/ (exhibiting higher signal values), a shift towards the negative area would be apparent. Similarly, for /t/, more positive values would be displayed if the velum were raised to a higher extent in the moderate condition (as a higher position is reflected by lower signal values). Although no systematic pattern is evident in general, a small tendency is observable for the nasal targets when preceded by lax vowels (fig. 4.8 right, light values): with moderate speech, the velum appears to exhibit slightly higher signal values. Comparing the position of the tongue tip during fast and moderate speech, both stops in both tensity conditions exhibit a slight tendency towards more negative difference values, reflecting higher signal values and thus a higher tongue position in the moderate rate.

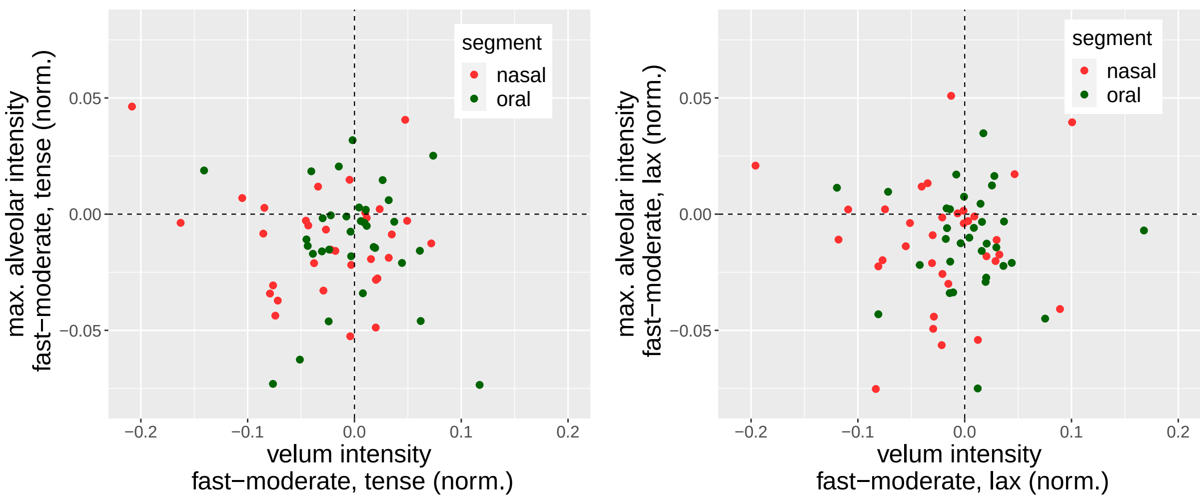


Figure 4.8: X-axis: differences of the max. (nasal) or min. (oral) velum signal intensities between fast and moderate speech in the post-vocalic stops. Y-axis: differences of the maximum alveolar signal intensities between fast and moderate speech in the post-vocalic stops. Left: stops preceded by tense vowels; right: stops preceded by lax vowels. Normalized mean values are given for the fast–moderate differences, separated by the speaker.

Moreover, the question is considered of whether the differences between the gestural positions are comparable for both stops, i.e. whether the velum and the tongue show deviations induced by the speaking rate for both stops to the same extent. Considering the nasal and oral difference values displayed in fig. 4.8, it can be observed that the oral differences tend to be located closer to the center than the nasal differences, suggesting that during the oral targets, the gestures are less affected by the speaking rate and thus yield smaller difference values than during the nasal stop. This is slightly more evident for the velum gesture than for the alveolar signal in both tensity conditions.

Statistical analysis

Separate models were run for the tense and lax vowel contexts. The maxmin velum parameter and the maximum alveolar signal were selected as the dependent variable, respectively, and the consonant and rate condition were defined as the fixed effects. The word stem and the speaker were selected as the random effects.

Tense vowels Considering the velum position with tense vowels being involved, results suggest a main effect of the consonant ($F[1,33]=837.65$, $p<0.001$) as well as an interaction between the consonant and the speaking rate ($F[1,1298]=7.17$, $p<0.01$). Post-hoc corrected bonferroni tests revealed a significant effect of the rate for /n/ ($p<0.05$) but not for /t/. Moreover, no significant difference was reported with respect to the velar deviation in /n/ versus /t/ in fast versus moderate speech, suggesting a similar amount of gestural deviation between the stops in both rate conditions.

With respect to the tongue tip position, results suggested a main effect of both the consonant and the speaking rate ($F[1,42]=94.02$, $p<0.001$ and $F[1,1332]=12.07$, $p<0.001$), with a higher position during /t/ and during moderate speech. No interaction was reported, indicating that the distance between /n/ and /t/ was not found to be distinct for fast and moderate speech and similarly, the tongue tip distance between fast and moderate /n/ was not found to be significantly different from that in fast versus moderate /t/.

Lax vowels With respect to the velum behavior with lax vowels being involved, results reported a main effect of the consonant ($F[1,35]=871.83$, $p<0.001$) as well as an interaction between the consonant and the rate ($F[1,955]=11.53$, $p<0.001$). While the rate condition had a significant effect on /n/ ($p<0.05$), suggesting a lower velum in moderate speech, no effect was reported for /t/. As with the data on tense vowels, no further significant difference was indicated with respect to the distance between /n/ and /t/ in fast versus moderate speech, suggesting that the velum gesture showed a similar displacement between these stops in both speaking rates.

Findings on the tongue position reflected those from the tense condition: a main effect was suggested for the consonant ($F[1,43]=24.54$, $p<0.001$) and the rate condition

(F[1,32]=9.41, $p<0.01$), indicating a higher tongue position in the oral stop as well as in moderate speech, with no further interaction reported.

In summary, the patterns found were quite similar in both vowel tensity contexts: the tongue tip clearly exhibited a higher position in moderate speech and was also found higher in the oral stop. The tongue position differences between the two stops were comparable for both speaking rates; similarly, the position differences between fast and moderate speech were comparable for both stops. With respect to the velum gesture, speaking rate had an effect on the velum position when /n/ was involved, with a significantly lower position in moderate speech compared to fast speech. Alterations of this kind were not observed for the oral stop. Moreover, no differences were reported for the maximum distances between /n/ and /t/ in the two rates.

Velocity during nasal stops in moderate and fast speech

The results above comprised data referring to the spatial aspect, i.e. to the position of the tongue and velum in moderate and fast speech. Findings showed that the gestural positions were affected by both the speaking rate and the nature of the consonant. Following prior studies concerned with alterations of both the amplitude and velocity of the velar gesture, the data at hand were additionally examined for speaking rate effects on velocity patterns. As the velum generally shows clear alterations in its position especially during sequences containing a nasal stop, the analysis solely refers to CVNV sequences. The velocity parameter was derived from the basic velum signal and is indicated by normalized positional change per second. The data capture the maximum velocity during the onset of velum lowering. The closing gesture was not considered because less consistency was expected with respect to the timing and velocity during velar closure due to the nature of the post-nasal vocalic context.

Differences in the velocity and velum position during the nasal stop are shown in fig. 4.9. As just outlined, the velum exhibits a significantly lower position in moderate speech compared to fast speech during the nasal stop. In contrast, the maximum velocity during the opening gesture appears to be unaffected: in both tense and lax contexts, no tendency is evident towards one rate condition. The impression received from fig. 4.9 is compatible with the statistical reports: with speaking rate as the fixed effect, velocity was not reported to be significantly affected (tense vowels: F[1,32]= 0.25, $p=0.619$; lax vowels: F[1,362]=0.31, $p=0.575$).

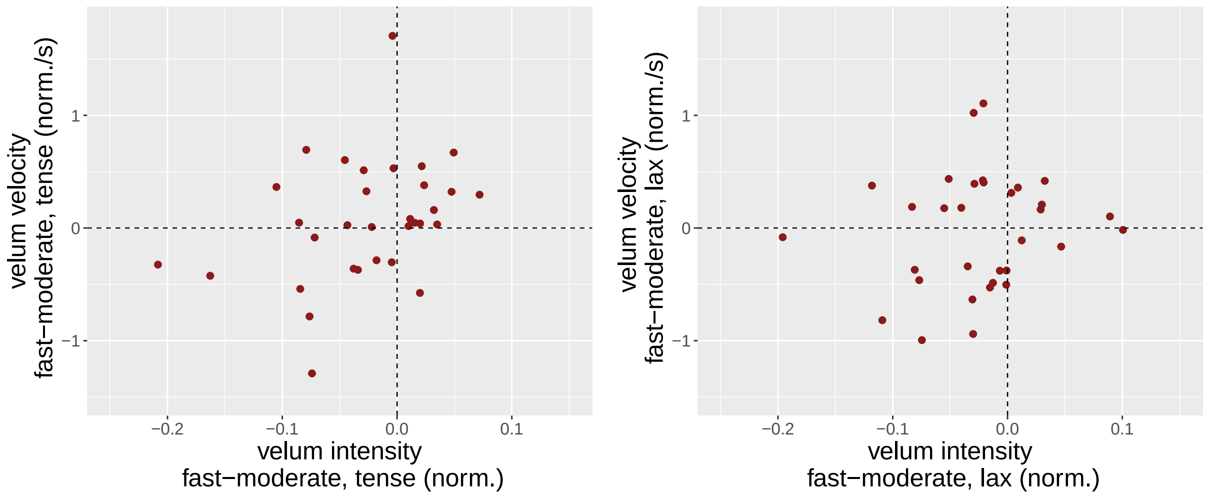


Figure 4.9: Differences in the maximum velocity during velum opening (y-axis) as a function of the maximum velum signal captured during /n/ (x-axis). Normalized mean values are plotted for the fast–moderate differences, separated by the speaker. Left: /n/ preceded by tense vowels; right: /n/ preceded by lax vowels.

4.3 Summary and discussion

The results reported in this chapter can be summarized and related to the previously stated hypotheses as follows. In general, the tongue tip was found lower in /n/ than in /t/ in both speaking rate conditions, which is compatible with H1. Considering the tongue movement patterns in fast versus moderate speech, findings generally suggested a higher tongue tip in the moderate condition speech for both stops, which is in agreement with H2. Otherwise, no interaction effects were reported, indicating that the tongue tip deviations between the nasal and oral stop were similar in both speaking rates. Likewise, the lingual displacement between moderate and fast /n/ was comparable to that for moderate and fast /t/. Accordingly, these findings are not compatible with the expectations expressed in H3 and H4. With respect to the velum position affected by the speaking rate, prediction H5 is in accordance with the findings: during /n/, the velum exhibited a lower position in moderate speech, whereas no significant differences were found for the velum position during /t/, which is not compatible with H6 but is in agreement with H7: the speaking rate significantly affected the velum deviation during /n/, but not during /t/. Moreover, the positional deviations between the oral and nasal stop were found to be similar in both speaking rates, contradicting H8. Finally, results on the maximum velocity during velum opening did not reveal any significant differences between fast and moderate speech, which is in support of H9.

Evidently, some of the findings are in agreement with reports from prior studies. For example, the tongue tip gesture in our data was found to be spatially reduced in /n/ relative to /t/ in both fast and moderate speech (see also Jaeger and Hoole (2011) for the effect of word frequency on tongue and velum position in /n/ vs. /t/). Similarly, in the fast rate condition, an overall amplitude reduction of the gestures was evident, except for the velum gesture during /t/. Thus in these data, the speaking rate had a stronger effect on the velum when a larger amplitude was involved, as was the case during /n/. Apart from this, however, the gestural movement patterns turned out as less distinct than initially expected. In particular, it is somewhat surprising that the tongue tip displacement between moderate /n/ and fast /n/ did not significantly differ from the range of the oral stop. If /n/ tolerates some disregard of lingual closure that is even enhanced with fast speech, but /t/ does not (at least not to the same extent), the difference between the two positions in moderate and fast speech is expected to differ for /n/ and /t/. However, results from the current data suggest that this was not the case. Similarly, the spatial distance between the oral and nasal stop did not differ for the two speaking rate conditions, which was initially expected to increase with fast speech. Thus, the reduction of the tongue tip gesture in /n/ in fast speech was less pronounced than expected.

Interestingly, the velocity during velum lowering was not increased with fast speech, while at the same time, the velum showed reduced lowering patterns. This is compatible with the assumptions about speakers' articulatory strategies during fast speech suggested in prior studies (Kent et al., 1974; Kuehn, 1976). In those experiments, some speakers showed increased velocity but a stable amplitude of velum lowering in fast speech, while for others, velocity was unaffected, whereas the velum gesture exhibited a reduced amplitude. The data at hand, obtained from a relatively large number of participants, generally suggest that velocity does not significantly change with different speaking rates, but that the articulatory events in fast speech are achieved by a general reduction of the lowering amplitude.

Overall, findings allow for considerations about some conceivable consequences for coarticulatory vowel nasalization. Velum lowering during the nasal was accompanied by a reduced tongue tip closing gesture relative to /t/ in moderate speech (fig. 4.7). With fast speech, amplitude reduction during /n/ was enhanced for both gestures: the velum exhibited less lowering and the tongue tip showed a lower extent of raising (fig. 4.8). Under the assumption that coarticulatory vowel nasalization is enhanced with a consistently lowered velum and a concurrent incomplete oral closing gesture, the patterns found for fast speech effectively do not suggest that nasal coarticulation is enlarged with an increased speaking rate: a higher velum position in fast speech is rather compatible with less instead of more nasality. This conclusion, however, may be drawn with respect to the spatial aspect. Considering the temporal extent of vowel nasalization, nasality might still be extended in fast speech, just because the tongue tip gesture is reduced during fast speech and the velum is lowered for the nasal stop, albeit not as strongly as in moderate speech. This combination may result in a less well defined boundary between the pre-nasal vowel and the nasal segment, such that the segments become more similar to each other, which is otherwise prevented by a more distinct closing gesture of the tongue tip in moderate speech.

Chapter V

Vowel nasalization in production and perception

Abstract

This chapter investigates how German listeners perceive and interpret vowel nasalization in synthesized acoustic stimuli. In this experiment, two nasality conditions are presented to the listeners: one with a constant nasalized interval that is shifted as a whole across the vocalic and nasal segments of the target word, and one in which the temporal extent of the nasalized interval is altered within the vowel. Results indicate that German listeners have less difficulty in determining the target word when the overall nasalized portion is shifted rather than temporally modified. This finding is discussed in terms of considerations about other factors contributing to this effect independent of vowel nasalization. Moreover, the hypothesis is tested whether a correlation is evident between individual speakers' coarticulatory patterns in production and their specific sensitivity to fine acoustic details in perception. Previous studies in this field have shown tendencies towards such an effect, namely that speakers with extensive anticipatory vowel nasalization were more likely to make use of acoustic cues, for example to prematurely predict the upcoming segment. In contrast to these findings, however, no such correlation effect is evident in our data.

5.1 Vowel nasalization in production and perception

5.1.1 Introduction

Perception in phonetic research is concerned with the interpretation and processing of heard instances on the part of the listener. While the content of this thesis so far has focused on production data, the current chapter investigates whether German listeners are aware of fine acoustic details in terms of nasalization. Acoustically, the spectrum of nasalized vowels constitutes a complex interplay of resonant frequencies affected by the pharyngeal, oral and nasal cavities, which will be roughly sketched in the following.

Considering the acoustic characteristics of a plain nasal stop, the first formant (F1) generally shows low frequencies due to a longer vocal tract passage. In addition, F1 exhibits an increased bandwidth, i.e. increased width of the resonant peak, which is due to the fact that by coupling the nasal cavity to the pharyngeal and oral cavities, the soft wall surfaces within the vocal tract are enlarged, while at the same time the vocal tract is more restricted at the nostrils compared to mouth opening. These factors cause an overall increase in damping of the sound energy, resulting in a decreased amplitude (Stevens, 2000).

Moreover, nasal stops are produced with an open velopharyngeal port but also with oral closure, such that the oral cavity functions as a side branch of the overall resonant tube. Thus, frequency components similar to those resonance frequencies that are characteristic for the side branch do not appear in the acoustic output, because they are absorbed in the side branch. Which of the specific frequencies are affected depends on the particular place of oral closure. As a result, characteristic low intensities occur in the respective frequency regions during nasal murmurs. Therefore, in contrast to formants, which constitute peaks in the acoustic energy, these anti-formants conversely appear as valleys in the spectrogram with low energy and additionally cause further reductions of those formant amplitudes which are above the anti-formants.

In the case of nasalized vowels, however, the pharyngeal, nasal and oral cavities are coupled as with the nasal stops, but the sound energy may escape both through the mouth and the nostrils. This causes some highly complex spectral modifications due to the combination of the acoustic characteristics of both oral and nasal sounds. Basically, in contrast to one distinct formant F1 in nonnasalized vowels, there are three spectral prominences in the region of F1 interacting in nasalized vowels, including an oral formant, a nasal formant and an anti-formant. While the nasal formant shows constant low frequencies, the frequency of the anti-formant in nasalized vowels is related to the degree of velopharyngeal port opening, with its frequency increasing with larger opening (Maeda, 1993). Depending on the degree of velum opening, the anti-formant will thus occur at different frequencies such that it may cancel out one of the other formants to a certain extent. Accordingly, in the case of a heavily nasalized low vowel a decrease of F1 will occur compared to a nonnasal vowel, because the high anti-formant cancels out the high oral formant, allocating the spectral dominance to the lower nasal formant (Feng and Castelli, 1996; Fujimura and Lindqvist, 1971; Maeda, 1993; Serrurier and Badin, 2008). Correspondingly, high nasal vowels involve a low first formant, which is affected not by the high anti-formant but by the nearby low nasal formant causing an overall slight increase of F1. In summary, nasalized vowels are characterized by an increased bandwidth of F1 as well as a decreased amplitude due to the presence of an anti-formant. In addition, spectral changes are introduced by the presence of a nasal formant and anti-formant insofar as F1 is decreased in low vowels while it is increased in high vowels.

Much research has been done on the question of how listeners perceive these specific spectral modifications of nasal sounds. While some studies addressed the direct relationship between F1 shifting and perceived nasalization (e.g. Kingston and Macmillan, 1995; Krakow et al., 1988; Wright, 1975), others focused on related questions such as how much nasal coupling was required for a specific vowel to be perceived as nasalized (House and Stevens, 1956; Krakow, 1993; Whalen and Beddor, 1989). Another issue concerns perceived vowel nasalization as a coarticulatory effect, i.e. if and how listeners use nasal cues in the vowel to predict the contextual environment and whether individual listeners' coarticulatory patterns in production are also reflected in perception. The two latter aspects are the focus of the following review on perception studies as well as of our own experiment outlined in this chapter.

Prior research on the perception of nasalization

In an early experiment, House and Stevens (1956) created synthetic sounds by using an electrical vowel and nasal analog via an acoustic tube and vocal tract model. The stimuli consisted of the vowels /i, ɛ, ɑ, ɔ, u/ and each of these were provided with four different degrees of coupling to the nasal tract analog. In a first test, 34 American English listeners were presented with pairs of identical stimuli and asked to judge each pair as either characterized by nasality or as containing non-nasal vowel-like sounds. As a result, listeners required much less nasal coupling in pairs with /i/ and /u/ vowels to judge them as nasalized compared to the low vowel /ɑ/. This finding was confirmed in a second experiment, in which the pairs consisted of vowels of the same categories but different nasal coupling. Participants were instructed to indicate whether the first or second item of a pair sounded more nasal or whether there was no difference in nasality. Again, each vowel was found to have its own coupling range: for /i/ and /u/, responses on nasality identification started earlier with only little coupling present, followed by pairs of /ɛ/ and /ɔ/. The highest degree of nasal coupling was required for /ɑ/ to be perceived as nasal. The authors noted that their data were based on synthetic rather than natural speech, but referred to previous research on this issue and postulated that this fact "should not obviate the principle conclusion" (House and Stevens, 1956, p. 230).

Other studies were concerned with the question of whether speakers of different languages with and without contrastive nasalization showed different perception thresholds when judging a vowel as nasal, and whether there were specific acoustic properties related to the 50% crossover points from non-nasal to nasal judgements in listeners' decisions. For example, Hawkins and Stevens (1985) generated synthetic stimuli of non-nasal /tV/ syllables and introduced pairs of formants and anti-formants ('poles' and 'zeros') to the spectra. This was done by interpolating between the values of the frequencies of the poles and zeros and F1 for the non-nasal and nasal extremes, such that for each of the five vowels tested, continua of nine steps were created. These stimuli were presented to Hindi, Gujarati, Bengali and American English listeners who were told to judge each stimulus as either nasal or non-nasal. Results were quite similar across the language groups, as listeners' crossover points for judging a stimuli as nasal occurred at similar levels of frequency modifications. Interestingly, these crossover points involved frequency spaces in the range of 75-110 Hz across the vowels, that is, listeners started to rate stimuli as nasal when these provided pole-zero differences within this range. This lead the authors to the conclusion that listeners relied on certain acoustic properties rather than on linguistic experience to judge a vowel as nasal. In a second experiment, participants' performances of discriminating stimuli pairs with pairs differing in one and two steps on the nine step continuum were tested, such that some pairs were close to the 50% crossover points of the continuum (e.g. stimuli 5 vs. 6) and others were located more at the oral and nasal extreme ends (e.g. 1 vs. 2 or 8 vs. 9). A general effect was found for the location within the continuum, with more correct responses for pairs in the vicinity of the crossover points rather than the extremes. In addition, results showed differences for the two language families: while the Indian participants were consistently more accurate in discriminating pairs in the vicinity of the crossover point, the American listeners did not show such consistency and had more variable response patterns depending on the vowel. From the overall results, the question was discussed of whether there was one particular acoustic property leading the listeners to rate a vowel as nasal. The authors pointed to the fact that the introduction of a pole-zero pair with a certain spacing would cause a specific maximum deviation in decibel compared to the spectrum of a non-nasal vowel. The closer the pair was to F1, the more salient the deviation, such that a pair which was located further away from F1 would also require more spacing to elicit a nasal response. By this, the maximum deviation in dB from the spectra, caused by the spacing, was suggested to be related to listeners' decisions. However, the authors raised concerns that this strategy would presuppose that a listener would have to compare the perceived spectrum with a memory of a non-nasal vowel spectrum, which they rejected. As an alternative, they suggested that by introducing a pole-zero pair, the prominence of the spectral peak of F1 was reduced due to the additional peak in the low-frequency region, which had a similar effect to broadening the bandwidth of F1, without adding a pole-zero pair.

Similar findings were reported by Stevens et al. (1987), who ran comparable experiments in which pole-zero pairs with varying spacing were introduced into the spectra of synthetic non-nasal vowels. Here, the crossover points occurred when the maximum deviation was in the range of 6-9 dB relative to the non-nasal spectrum. The deviation was achieved by either adding an extra peak near F1, by decreasing the amplitude of the F1 peak or both. As with the data presented by Hawkins and Stevens (1985) above, the spectral deviation for non-low vowels was above F1 or centered and for the low vowel /a/ it was below F1. These findings were further considered by Stevens (1989), who interpreted the results as evidence for a non-monotonic relation between acoustic and auditory parameters in the sense that at some specific point, small acoustic modifications cause large auditory changes (and likewise, small articulatory modifications cause large acoustic changes), such that "the auditory response shifts from one type pattern to another." (Stevens, 1989, p. 4).

Rather than examining perceived vowel nasalization as a function of frequency modifications, Malécot (1960) investigated vowel nasality as a distinctive feature in American English. To test the hypothesis that for native speakers of American English the cue of vowel nasalization was sufficient to perceive a following nasal consonant, the author manipulated pre-recorded word pairs such as *cap–camp* for a tape-cutting experiment by extracting and recombining the segmental parts. An oral vowel in e.g. *cap* was substituted with the nasalized vowel from *camp*. Along with the original oral items, the stimuli were presented to 25 participants who were asked to specify on a sheet of paper the word they had heard, with all items possible. In 95% of those responses referring to the manipulated stimuli with a nasalized vowel but omitted nasal, the participants marked the words which contained a written nasal. Next, the author tested the prediction that the presence or absence of a nasal segment was an important cue for the voicing of the following stop. To examine this, the word pairs *amble–ample*, *candor–cantor* and *anger–anchor* were edited such that the nasal in words containing a voiced stop was removed. For those words with a voiceless stop, separately recorded nasals were inserted before the stop. Along with the unaltered variants of each word, the stimuli were presented to the same participants who had taken part in the first experiment. As expected, the participants judged those items without the nasal segment as the voiceless variant, and conversely the items with inserted nasals as words with a voiced stop. Based on these results, Malécot concluded that

> "in the limited conditions present in camp, hint, bunk, and the like, the nasal consonants as separate segments have virtually disappeared, possibly because of the silent interval necessary as a cue for the voicelessness of the following stop, with the result that now, in both manner dimensions, the perceived [m], [n], or [n] is conveyed almost entirely by cues contained in the preceding vowel. [...] Vowel nasality is thus the principle distinctive feature in the cases of O/N in question." (Malécot, 1960, p. 228f).

Similarly, Ali et al. (1971) ran a perception experiment to investigate listeners' ability to predict a post-nasal consonant solely by means of the acoustic cues in the preceding vowel. Natural stimuli of CVC and CVVC were edited such that the final consonant (either nasal or oral) as well as the vowel-consonant transitions were eliminated. The stimuli were presented to 22 American English listeners who were asked to determine whether the upcoming consonant was nasal or oral. Participants were largely correct in their predictions of the post-vocalic consonant based on the acoustics of the vowel. Interestingly, the consonant was erroneously predicted to be nasal more often for oral stops than fricatives. Moreover, vowel quality was another factor insofar as consonants following the low vowel /a/ were more often predicted to be nasal than after /i/, /eɪ/ and /u/.

Further findings on the predictability of an upcoming consonant were provided by Ohala and Ohala (1995), who ran a replication experiment following Lahiri and Marslen-Wilson (1991, 1992). For this, they modified several aspects of the original procedure and created stimuli pairs of existing words from English and Hindi, a language with contrastive nasalization. The Hindi words contained doublets of CVC and CVN sequences as well as triplets of CVC, CṼC and CVN sequences. For English, CVC and CVN sequences were tested. The stimuli were edited by gating out both the final consonant and parts of the vowel at different points. The participants, 39 Hindi and 44 English listeners, were presented with single word trials and were asked to indicate the heard word via forced choice responses: the English listeners selected either CVC or CVN, while for Hindi, alternating responses with doublets of CVC, CVN and triplets of CVC, CṼC and CVN were provided. The results for the English listeners clearly showed largely correct responses for both the CVC and CVN conditions: for words with the final consonant gated out, listeners relied on the vowel for making their decision. For the Hindi listeners, however, responses differed such that when a CVN stimulus word with an omitted nasal was to be assigned to a triplet option, this word was by majority identified as CṼC instead of CVN. When only two options were given, the CVN stimulus was unexpectedly equally assigned to the CVC and CVN option. In contrast, when the CVC stimuli were to be identified, the majority responses were correct in both the triplets and doublets options. For the CṼC condition, responses were again correct a majority of the time. The authors assumed that the response confusion involved with the Hindi CVN stimuli was caused by the way perception was processed in analogy to other cognitive processes: "Apparently, in perception, the default strategy is to pay most attention to what is present rather than what is absent but might have been present." (Ohala and Ohala, 1995, p. 56). In this sense, the Hindi listeners chose the CṼC answer because the fragment CṼ omitted crucial transition cues during the release of a nasal vowel into a nasal consonant. That is, when listeners heard CṼ without the relevant transition cues, they associated the fragment with a word form which offered the greatest resemblance.

More research on vowel nasalization as a predictive cue was provided by Beddor and Krakow (1999), who investigated the perceptual compensation of vowel nasality in nasal contexts for English and Thai, following the hypothesis for compensation considered by e.g. Fowler (1996); Fowler and Smith (1986); Krakow et al. (1988) and Ohala (1993). According to this hypothesis, listeners compensate for coarticulatory effects if they detect the source for this effect and assign the effect to the source rather than to the actual segment. Thus, in the case of compensation for coarticulatory vowel nasalization, nasalized vowels are predicted to be rated as less nasalized when surrounded by nasal consonants to which the effect of nasalization is assigned. If such a source is lacking, however, the vowel is expected to be perceived as nasalized. To test this assumption, Beddor and Krakow (1999) ran two perception studies with one group of English listeners and one group of Thai listeners. In the first experiment, a discrimination test was constructed by cross-splicing oral and nasalized vowels from the appropriate consonantal context to an inappropriate context, and in addition, the vowels were isolated. The manipulated stimuli were presented as trials of two pairs, with one pair differing in vowel nasality. Listeners were asked to select the pair that involved differing vowels. For both language groups, responses were least correct when the pair in question exhibited vowels in nasal contexts, especially nasal vowels in nasal contexts: listeners more often judged those vowels as different which acoustically had the same nasality but occurred in one appropriate and one inappropriate consonantal context (e.g. $\mathrm{N\tilde{V}N}$–$\mathrm{C\tilde{V}C}$), suggesting that a nasal vowel in nasal context was rather perceived as oral. Conversely, acoustically different vowels were erroneously judged as the same when surrounded by an appropriate condition (e.g. $\mathrm{N\tilde{V}N}$–CVC).

Building on this first test, the authors investigated listeners' decisions in rating the nasality of two words in relation to each other. The stimuli used in this experiment were the same as for the first one and were presented in single pairs. Again, most difficulties in discriminating occurred when a nasalized vowel was surrounded by nasal context. Listeners' responses were largely accurate when the rating concerned a nasalized vowel in isolation or in oral context relative to an oral vowel, but when placed within a nasal context, participants had difficulties in discriminating a nasalized vowel in nasal context from an oral vowel, independent of the oral vowel's context. Considering language-specific differences in the usage of anticipatory nasalization, another hypothesis was postulated: due to less experience with extensive coarticulatory nasalization, the Thai speakers were expected to compensate less for the perceived effects and thus to show better performances in rating a nasal vowel in nasal context. Besides large parallels for the results of the two language groups, several differences also became apparent. The Thai listeners, just like the English group, were overall less accurate in their discrimination responses when the nasal vowel was embedded in nasal contexts compared to oral contexts. However, Thai listeners were better at discriminating comparisons when the isolated matched vowel was oral rather than nasal, such that the pairing of NVN–$\tilde{\mathrm{V}}$ $\sim$ NVN–V was discriminated more accurately than that of NVN–$\tilde{\mathrm{V}}$ $\sim$ $\mathrm{N\tilde{V}N}$–$\tilde{\mathrm{V}}$. From this, it seemed that Thai listeners were even more instead of less likely to compensate for vowel nasalization, since they showed even greater difficulties than the English listeners in rating the nasal vowels in nasal context equal to nasal vowels in isolation. As an alternative interpretation to the compensatory explanation, the authors proposed that since Thai listeners were overall poorer in discriminating nasal vowels than oral ones, this might point to a more general language-specific issue with distinguishing nasal vowels compared to the English listeners.

For the rating experiment, the Thai listeners showed similar response patterns compared to the English listeners, again with poor performances in judging the nasality of a nasal vowel in nasal context. However, unlike the English participants and unlike the findings from the discrimination task, the Thai listeners were less likely to rate the nasal vowel in isolation (Ṽ) or in CṼC as more nasal than in NṼN, indicating less compensation. Conversely, in pairs with nasal and oral vowels in appropriate contexts, the Thai participants rated the vowels less frequently as the same than the English listeners.

The authors concluded that their hypothesis was largely confirmed in that listeners' performances in both language groups were poor when judging nasal vowels in nasal contexts. As they pointed out, though, the generalization was not legitimate that listeners perceived these vowels just as oral as an oral vowel in the appropriate context, because responses were not consistent enough to allow for this conclusion. Instead, the authors suggested that

> "the overall results are more compatible with the interpretation that these listeners were *partially* compensating for contextual nasalization, that is, that listeners were attributing some, but not all, of coarticulatory vowel nasalization to the nasal consonant context." (Beddor and Krakow, 1999, p. 2884).

Similar to this study, Fowler and Brown (2000) examined whether listeners assigned nasality directly to the vowel or perceived it as the onset of the following nasal. In a first experiment, stimuli of naturally produced CVCV and CVNC sequences were created with the first vowels being spliced and cross-spliced across the oral and nasal contexts. Responses from 18 American English listeners were captured by measuring reaction times in an online experiment in which participants were to press the correct key on a keyboard for the respective context as fast and as accurately as possible. As expected, response times were significantly slower in the cross-spliced conditions for both nasal and oral conditions, when listeners were confronted with misleading cues in the vowel relative to the following consonant.

In the follow-up experiment, the question was addressed which particular factors contributed to the slower reaction times. In this task, listeners heard trials of stimuli triplets with manipulated nasality degrees of the first vowels and were asked to rate the middle target relative to one of the flanking targets with respect to its vowel. Results showed that listeners were more likely to rate a nasalized vowel followed by a nasal consonant more similar to a less nasalized vowel followed by an oral consonant. Likewise, when judging a nasalized vowel in oral condition, this vowel was found to be more similar to a more nasalized vowel followed by a nasal consonant. The authors interpreted this result as consistent with the hypothesis that listeners assigned the perceived vowel nasality to the nasal rather than to the vowel. However, this finding was opposed to the results of a third experiment, in which the authors investigated to what extent listeners perceived and assigned vowel nasality if they heard naturally produced word pairs with the vowel providing nasality at its endpoints, depending on the following context, rather than manipulated steps. Stimuli pairs were edited with varying conditions of vowel nasality and appropriate or inappropriate consonantal contexts. Listeners were told to judge whether the vowels in each stimulus pair were the same or different and to give their response as quickly as possible by pressing the respective key. This time, listeners responded most rapidly when both the nasality condition and the appropriate context were the same, slower when the nasality type was held constant but the context varied, and slowest when the vowels differed in nasality but were followed by the appropriate consonantal contexts. According to this ranking, a nasalized vowel followed by a nasal was perceptually closer to a nasalized vowel in oral context than to an oral vowel in oral context. The authors therefore concluded that parsing of vowel nasality was incomplete: listeners did not completely assign the nasal information from the vowel to the following nasal and hence did not merely perceive that vowel as oral. While the second experiment with synthesized degrees of nasality showed tendencies towards such an account, response times in the third experiment with natural stimuli were lowest when two vowels with different nasality occurred in appropriate contexts. This indicated that a nasal vowel in a nasal context was not simply interpreted as oral, contradicting the hypothesis that parsing of vowel nasalization was complete.

Later research of Beddor and colleagues (e.g. Beddor, 2007, 2009, 2015) focused on the interplay of coarticulation in production and perception. The authors related their findings to the more general question of which stages need to be passed for coarticulatory vowel nasality to become contrastive in a language. As the basic idea, Beddor considers the temporal extent of overall nasalization as a roughly constant gesture across the vowel and the nasal in production, which is expected to have consequences for perceptual processing. Some aspects were already outlined in chapter III, illuminating the effect of vowel length on perceived vowel nasalization in her experiments, which will not be repeated here in detail (Beddor, 2007, 2009; Whalen and Beddor, 1989). The concept of the trading relation hypothesis, however, is examined in terms of the perceptional aspect as well as further findings on the impact of the post-nasal context on perceived vowel nasalization. The basic assumptions on the trade-off relation are recapitulated below.

The idea of an inverse relation between vowel nasalization and nasal duration implies that the velum lowering gesture is held roughly constant in size but may be shifted along the vowel-nasal sequence, such that a strongly nasalized vowel is followed by a short nasal and a weakly nasalized vowel precedes a longer nasal segment. This idea has arisen from the finding that the consonantal voicing of a post-nasal stop or fricative affects the extent of vowel nasalization: in the context of a voiceless obstruent, the vowel is nasalized to a greater extent and the nasal is shorter in duration, whereas preceding a voiced stop, the vowel is less nasalized but the nasal shows longer duration (Beddor, 2009, among others). Based on this observation, the hypothesis was formulated that depending on the following context, nasalization of the vowel is not increased per se, but that the velum lowering gesture is shifted to an earlier onset point relative to the raising of the tongue tip for the nasal. From a gestural point of view, the shifting of the gesture and the earlier velar closure in voiceless contexts is likely to be due to conflicting phonetic properties of a nasal and a voiceless oral stop in production. Shifting the velum gesture into the preceding vowel may help resolve this conflict. If such a trade-off takes place in production, then listeners on the perceptional side are assumed to perceive the acoustic characteristics of the lowered velum gesture instead of separately assigning nasalization to the vowel and nasal. In other words, listeners, at least some, are expected to treat the heard cues as perceptually equivalent.

From this perspective, the experimental findings from Beddor (2009) are considered once again. From the production test with data from six American speakers, a clear negative correlation between the extent of vowel nasalization and the duration of the nasal stop in English CVNC words was evident. Likewise, the voicing condition of the post-nasal stop affected the length of the nasal, which was more extended when followed by a voiced rather than voiceless obstruent. Moreover, voicing also affected the remaining gestures: when followed by a voiced stop, the preceding gestures for vowel length, vowel nasalization across ṼN and alveolar constriction across NC were slightly longer compared to the voiceless condition. The subsequent perception experiment with manipulated Ikalanga CVNCV words showed that listeners' performance was best when discriminating those pair members with different extents of overall temporal nasalization. Most difficulty appeared with similar-nasality pairs, that is, pair members with similar overall nasalization that was distributed variably across the vowel and the nasal segment. Considering both the production and perception results as a whole, some discrepancies were conspicuous: on the one hand, listeners attended to the acoustics of the lowered velum rather than to its precise distribution, treating the nasalized vowel and the nasal as perceptually equivalent. On the other hand, the production data showed a systematic earlier onset of velum lowering in the context of a voiceless stop, with longer nasalization of the vowel and a shorter nasal. If listeners attended to the overall nasality, but showed systematic patterns in production depending on the following stop voicing, stop voicing was expected to influence listeners' decisions as well. Two further perception experiments were created to illuminate this issue. In the first one, natural utterances of *bed, bet* and *mend* were edited by cross-splicing the segments to create continua series of *bed–bẽnd* and *bet–bẽnt*, for which the nasal was edited with respect to duration and the vowel in terms of nasalization percentage. The nasal consonant continuum consisted of ten steps from zero to 85 ms, and the continuum for the vowel consisted of three steps from zero to 66% nasalization. The nasal and vocalic segments were combined such that three groups of stimuli were classified: a group with pairs only differing in nasal duration, pairs which differed in the overall nasality extent and pairs with an overall similar nasality (i.e. similar overall nasality extent but distributed differently across the vowel and nasal). The stimuli pairs were presented to 32 native American English listeners who were asked to decide whether the pair members sounded the same or different. As predicted, the *bed–bẽnd* stimuli showed the most correct responses when presented in pairs with different overall nasality, followed by varying nasal duration and similar overall nasality. In contrast, stimuli pairs with a voiceless stop were discriminated most accurately when the pairs showed different but also similar overall nasality, followed by pairs with nasal stop duration differences. According to Beddor, this was expected, since in production vowel nasalization was found to be much more extended in voiceless contexts and thus potentially might have been sufficient as a cue for nasality in perception. Hence, at least some listeners were assumed to use vowel nasalization as the relevant cue rather than overall nasalization for making their decisions, such that differences in vowel nasalization alone (also present in the pairs with overall similar nasality distributed variably across the segments) might have been the crucial factor for their decisions. Indeed, results were highly listener-specific: while some listeners treated nasalization on the vowel and the nasal as perceptually equivalent in both voicing contexts, others showed high discrimination performances for this group, irrespective of the voicing condition. Thus, those listeners with the more accurate results were assumed to precisely attend to vowel nasalization, perceiving fine details in distribution and therefore being able to differentiate between pair members despite their similar overall nasality. A third group of listeners showed mixed results, with difficulties in judging pairs with similar nasality in the voiced condition but doing better at discriminating pairs with a voiceless context.

Results for the subsequent perception experiment provided additional evidence for listeners' assessment of coarticulatory cues. In this four-choice test, listeners were asked to identify the presented words either as *bed, bend, bet* or *bent.* The stimuli were edited similarly to those in the discrimination test with three different degrees of vowel nasalization and ten steps of nasal duration. Results revealed that listeners required a longer nasal to identify a CVNC word when the final C was voiced rather than voiceless. This was true for all degrees of vowel nasalization. Second, for both voicing conditions, an increase of vowel nasalization compensated for shorter nasal duration: less duration of the nasal was required to identify CVNC when vowel nasalization was extended. In the extreme form with 66% of the vowel nasalized, often no nasal was needed at all to identify CVNC in the voiceless condition. Considering these findings in a more differentiated way, listeners turned out to show specific but systematic patterns in weighting the available cues. For instance, the same acoustic information signal with a shorter nasal and moderately nasalized vowel was perceived as *bet* for some participants and as *bent* for others, presumably those who heavily relied on vowel nasalization as the relevant cue. That is, listeners showed patterns which could be classified as relying on perceived equivalence or solely on vowel nasalization:

some listeners referred to the overall cue of nasality for making their decisions, while others were found to rely on vowel nasalization instead.

From all these results, Beddor concluded that first, from an articulatory perspective, speakers of different languages vary in the onset of velum lowering depending on the following context. The velum gesture, which usually is assigned to the consonantal constriction for a nasal, is associated with a vocalic configuration, allowing for covariation between the nasalized vowel and the nasal consonant. Thus, the velum lowering gesture variably comprises the vowel and nasal consonant across speakers. In some contexts, however, the distribution of this gesture turns out to be systematic and context-dependent, especially in those contexts favouring earlier lowering in the vowel and a shortened nasal, which leads to a temporal trade-off relation between these two segments. Second, in perceptual terms, the listener has experience with these patterns and perceives the given nasality of the vowel and nasal as equivalent, or, in specific contexts, attends more to the nasalized vowel than to the nasal consonant. Overall, the acoustic signal provided by the speaker can be differently weighted in perception by different listeners, such that some may arrive at a different phonological interpretation than intended by the speaker.

To illuminate the dynamic perceptual processing unfolding over time, Beddor et al. (2013) ran a series of eye tracking experiments in which American listeners were to match an acoustic stimulus with an appropriate picture. According to the hypothesis tested, listeners were expected to dynamically use the unfolding coarticulatory information of a word to predict its outcome. In a first step, vowels from natural CVC and CVNC words were cross-spliced and edited, such that the vowels in nasal condition were provided with two degrees of temporal vowel nasalization and the stimuli ended either in a final voiced or voiceless stop. In addition, stimuli of C$\tilde{\mathrm{V}}$C were created to test whether listeners treated the C$\tilde{\mathrm{V}}$NC and C$\tilde{\mathrm{V}}$C sequences similarly when the final stop was voiceless rather than voiced. Overall results indeed showed that listeners relied on coarticulatory cues unfolding over time as soon as such cues were available: participants fixated the correct picture more quickly when early coarticulatory nasalization was present, but only in those picture pairs where it served as disambiguating cue (i.e. pictures referring to words of C$\tilde{\mathrm{V}}$NC–CVC, but not in C$\tilde{\mathrm{V}}$ND–C$\tilde{\mathrm{V}}$NT). With respect to coda voicing, listeners were more likely to look at the CVNC image in pairs of C$\tilde{\mathrm{V}}$NC–C$\tilde{\mathrm{V}}$C for both the voiced and voiceless conditions. However, the temporal difference in fixations was greater when the coda was voiced, which indicated that "vowel nasalization alone, without a nasal consonant, is a more convincing instance of a CVNC word in voiceless than in voiced contexts." (Beddor et al., 2013, p. 2361). As with the previous experiments, the authors emphasized that the data analysis for individual participants clearly showed listener-specific differences in the usage of the coarticulatory cues for predicting and identifying the upcoming word.

In their subsequent research, Beddor and colleagues (2015) were more engaged with a direct link between production and perception within the same language user. One main question was whether those speakers who showed extensive anticipatory vowel nasalization in production were also sensitive to this cue in perception, and, conversely, whether speakers who nasalized the vowel less also poorly assessed nasalization perceptually. To test this, a replication experiment of the eye tracking study from Beddor et al. (2013) was run with 32 new participants, all native speakers of American English. Again, overall results showed that listeners started to fixate the picture referring to a CVNC word rather than CVC shortly after vowel nasalization had become audibly available. As expected, the extent of assessing Ṽ as a cue for an upcoming nasal differed across listeners: while for some, the nasalized vowel was sufficient even without the nasal stop to fixate the CVNC picture, others relied on the nasal stop for making their decision. To obtain the production data, the same participants were acoustically recorded while reading out words from the perception experiment. The recordings were analyzed with respect to the main acoustic properties of nasality, a decrease of F1 for non-high vowels and an increase of the amplitude for the low-frequency nasal formant. Overall analyses clearly showed more nasalization in CVNT sequences across the vowel relative to CVND contexts. However, as in the perception experiment, individual productions highly differed with respect to anticipatory coarticulation: some speakers showed heavy vowel nasalization, while for others, vowel production was hardly affected by the following context. In these cases, vowels followed by a nasal stop showed similar acoustic properties to those obtained for oral vowels. Considering a possible link between production and perception, results for individual participants indeed indicated that speakers' usage of anticipatory nasalization in production also reflected their perceptual sensitivity towards nasalization.

Following the basic concept of such a correlation, a similar study was conducted for Afrikaans, which traditionally is seen as displaying a phonological contrast between prevoiced and voiceless unaspirated stops. However, an ongoing sound change is assumed based on the observation that the voiced stops in younger speakers' utterances tend to lack prevoicing, thus approaching the voice onset time of the voiceless stops. The voicing contrast, though, is still preserved by a difference in fundamental frequency (f0) of the following vowel: in post-stop position, f0 is lower when the stop is phonologically voiced rather than voiceless. Following the production-perception hypothesis, speakers with greater differences in the post-stop fundamental frequency were expected to attend more to f0 in perception than speakers who produced similar f0 in both voicing conditions. In fact, acoustic analyses of recordings from older and younger Afrikaans speakers revealed that the older speakers were much more likely to produce prevoiced stops than the younger group. Unexpectedly, however, both groups showed systematic differences in fundamental frequencies of the post-stop vowel, irrespective of stop voicing: higher f0 occurred after phonological /p/ rather than /b/, regardless of whether /b/ was realized as prevoiced [b] or voiceless [p]. Thus, while the voice onset time varied across age groups, the differences in f0 seemed to depend on the phonological voicing and turned out to be stable in both groups. For the perception test, the same speakers participated in an identification task in which the presented stimuli differed both in stop voicing and f0 on a continuum scale. For both age groups, when no voicing was present, f0 was sufficient for the /b/–/p/ voicing contrast. Provided with some voicing, listeners were more likely to identify initial /b/, even when f0 was high. This effect was more apparent for the older participants, who had a clear tendency to rely on prevoicing when making their decisions. Thus, f0 as a distinctive cue turned out to be stable both in production and perception, but the data indicated that older speakers-turned-listeners were more likely to produce /b/ with a prevoiced stop than younger participants, and accordingly in perception, older listeners relied on the prevoicing cue rather than on f0 to make their decisions. The relation pattern between production and perception of voicing and f0 also held for several individual speaker-turned-listeners. However, rather than postulating a strict correlation between perception and production, Beddor assumes that flexibility in perception is more expected than in production, such that speakers who are inconsistent in producing a target property may as listeners nonetheless be sensitive to it. This is even expected because "speech perception is malleable and adaptive; listeners perceptually retune depending on phonetic context, speaker, speaking rate, novel experiences, and much more." (Beddor, 2015, p. 7). In agreement with this proposition, the experiments revealed that some of those speakers who did not use the target property in production (vowel nasalization or voicing, respectively) were sensitive to it in perception, and some were not. In turn, speakers who used the target cue consistently in production were also always found to be perceptually sensitive to it.

Similarly, Zellou (2017) investigated the relation between individual American English participants' production patterns and their perceptual compensation for vowel nasalization. For this, 39 speakers participated in a reading task with CVC, CVN and NVN words. The acoustic recordings were analyzed in terms of vowel nasalization, estimating the degree of nasalization in CVN relative to the other contexts for each speaker individually. As expected, speakers highly differed in their usage of coarticulatory nasalization. Next, the speakers took part in a four-interval-forced-choice discrimination perception test, in which series of two stimuli pairs were presented. Participants were asked to indicate the pair which exhibited the more differing vowels. The stimuli consisted of CVC–CVN–NVN triplets for two vowels (*bed–ben–men* and *bode–bone–moan*) with three different degrees of nasalization (zero, light and heavy), depending on the surrounding context. These triplets were edited by cross-splicing the vowels within each vowel category, such that each context was assigned three new vowels of the same category but with three different degrees of nasalization. Afterwards, stimuli were paired into groups of two, for each vowel category separately. The prediction was that speakers who showed a large extent of coarticulation in production would compensate for nasalization in perception, i.e. assign the perceived nasality more often to the nasal source if available. In terms of response accuracy, these listeners were therefore expected to be less precise than those who showed less coarticulation in production, who instead were predicted to compensate less for nasalization and be more sensitive for differences. Results were twofold: first, listeners had indeed overall more difficulties in discriminating vowels if stimuli pairs involved varying contexts (e.g. CVC–CVN ∼ CVC–CṼN) rather than identical conditions (e.g. CVC–CVC ∼ CVC–CṼC). This was ascribed to the presumed compensation effect: nasality differences between the vowels were perceived more easily if the context was held identical across all stimuli, such that a direct comparison of the vowel acoustics was possible. In contrast, in pairs with varying contexts, nasality of vowels in nasal contexts was perceptually assigned to the nasal stop itself, such that the vowel was rather perceived as oral and therefore as similar to the oral vowel in the counterpart stimulus. Second, with respect to the predicted correlation between production and perception, Zellou indeed found that speakers who showed heavy nasalization patterns also tended to have more difficulties in discriminating the stimuli pairs, indicating that they compensated more for nasalization than those speakers who showed less coarticulation in production. In a third experiment, listeners participated in a rating task and were asked to judge perceived nasality in single stimuli pairs by indicating which word sounded more nasal. Again, results revealed overall compensatory effects in which rating accuracy was poorer if a nasalized vowel was surrounded by nasal context and thus interpreted as oral. However, contrary to the discrimination task, participants' performances in the rating test did not correlate with their coarticulatory production patterns.

The experiments described above show that coarticulatory vowel nasalization is interpreted in a complex way on the part of the listener. On the one hand, perceived nasalization can be used to predict the upcoming context, especially in languages with extensive coarticulatory vowel nasalization. On the other hand, these cues are partially compensated and ascribed to the nasal consonant, if present, such that the vowel is interpreted more as oral rather than nasal. Thus, coarticulatory vowel nasalization seems to serve a twofold purpose (Zellou, 2017, p. 15): on the one hand, it is used to efficiently predict the intended utterance if provided in appropriate context, on the other hand, partial compensation is used to identify what was said. Furthermore, evidence has been provided for a correlation between the usage of vowel nasalization in production and perception: across different language groups, those speakers with less anticipatory vowel nasalization are found to compensate less for this effect in perception. Within speakers of the same language group, similar effects are observed: strong coarticulation in production is reflected by higher compensation patterns in perception (although this pattern is not entirely consistent). Correspondingly, listeners who in perception use vowel nasalization to determine the following context are often those with more extensive anticipatory vowel nasalization in production. However, there are also studies which could not confirm such a correlation of production and perception (e.g. Grosvald and Corina, 2012 and also Zellou, 2017 for some of the data). It has been suggested that speakers who do not make use of such coarticulatory cues may nonetheless be sensitive to them in perception, such that language users are assumed be more flexible in perception than in production (Beddor, 2015). In addition, the findings from American English, Hindi and Thai indicate that experience with coarticulatory vowel nasalization and its interpretation are highly language-specific and must be taken into consideration.

5.1.2 Experiment VII: Staircase perception experiment

The findings from the experiments above show that at least for some languages, individual speakers may be classified as 'innovative' (cf. Beddor, 2012, p. 53) by attending to vowel nasalization as a relevant distinguishing cue in perception (Beddor, 2009; Fowler and Brown, 2000; Ohala and Ohala, 1995). Moreover, the individual extent of vowel nasalization in production has been found for some speakers to be related to the perceptual usage of these cues (Beddor, 2015; Zellou, 2017). However, a majority of these studies involved speakers and listeners of American English, a language with rather extensive anticipatory vowel nasalization, depending on the geographical area.

To investigate whether similar correlation patterns can be found for German as well, and to generally test the perceptual sensitivity of German listeners with respect to vowel nasalization, an adaptive staircase experiment (Gerrits and Schouten, 2004; Kaernbach, 1991) was run with German participants who had previously participated as speakers in the MRI study.

Predictions

Following the studies outlined above, the questions arise whether and how vowel nasalization is perceived and used as a perceptual cue in other languages, and whether a correlation between coarticulation in production and perception may be also evident. Taking into account the theoretical background of the discrimination experiments described in Beddor (2009, pp. 799-810), German listeners were tested for perceived differences in stimuli when the overall velum lowering extent was either temporally modified or shifted as a whole along the vowel and nasal segment. Following the findings from American English, the hypothesis was adopted for German that alterations of the overall nasality should be detected more easily than the shifting of a constant nasality interval. In a second step, it was further examined whether a correlation was apparent between individual speakers' nasalization patterns in production and their specific sensitivity to nasality cues in perception. However, in terms of predicting the direction of correlation direction, it was not straightforward what to expect: one possibility is that participants who show heavy anticipatory nasalization patterns may provide better discrimination performances due to their more extensive linguistic experience with coarticulatory nasalization. A different possibility suggests that such listeners would give less accurate responses due to compensatory effects. As pointed out, listeners can be sensitive to nasalization patterns in various ways: on the one hand, they may use perceptual cues to attribute the coarticulatory effect to its source, such that nasalized vowels in nasal contexts are rather perceived as oral. On the other hand, the acoustic effects are used to derive information about the source itself, as exemplified by experiments with American English, in which listeners largely completed CṼ sequences by adding a nasal rather than oral stop. Thus, given the ambiguous options for the usage of perceptual cues, the following predictions for the perception test with German listeners include both directional and non-directional hypotheses:

Hypotheses: Vowel nasalization in perception and production

H1 German speakers show individual temporal degrees of anticipatory vowel nasalization in production.

H2 German listeners show higher discrimination abilities when the overall temporal extent of nasalization is modified rather than shifted along the vowel and nasal.

H3 For German participants, the temporal extent of anticipatory nasalization in production is correlated with their perceptual sensitivity to nasality cues.

In the following, findings from the perception test are presented first and then combined with the production data, exploring the question of whether a correlation is apparent between vowel nasalization in production and perception for speakers-turned-listeners as native language users of Standard German.

Stimuli

Stimuli patterns of AABA, ABAA, BBAB and BABB sequences were created, with the target stimulus in second or third position. As stimuli, the word [baːntə] was created and edited with regard to the velum lowering gesture, which was done by means of the software tool VocalTractLab (Birkholz, 2013), version 2.2 API for Windows (10 November 2017). The download link is provided in the bibliography.

The discrimination experiment comprised two conditions of nasalization patterns: one in which a constant velum lowering interval was shifted along the vowel and nasal consonant, and one which contained a consistent nasal consonant but a temporally modified portion of vowel nasalization. For both conditions, stimulus A was constructed in the same manner, representing a near-natural utterance of the target word with a nasal consonant of 80 ms and no vowel nasalization present. Although in natural speech, anticipatory vowel nasalization is always present to some extent, this stylized construction allowed for more controlled modifications in a systematic way.

Condition one (henceforth the ‘constant condition’) implied a constant interval of a total of 80 ms velum lowering. At the start, the difference (‘delta’) between the A and B stimuli corresponded to this maximum of 80 ms. For example, three A stimuli with 80 ms of nasal consonant duration and 0 ms of vowel nasalization were opposed to one B stimulus with 0

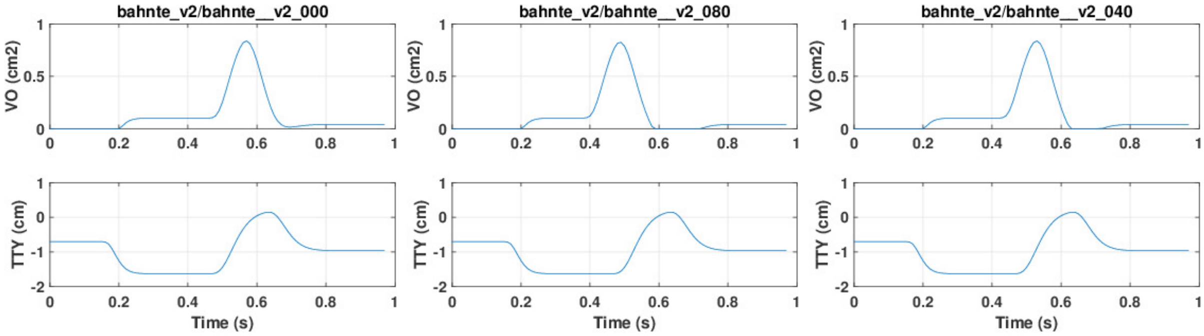


Figure 5.1: Shifting pattern for the constant condition. Left: basic configuration in stimulus A: tongue tip (TTY) and velum opening (VO) start at the same time point. Middle: VO interval shifted into the vowel by 80 ms (stimulus B): delta of stimuli A and B is at its maximum. Right: VO interval shifted into the vowel by 40 ms (stimulus B); delta of stimuli A and B is half of its maximum.

Figure 5.2: Sounds and spectrograms of two stimuli of the constant condition with delta at its maximum. Left: velum opening (VO) of 80 ms after vowel offset. Right: VO of 80 ms shifted into the vowel.

ms of nasal duration but 80 ms of nasalization of the last portion of the stressed vowel. In the course of the experiment, delta was decreased by shifting the nasalized interval as a whole in the B stimuli (details in the procedure section). Figure 5.1 exemplifies the shifting procedure for the constant condition. Corresponding sounds and spectrograms for the left and middle panel are given in fig. 5.2.

In contrast, condition two (the ‘extended condition’) involved B stimuli with a constant nasal consonant of 80 ms but varying temporal nasalization on the vowel, which ranged from zero to full nasalization of 350 ms. Figure 5.3 exemplifies three trial modifications of the nasalized time span during the vowel. At the beginning of the experiment, and as with the constant condition, delta corresponded to the maximum nasalization difference between A and B, which was 350 ms. Figure 5.4 shows corresponding sound and spectrogram patterns for the left and middle panel of fig. 5.3. Further, the stimuli were constructed by implementing the post-vocalic consonantal portion (/n/+/t/) with a tongue-tip gesture of a total of 160 ms. For stimulus A, the velum lowering duration was set to 80 ms, such that the nasal stop resulted from the synchronous onset of the tongue tip and the velar gesture. For all stimuli, the post-nasal voiceless stop was generated by placing a glottal gesture with the VocalTractLab specification “slightly breathy” over the remaining part of the tongue-tip gesture.

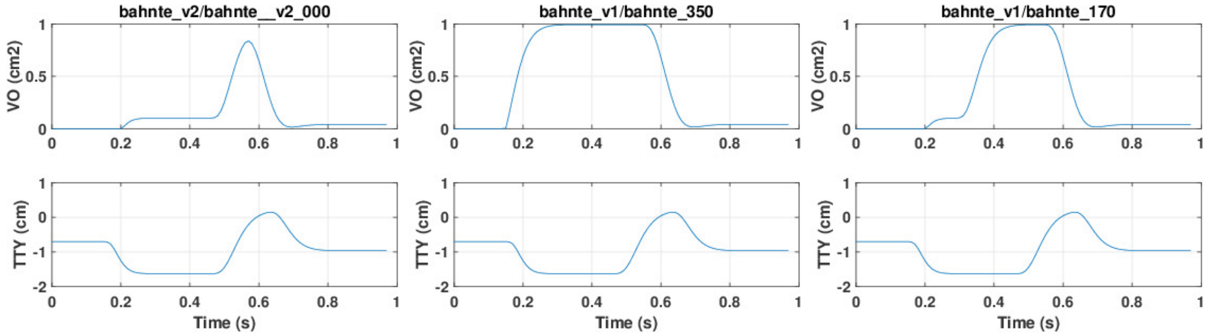


Figure 5.3: Shifting pattern for the extended condition. Left: basic configuration in stimulus A: tongue tip (TTY) and velum opening (VO) start at the same time point. Middle: overall VO of 350 ms during the vowel (stimulus B): delta of stimuli A and B is at its maximum. Right: VO of 170 ms during the vowel (stimulus B); delta of stimuli A and B is about half of its maximum.

Figure 5.4: Sounds and spectrograms of two stimuli of the extended condition with delta at its maximum. Left: vowel nasality = 0 ms. Right: vowel nasality = 350 ms.

Participants and procedure

Data are presented for 20 listeners who had previously participated as speakers in the MRI measurements. The perception experiment was run roughly two years after the MRI study, for which reason not all former speaker-participants were available for the perception experiment. As with the MRI recordings, listeners were paid for participation.

Listeners were tested individually in a quiet room and were familiarized with the task by listening to several audio trial examples of each condition, presented via a notebook. In addition, they were informed that the differing stimulus was either in second or third position and that there were two conditions which would be tested separately. After each practice trial, listeners were asked to indicate the target stimulus and were given the correct answer immediately. For the main experiment, listeners were informed to press the key with either the number 2 or 3 on the keyboard and to guess if necessary. Participants wore Sennheiser HD TV 65 headphones and were allowed to adjust the volume as desired. The acoustic trials were presented via a Matlab script. Listeners had no time limit for making their decisions. Information about the outcome appeared immediately on the screen after the entry. All participants started with the constant condition, since during practice, stimuli differences turned out to be perceived much more easily in this condition than in the second version, where variable nasalization affected only the vowel.

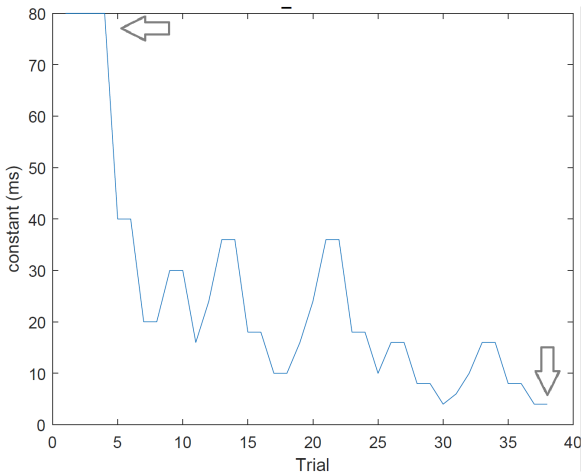


Figure 5.5: Exemplary staircase plot for one participant (constant condition). Delta is indicated by the y-axis, the trial number by the x-axis. The plot shows 12 turning points: one turn after the first two correct responses (left arrow) plus 10 turns visible in the curve plus one final turn when the twelfth turning point is reached (right arrow).

The overall advantage of the adaptive staircase method (Gerrits and Schouten, 2004; Kaernbach, 1991) is an ongoing adjustment to listeners' discrimination thresholds: depending on each response outcome, delta is either increased or decreased. In this experiment, delta was divided by 2 after two consecutive correct responses and immediately increased by 50% if one answer was wrong. Thus, the gradual delta modification was expected to approach listeners' individual discrimination thresholds after several trials and was assumed to stay closely above and below this value. Since only delta was adapted, the trial patterns randomly alternated with respect to the direction of the nasalization approach. For example, if B in the trial ABAA was discriminated correctly, the next trial for confirmation could involve BBAB with the same delta but reverse trial. As the listeners were expected to finally oscillate around their threshold boundaries, the experiment ended after 12 turning points but with a fixed upper limit of 50 trials. Figure 5.5 exemplifies the staircase procedure with the constant condition for one participant. A fictitious example of the trial adjustments depending on the outcome is additionally provided in the appendix (table B.1).

Results: Perception

The determination of a specific by-participant threshold value involved the following considerations. Simply calculating the mean across all turning points would have also included the higher values in the beginning of the experiment, when participants were about to reach the lower delta step-by-step. On the other hand, averaging across the lowest three or four overall turning points would have led to values which rather indicated incorrect responses: as these values elicited an immediate increase of delta, they actually did not belong to the participants' threshold range. In addition, if low threshold values occurred as sporadic events in the course of the experiment, these might have rather represented chance level instead of systematic perception skills. For this reason, a fixed number of reversals was selected that corresponded to the six consecutive turning points with the lowest threshold values. Thus, listeners' individual perceptual thresholds referred to a period during which the listeners were relatively constant in their decisions. Since there was a maximum of 12 reversals, seven intervals with each containing six turning points were considered (reversals 1-6, 2-7, 3-8 etc.). The interval with the lowest overall mean was selected, which most often referred to one interval during the middle rather than the end of the experiment.

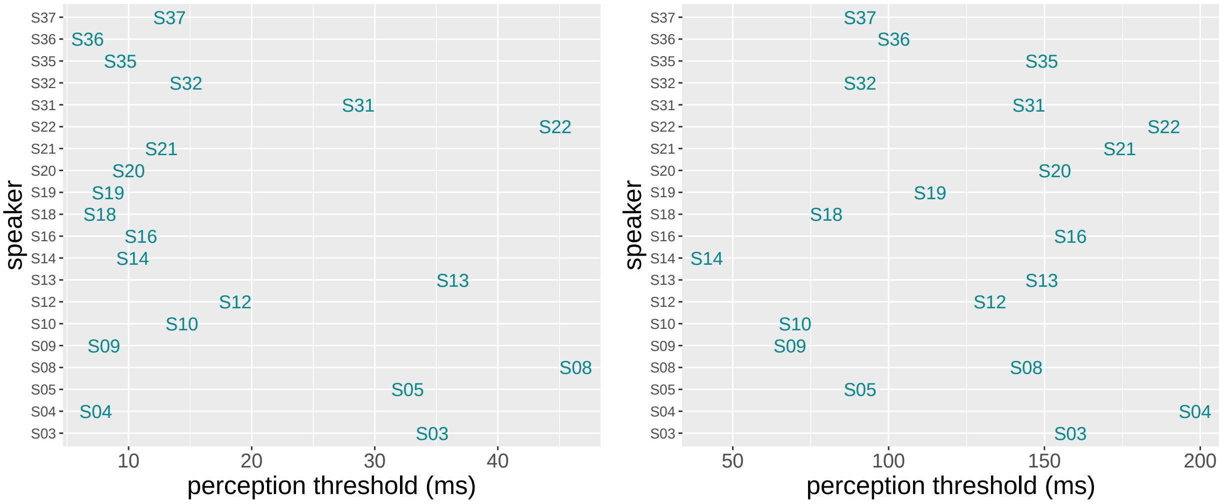


Figure 5.6: Perception threshold means of the lowest six consecutive turning points for 20 listeners. Left: thresholds for perceived differences when the lowering velum gesture was shifted (constant condition). Right: thresholds for perceived differences of the nasality extent within the vowel (extended condition).

Figure 5.6 shows the results for all participants for the constant and extended condition, respectively. Lower threshold values indicate smaller nasality differences between the stimuli, suggesting that participants with low values noticed the fine distinctions more precisely. For a better comparison of the conditions, fig. 5.7 shows the means for the constant condition as a function of the means of the extended condition. Evidently, most participants showed overall finer discrimination abilities in the constant condition compared to the extended type. In addition, a more consistent pattern is evident for the constant condition, where 14 of the 20 data points are located below the value of 20 ms in contrast to a more sporadic distribution for the extended condition. Moreover, responses for the two conditions did not necessarily correspond for individual listeners, as fig. 5.7 indicates. For instance, while participants S14, S09 and S10 showed relatively low perception thresholds in both conditions, participant S04 had one of the lowest values in the constant condition but the highest in the extended version. Conversely, S05 exhibited a relatively high threshold in the constant condition but had less discrimination difficulties in the extended condition compared to most other listeners.

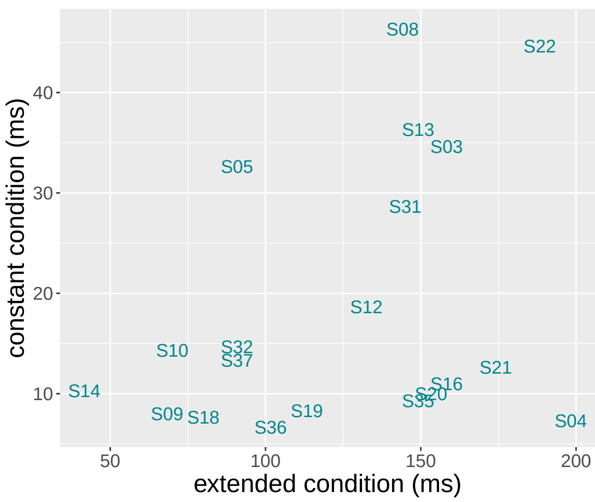


Figure 5.7: Means of the lowest six consecutive turning points for 20 listeners. X-axis: extended condition. Y-axis: constant condition.

Interim discussion

The perception threshold differences between the constant and extended condition are to be considered in more detail. As suggested by figs. 5.6 and 5.7, listeners were poorly sensitive to nasality differences if these affected solely the vowel. In contrast, when the gesture was shifted as a whole, differences between the stimuli were perceived more distinctly. This finding is in direct contrast to H2, which claims that vowel nasality differences are detected more easily than differences between stimuli with varying distributions of a constant gesture. If listeners relied on vowel nasalization as a relevant cue, the threshold values in the extended condition would have been generally lower and more consistent across listeners. As it turned out, though, German listeners were highly insensitive to alterations of the temporal extent of vowel nasalization and probably to the presence of coarticulatory vowel nasalization at all. Consequently, the low thresholds in the constant condition do not indicate listeners' perception of vowel nasality but rather their sensitivity to other acoustic modifications coming along with the overall shifting. For example, listeners may have attended to the silent interval between the nasal offset and the offset of the /t/, which increased the more the 80 ms interval was shifted into the vowel. Bringing the oral stop closer to the vowel by the same amount was not possible because this would have reduced the overall word duration, obscuring the virtual effect under consideration. Therefore, the increased silent interval was presumably interpreted as a longer time span of alveolar stop closure for /t/, inducing the effect of a rather 'strong' /t/ compared to an oral stop with a short closure period similar to a weak /t/ or even /d/. As German language users are familiar with alveolar stop closure differences between /d/ vs. /t/, this might have helped them in identifying the target word more precisely.

A similar but slightly different approach considers the extent of the nasal portion itself: listeners may have relied on the stepwise absence (or in reverse trials on the increasing presence) of the nasal consonant. This idea is supported by the evidence from the production data in chapter III, which showed that the interval of the overall velum lowering gesture (OVL) in CVNCV contexts was not just shifted more into the vowel compared to CVNV contexts, but also reduced in its temporal extent, resulting in a shortened nasal consonant. As mentioned previously, similar results were provided by Carignan et al. (2019) and Carignan et al. (2021), who investigated nasalization patterns in German /Vnt/ versus /Vnd/ sequences in the style of Beddor's suggestions for American English. Instead of a constant gesture being shifted along the segments, they reported that the velum lowering gesture in /Vnt/ was generally truncated (with a slight shift evident, though), that is, velum lowering was reduced both in its spatial and temporal amount. Thus, if articulatory modifications of the nasal consonant systematically occur in production, German listeners might be sensitive to such nasal weakening patterns in perception as well, at least more than to alterations of vowel nasalization. In fact, the two conditions examined provided some appropriate contexts to test this approach: listeners turned out to be quite sensitive to nasal consonant weakening (i.e. temporal shortening or absence), but much less to variations of vowel nasalization.

Results: Perception and production

Following the idea that speakers' behavior in production may be reflected by their perceptual sensitivity, H3 as defined in the hypotheses of this chapter was examined by combining the perception results with the production data of the individual participants. As demonstrated by the prior chapters, the largest extent of vowel nasalization was found in pre-nasal tense vowels. Thus, to elaborate a ranking in terms of by-speaker temporal vowel nasalization, 26 words in carrier phrases with broad focus (for details on the focus condition see section 1.6.1) were analyzed with tense vowels in pre-nasal position. The post-nasal context involved both N (n=303) and NC (n=215) contexts. The detailed list of the target words is attached in the appendix (table A.11).

To determine the individual by-speaker ratios of vowel nasalization relative to OVL, the amount of vowel nasalization was divided by OVL for each target word produced by each participant. Thus, higher ratios corresponded to a larger extent of vowel nasalization relative to OVL. As with the previous analyses, vowel nasalization was defined as the time span between the point of maximum velocity during velum opening and the acoustic vowel offset, while OVL refers to the interval between the points of maximum velocity during the velum opening and closure gesture (see introduction 1.6.4). Figure 5.8 shows the by-speaker ratios ranked by the median based on the production target words. The largest difference between the ratio medians on the extreme ends occurred for speakers S13 and S31: for S13, fifty percent of the ratio values were larger than 0.508 (mean=0.473, sd=0.17), while for speaker S31, fifty percent were above 0.186 (mean=0.189, sd=0.21). To investigate the relationship between perception and production, data sets were combined by calculating the by-speaker ratio means, which were then defined as a function of the individual delta means from the perception test. In case of dependency, data were expected to indicate a tendency either in terms of a positive correlation (if speakers with higher ratios were

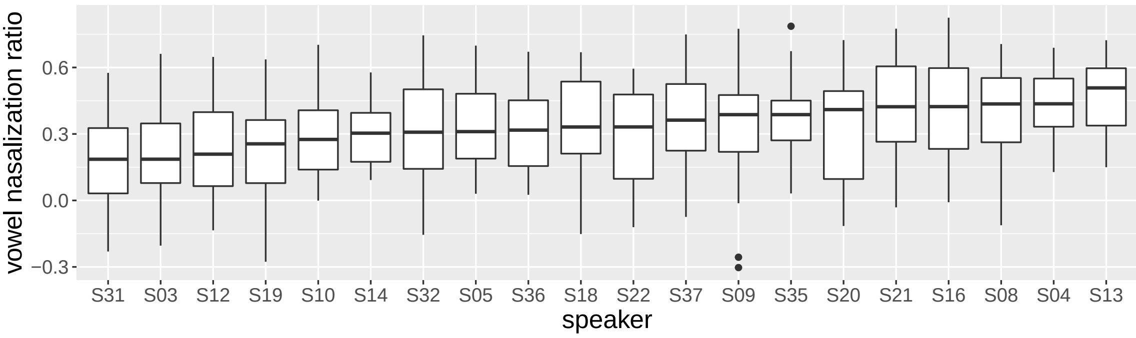


Figure 5.8: Ratio of vowel nasality to OVL (vow.nas/OVL) of target words with CVNV and CVNCV sequences for 20 speakers.

less sensitive, possibly due to compensatory effects) or a negative correlation (if speakers with high ratios were more sensitive, possibly due to coarticulatory experience). Figure 5.9 shows the combined data sets for both experimental settings separately.

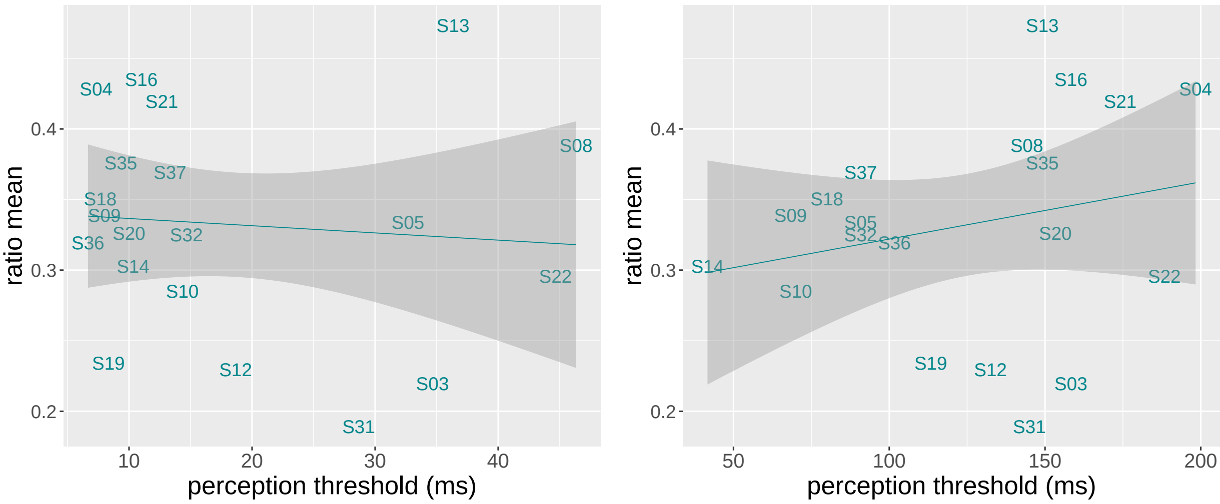


Figure 5.9: Means of the nasality ratios (vow.nas/OVL) and the perception thresholds of the lowest six consecutive turning points for 20 participants. Results are shown for the constant (left) and extended (right) condition.

Considering the constant condition, fig. 5.9 (left) suggests a very weak relationship between nasalization ratios and discrimination performances in perception. On the other hand, there might be a weak positive relationship for the extended condition (fig. 5.9, right). For statistical verification, a Pearson correlation test was applied to the data referring to the extended condition. Results revealed no correlation between the production and perception values (r=0.230, p=0.33). For the constant condition, a Spearman correlation test[1] was run, which also did not reveal any significant correlation effect (r=-0.158, p=0.51).

It should be noted that at first glance, the outcome might be considered somewhat oversimplified as this analysis refers to production data that include target words with two different coda contexts and seven different vowels, such that some possible correlation effects for specific contexts might be obscured. For example, the temporal extent of vowel nasalization has previously turned out to depend on the vowel category, with tense /a/ showing the largest portion in particular (see chapter III). Therefore, for some vowels a correlation could be evident, which may be disguised as the means refer to the overall data across the vowels. Similarly, as demonstrated in chapter III, the coda condition has been found to affect the extent of vowel nasalization, such that for words with CVNCV sequences, the ratios are expected to show higher values than for words with CVNV sequences. Thus, correlation effects might be visible more clearly in one particular condition rather than in the combined data. In fact, during the course of this survey, the data have been grouped by the contexts mentioned and considered in a more differentiated way. Detailed plots

[1]The Spearman method was applied for the constant condition because data deviated from normal distribution.

for each environment are attached in the appendix (section B.2). A clear difference was indeed apparent for the ratio values of words with CVNV sequences compared to CVNCV sequences. However, broad inspection of the relation between the production and perception data did not reveal any systematic patterns, irrespective of the coda environment: for each perception condition, each coda condition and each vowel category, no context gave reason to consider an adapting analysis in addition to the overall data presented above. Thus, the finding that no correlation was evident between speakers' production data and their individual perception thresholds can be considered as representative for the more differentiated contexts.

5.2 Summary and discussion

This chapter investigated the perceptual sensitivity of German listeners to vowel nasalization as well as the relationship between speakers' nasality patterns in production and their discrimination skills as listeners in perception. Results revealed that on the one hand, listeners were overall more precise in identifying the target words when the constant interval of the nasalized portion was shifted as a whole across the vowel and nasal segment. More difficulties arose when the overall temporal extent of nasality solely varied within the vowel. On the other hand, for both perception conditions, no correlation was found with respect to the means of participants' nasalization extent in production and their threshold boundaries in perception. Thus, for these data no relationship was apparent for individual German speakers to make use of coarticulatory vowel nasalization in production and its acoustic cues in perception.

As suggested in the interim discussion, the finding that listeners had less difficulties in identifying items with an overall shifted nasality interval might rather have to do with the perception of alterations in the silent intervals between the vowel and the oral stop in the target word /baːntə/, such that a larger time span of silence was interpreted as a longer time span of alveolar closure, indicating a clear /t/. After all, German listeners are familiar with acoustic differences between /t/ and /d/, which are contrastive sounds in German, in contrast to vowel nasalization. Alternatively, listeners might have perceived differences in nasal presence, such that when a target word with a heavily nasalized vowel but only a small nasal portion was opposed to a word involving a full nasal, listeners may have primarily relied on the alterations concerning only the nasal portion. As with the extent of alveolar stop closure, German listeners can be assumed to have some experience with nasal weakening in certain contexts as well, as indicated by the production data in chapters III and IV. Correspondingly, prediction H2 was rejected: German listeners did not perceive differences in vowel nasalization alterations; instead, they relied on modifications they were more familiar with as native speakers.

Considering speakers' individual relationship between vowel nasalization in production and perception, participants indeed showed individual levels of anticipatory vowel nasalization in production, confirming H1. However, no correlation to their perceptual performances could be attested. Thus, the finding that across listeners, perception skills might be related to some general experience with nasal weakening patterns in production cannot be applied to the individual language user in our data, contradicting prediction H3. Possibly for German, the variation patterns within a specific speaker's own production are too inconspicuous to systematically affect their perceptual sensitivity. Moreover, as outlined by Beddor (2015), the assumption of a correlation between production and perception patterns needs to include "a mechanism that allows for greater variation in perception than in production" (Beddor, 2015, p. 7), because listeners need to constantly adjust to the linguistic context they are exposed to, such as varying speakers or speech rate. By this, they are "expected to be sensitive to information that is available in the unfolding acoustic signal and should [...] exhibit greater flexibility in perception than in production." (Beddor, 2015, p. 7). In terms of the German data presented in this chapter, this point might have contributed to the finding that on the one hand, a clear tendency was apparent for a perceptual difference with respect to the two conditions tested, but at the same time, no strict correlation between production and perception patterns was found for individual speakers.

Apart from the correlation aspect in this survey, the perception test contributed further evidence to the basic research field of determining individual perceptual thresholds with respect to listeners' discrimination skills. In particular, not much is known about the temporal thresholds in perception, i.e. the temporal extent of nasalization that must be present in a vowel to be perceived as different from an oral vowel. Studies on the perception of nasality have typically explored whether listeners use acoustic cues to predict an upcoming segment or compensate for coarticulatory effects and whether different language groups show different response patterns. However, less is known about the particular acoustic and auditory boundaries which lead a listener to judge a stimulus to be different in a very specific constellation. As described in the introduction of this chapter, some evidence was provided by Hawkins and Stevens (1985), who tested different language groups for their 50% crossover points when listening to stimuli of a synthetic oral–nasal continuum. Spectral modifications were generated by introducing pole-zero pairs into the spectrum of a non-nasal vowel. Listeners' 50% crossover points for nasal responses occurred when the spacing of the pole-zero pair was in the range of 75-110 Hz. In a similar experiment, Stevens et al. (1987) found that listeners' crossover points occurred when the maximum perturbation of an additional pole-zero pair near the first formant was within a range of 6-9 dB relative to the non-nasal F1 spectrum. Moreover, Beddor and Strange (1982) tested Hindi and American English listeners for categorical discrimination when judging vowels on an oral–nasal continuum generated by means of synthesized velopharyngeal port opening. The 50% crossover points for identifying the vowel as nasal occurred when the velar port opening degree was around 12 mm^2, given continuum steps of 2.4 mm^2 in a range from 0 to 24 mm^2. When this range was extended from 0 to 36 mm^2 and steps were 3.6 mm^2, the crossover boundary shifted towards more nasality for both language groups. In a subsequent four-step discrimination test with the same stimuli from the vowel continuum, Hindi listeners were most accurate at distinguishing those stimuli pairs which differed in velum port opening degrees close to the identification boundary rather than to the extreme oral or nasal ends, indicating that they perceived the oral–nasal contrast as categorical.

The experiment presented in this chapter provides further insights into terms of the question of which fine perceptual details listeners are able to detect when confronted with different nasality patterns. The present study, however, investigated perception boundaries in temporal terms rather than adjusting the spectral shape in multiple steps and thus explored how much temporal extent of nasalization was necessary for a vowel to be perceived as different from an oral vowel. Results from our data showed that participants had much lower perception thresholds when modifications affected the coda rather than the nasal portion within the vowel. On average, participants recognized differences of ≈19 ms (sd=13.2 ms) between the stimuli of the constant condition, i.e. when the coda was affected. In contrast, differences between stimuli with varying vowel nasalization were perceived at an average of ≈125 ms (sd=43.66 ms). In other words, if stimuli differed less than 125 ms with respect to vowel nasalization, listeners were no longer able to differentiate between them. The threshold difference of more than 100 ms between the two conditions indicates that the German listeners were highly insensitive to coarticulatory vowel nasalization, but were remarkably sensitive to fine changes in the coda, irrespective of whether this was caused by the effect of /t/ enhancement or nasal stop weakening.

More generally, statements about these relatively fine differences are only feasible due to the overall concept of the staircase procedure: unlike previous perception tests, in which listeners were confronted with relatively coarse-grained steps on a continua, the staircase function allows for working toward the individual perceptual threshold of one specific listener. The question of how much temporal nasalization is necessary for a vowel to be perceived as different would be interesting to test for other languages as well, including languages with contrastive nasal vowels. The same holds for the exploration of discrimination thresholds for nasal weakening differences, when the nasal is followed by a voiced stop, a voiceless stop or located in word-final position. For example, when followed by a voiced stop, listeners might show poorer discrimination performances because a nasal and a voiced stop share more acoustic properties than a nasal and a voiceless stop (Ohala and Ohala, 1991). Thus, nasal weakening might be compensated to a certain extent by the acoustic parameters of the voiced stop, such that the stimuli sound more similar relative to each other. The staircase procedure may help in exploring such issues in future work.

Summary and conclusion

Coarticulatory vowel nasalization is present in vowels that are followed or preceded by a nasal consonant: the velum configuration required for a nasal stop is anticipated or maintained during the vowel. The extent of vowel nasalization is different for the individual languages, but is also influenced by more general linguistic parameters including the nature of the vowel, the surrounding context and prosodic factors. About one fifth of the world's languages exhibit nasal vowels that have a contrastive function: the presence or absence of nasality alone may change the meaning of a word. For most of these languages, the nasal vowels have evolved out of sound sequences of older language stages in which the vowel was followed by a nasal stop. It is generally assumed that at some stage, the vowel was coarticulatorily nasalized for some time, but became more and more nasalized over time until the coarticulatory effect of nasalization was associated solely with the vowel, while the actual source, the nasal stop, was lost. If this process is based on the general principles of the human articulatory and perceptual system, these mechanisms should also be verifiable in modern languages, even in those that do not exhibit contrastive nasal vowels. The aim of this thesis was to shed light on some of the principles involved in coarticulatory vowel nasalization by investigating the behaviour of the velum during fluent speech, especially during vowels in different segmental and prosodic contexts. The data were obtained from native speakers of Standard German, a language that currently does not exhibit strongly nasalized vowels and thus is appropriate for examining the very basic mechanisms of the gestural interplay during vowel-nasal sound sequences. Articulatory data were acquired via modern real-time MRI from more than 30 participants, which is a highly exceptional case in phonetic research.

Chapters II and III considered the inherent vowel properties affecting the velum position. While chapter II explored the effect of the vowel height, chapter III focused on the temporal extent of nasalization in vowels of different lengths when followed by different consonantal contexts. Chapter IV was concerned with the effect of different focus conditions on the velar behaviour during pre-nasal vowels and on the lingual and velar movement patterns during consonantal sequences in which a nasal stop was followed by an oral stop with different voicing. In addition, the effect of the speaking rate was considered for the tongue and velum position in post-vocalic oral and nasal stops. Chapter V investigated the perceptual aspect of vowel nasalization by examining how German listeners perceived nasality patterns that systematically varied either solely within the vowel or across vowel-nasal sound sequences. The present chapter provides a summary of all experiments as well as their basic findings, which are discussed with respect to implications for current approaches on the role of the speaker and listener in the evolution of contrastive vowel nasalization.

The effect of the tongue position on velum height Chapter II considered the impact of the vowel height on the velum position in tense and lax vowels preceding a nasal or oral stop. Studies on this issue suggest an inverse relationship between the tongue height and the velum position, as the low vowel /a/ has consistently been reported to be produced with a lower velum than the higher vowels. This pattern is evident for vowels in nasal and often even oral contexts (Amelot and Rossato, 2006; Bell-Berti, 1973; Clumeck, 1976; Lubker, 1968; Rossato et al., 2003). One approach to this finding considers the palatoglossus muscle inducing a pull-down effect on the soft palate when the tongue is in a low position (Dixit et al., 1987; Moll and Shriner, 1967). At the same time, perception experiments provide evidence that listeners perceive a high vowel as nasal with only little nasal coupling, while a larger opening of the velopharyngeal port (VP) is tolerated with the low vowel (House and Stevens, 1956; Lubker, 1968; Maeda, 1993; Ohala, 1975). One interpretation of the perceptual effect is that speakers show increased velum raising in high vowels to prevent these vowels from being affected too much by spectral changes that may elicit the perception of nasality.

Our findings are in partial agreement with the results from prior research: considering the position of the soft palate at the vowel midpoint in vowels followed by a nasal stop, the tense low vowel /a/ was clearly produced with a more lowered velum than the other vowels. However, this pattern was not observed for the lax vowels which overall exhibited velum lowering degrees that were more similar to those of high tense vowels. Moreover, results did not suggest a strict hierarchy of VP opening and vowel height with respect to the other vowels tested. Furthermore, the velum position was hardly affected by the horizontal position of the tongue, suggesting that velum lowering was similar in front and back vowels of the same height. Even less differences were observed in the oral contexts, in which the velum was found in a similar position for all vowels, irrespective of whether these were tense or lax. Thus, the only clear pattern evident was that the velum during pre-nasal tense (but not lax) /a/ was explicitly more lowered than during all other vowels, which leads to two considerations.

First, all pre-nasal vowels were nasalized to some extent, but except for /a/, the extent of VP opening was quite similar between the vowels. Although slight differences were apparent, with /o/ showing the second lowest velum position, these were not reported as significant. This pattern, however, may not be surprising, as the other vowels /i/, /ø/, /y/ and /u/ are generally produced with a relatively high tongue. Even if /ø/ is produced with a slightly lower tongue than /i/ and /y/, this tongue height may still be sufficient to result in the velum taking a similar position as in the high vowels. In addition, as the data on the temporal extent of nasalization show, the velum in /a/ exhibited a low position at the vowel midpoint because velum lowering (in our data the point of the maximum velocity during the velum lowering gesture) started before the midpoint, which was not the case for the other vowels. However, this still does not explain why speakers consistently show a lowered soft palate (and an earlier lowering onset) during the low vowel but not during the high vowels. That this pattern was not evident in the oral context might at the first glance argue against the idea that the palatoglossus is the primary factor, pulling down the soft palate when the tongue is low. On the other hand, the varying patterns for the oral and nasal contexts may be explained by the muscular interplay between the palatoglossus and the levator palatini (cf. Kuehn and Azzam, 1978, p. 356): for oral sounds, the levator strength may override that of the palatoglossus, inducing velum raising. This is not the case if the velum anticipates an upcoming nasal stop, for which the levator muscle activity is decreased. In these cases with a low levator activity and a low tongue position, the effect of the palatoglossus connection becomes visible. Alternatively, data can be explained by the perceptual approach: as it has become evident from the subsequent experiments, velum lowering starts much earlier in tense /a/ than in lax /a/ (proportionally to vowel length) and also earlier than in other tense vowels, which may indicate an articulatory pattern that is intended by the speaker. As the low vowel tolerates a higher degree of velum opening before its frequency spectrum is affected by the additional formants introduced, the speaker may anticipate the lowering gesture quite early in the vowel without running the risk of producing a sound that is perceived by the listener as unnaturally nasalized. However, as pointed out by Solé (1992, 2007), this pattern is likely to be language specific.

Second, lax vowels generally did not show the same velum lowering behaviour as tense vowels, suggesting that vowel length plays also a role. The reason for the different lowering patterns may be twofold: on the one hand, lax vowels are articulated less peripherally in the vocal tract than tense vowels. If the tongue is in a more centralized position instead of being distinctly retracted and lowered, this may also have some impact on the pull-down mechanism induced by the palatoglossus, which may be less dominant in lax low vowels. On the other hand, lax vowels are much shorter than tense vowels, suggesting that the velum simply may not have enough time to lower to such an extent that at the vowel midpoint, clear velum lowering differences become apparent for the different vowels.

Overall, data are largely compatible with findings from numerous past studies on different languages, suggesting that the low vowel exhibits a conspicuously low velum in nasal context. Data are also compatible with both a physiological and a perceptual account for the velum lowering patterns that occur during vowels of different heights.

Vowel duration and vowel nasalization Chapter III focused on the relation between the temporal extent of vowel nasalization and vowel duration of tense and lax pre-nasal vowels in CVNV and CVNCV sequences. The aim was to explore whether long vowels are preferentially nasalized compared to short vowels and whether the post-nasal context contributes to some systematic pattern of temporal nasalization. That vowel length may play some role for the emergence of contrastive vowel nasalization is suggested by the vowel length parameter proposed by Hajek (1997) and Hajek and Maeda (2000). This parameter implies that in the development of contrastive vowel nasalization, long vowels are always affected first before short vowels. The approach does not explicitly state that this is due to increased articulatory vowel nasalization induced by the speaker in particular, as findings from perceptual studies are also compatible with such an account (Delattre and Monnot, 1968; Hajek and Watson, 1998; Whalen and Beddor, 1989). Results from our analyses indicated that in absolute terms, tense vowels tended to exhibit a higher extent of vowel nasalization than lax vowels, which was significant for all vowels in CVNCV contexts and for /a/ and /o/ in the CVNV sequences. Proportionally to the respective vowel length, however, differences in nasalization were less distinct. A clear exception to this was the low vowel /a/: from all vowels tested, tense /a/ was clearly nasalized to the largest extent, thus exhibiting more nasalization than its lax counterpart both in absolute terms and relative to the vowel length. Apparently, velum lowering was initiated earlier compared to the lax vowel and also compared to the other tense vowels. The finding that velum lowering occurred so early in tense /a/ suggests some timing control on the part of the speaker, presuming that (s)he has knowledge about the marginal acoustic effects of introducing nasality to a low vowel. On the part of the listener, in turn, long vowels elicit the percept of nasality more easily than short vowels, which is especially true for the low vowel (Delattre and Monnot, 1968; Hajek and Watson, 1998; Whalen and Beddor, 1989). The combination of introducing nasality in production and perceiving this nasality with increased vowel length may lead to the phenomenon that across languages, long or tense /a/ is often affected first in the development of contrastive vowel nasalization (Hajek, 1997; Hajek and Maeda, 2000).

Next, the post-vocalic context was found to affect vowel nasalization: both tense and lax vowels were nasalized to a larger extent when the nasal was followed by a voiceless stop rather than unstressed vowel, with the exception of tense /a/. Vowel duration, in contrast, was not systematically impacted by the post-nasal context, which suggests that the post-vocalic context plays a key role for the extent of vowel nasalization. This assumption can be accounted for by a phonetic approach (Ohala and Ohala, 1991, 1993): a post-nasal oral stop requires a closed velopharyngeal port to ensure sufficient intra-oral pressure. Thus, the velum is lowered during the vowel in anticipation of a nasal that is in aerodynamic conflicts with the following stop, resulting in some weakening of the nasal. This conflict does not occur in contexts in which the nasal is followed by an unstressed vowel: the nasal stop may be articulated without the velum being constrained to early closure, such that the lowering gesture may be even extended to the unstressed vowel.

Consistent with these results, the temporal extent of the overall velum lowering gesture (OVL) was found to be reduced in CVNCV compared to CVNV sequences. Interestingly, OVL was increased in CVNV contexts, although the vowels were found to be less nasalized. Evidently, the velar gesture started later in the vowel and extended throughout the nasal stop to the following unstressed vowel. In CVNCV sequences, in contrast, the velum lowering gesture began earlier in the vowel, but was shortened overall, suggesting that in these sequences, vowel nasalization was proportionally increased relative to OVL compared to CVNV contexts. It was discussed that such changes in proportionality may cause listeners to re-weight the cues available, such that they start to pay closer attention to vowel nasality due to its proportional increase relative to OVL, implying that the nasal stop is weakened at the same time (cf. Carignan et al., 2021). It was further discussed that the difference in OVL in the two contexts is not perfectly consistent with the concept of a temporally constant lowering gesture that is shifted across segments, as supposed by Beddor (2007, and subsequent work). However, although OVL differed for CVNV vs. CVNCV contexts, data still provided evidence that short nasal segments are preceded by more extensively nasalized vowels and longer nasal segments are preceded by vowels that are less nasalized.

Considering OVL in contexts with tense vs. lax vowels, results indicated an increase of OVL when vowel nasalization was also increased. This was observed for tense vowels in CVNCV contexts and for tense /a/ in CVNV sequences, suggesting that within the specific contexts, OVL was positively correlated to vowel nasalization. These findings are different from those presented by Beddor (2007), who provided evidence for a trade-off relation between vowel nasalization and nasal duration in V:N and VN sequences. Our data (CVNCV and CVNV considered separately) suggest that the post-vocalic nasal interval was not adjusted to the extent of the preceding vowel nasality, because otherwise, OVL differences should have been less distinct between tense and lax vowels.

In summary, vowel nasalization, in absolute terms, was found to be more extensive in tense vowels compared to lax vowels, but not proportional to vowel length. Vowels were nasalized to a greater extent when the nasal was followed by a voiceless stop compared to an unstressed vowel. Moreover, the extent of overall velum lowering was found to differ for the CVNV versus CVNCV contexts as well as for tense versus lax vowels in the CVNCV condition, which is different from the concept of a constant velum lowering gesture that is variably shifted across the segments.

Stress and speaking rate Chapter IV investigated the temporal and spatial amount of vowel nasalization in pre-nasal vowels as well as in /nd/ and /nt/ sequences produced in two different focus conditions, in which the target word was produced with either contrastive or broad focus. For the consonantal sequences, additional data of tongue tip movement patterns were analyzed. Studies on the effect of stress suggest a general enhancement of the tongue tip (Cho and Keating, 2009; Farnetani and Vayra, 1996; Giot, 1977; Kent and Netsell, 1971; Meynadier et al., 1998; Mooshammer et al., 1999; Straka, 1963) but provide inconsistent findings with respect to the velar behaviour (Fougeron, 2001; Krakow, 1993; Vaissière, 1988). Moreover, the impact of the post-nasal voicing context on the velar behaviour in German has been recently addressed by Carignan et al. (2021), who investigated the temporal and spatial aspects of the velum gesture considering /Vnt/ and /Vnd/ contexts. For the /Vnt/ contexts, results suggested some overall shortening of the lowering gesture as well as a decrease of the lowering amplitude compared to the voiced condition. In addition, the duration of vowel nasalization was found to be slightly greater and the nasal stop was a little shorter in the voiceless context. Based on their findings, the authors proposed that the reduced amplitude might have counteracted the slightly earlier timing of the lowering gesture, such that the overall degree of nasalization in the vowel was in fact not increased compared to the voiced context. Instead, the effect of the voiceless stop was primarily apparent within the nasal consonant, which was reduced in its duration and the amplitude of the lowering gesture was diminished during this interval.

Our data add to these findings by providing results on the effect of focus on the temporal and spatial patterns of velum lowering and tongue tip gestures during consonantal sequences consisting of /nt/ and /nd/ and during pre-nasal vowels in CVNCV and CVNC(C) contexts. Considering the consonantal contexts, the velum exhibited a lower position when the target sequences were produced with contrastive focus. Gestural enhancement was also observed for the tongue tip, which was found in a higher position. Moreover, velum and tongue position distinctions between the voicing contexts were explored. While the amplitude of the velum lowering gesture was consistently reduced in /nt/ compared to /nd/ independent of the focus condition (cf. Carignan et al., 2021), the tongue tip showed a lower position in /nd/ than in /nt/ with broad focus but not with contrastive focus.

With respect to vowels, only the velum position was considered, which was not significantly affected by the focus condition at the vowel midpoint. However, the temporal extent of vowel nasalization was clearly longer for tense /a/ when produced with contrastive focus. This was accounted for by the assumption that in both focus conditions, velum lowering in /a/ started before the vowel midpoint, such that the velum exhibited a similar degree of lowering at the vowel midpoint. The onset of the lowering gesture, however, occurred earlier with contrastive focus, leading to a more extended duration of vowel nasalization.

Based on these observations, it was concluded that the nasal consonant is generally more enhanced with focus and probably also with stress. In addition, the fact that stop voicing was clearly found to affect velum lowering independent of the focus condition supports the idea that the voicing of the post-nasal stop may be more relevant for nasal weakening than the stress pattern itself.

In addition to focus, chapter IV was also concerned with the effect of an increased speaking rate on the velum and tongue tip position in post-vocalic single oral and nasal coronal stops. The aim was to explore how the positions of the tongue and velum were impacted by fast speech compared to moderate speech, how the consonant affected the position of the tongue and how the tongue tip and velum gesture deviations between the oral and nasal stop were affected by the speaking rate. In general, while a low position of the velum was clearly evident in the nasal stop by nature, the tongue tip was also found in a lower position in the nasal in both speaking rate conditions. Considering the nasal stop in fast speech, both the velar and lingual gestures showed a reduced amplitude compared to the moderate speaking rate. With the oral stop, only the tongue tip gesture was reduced, whereas no differences were found for the velum position. However, the speaking rate did not affect the extent of the deviation between the oral and nasal stop, neither for the tongue or the velum. Furthermore, the tongue tip deviation between the nasal stop in fast versus moderate speech was comparable to that of the oral stop, whereas for the velum, the speaking rate significantly affected the velum deviation during /n/, but not during /t/. Chapter IV also considered the velocity of velum movement during the nasal stop, which, however, was found to be similar in both speaking rates. The lack of interactions between the consonant and the speaking rate (with the exception of the velum in /n/) was somewhat unexpected: the spatial distance between the oral and nasal stop was suggested to increase with fast speech for the tongue tip gesture, which, however, was not the case, indicating that the lingual gesture during fast /n/ was less reduced than predicted. Similarly, the velum position distances between moderate /n/–/t/ and fast /n/–/t/ were presumed to differ due to less pronounced gesture amplitudes in fast speech. However, the distances were found to be comparable for the two speaking rates.

Furthermore, the finding that the velocity during velum lowering in /n/ was not increased with fast speech was compatible with assumptions from prior research suggesting that with fast speech, speakers may either increase the velocity of the velum movements or reduce the amplitude of the gestures (Kent et al., 1974; Kuehn, 1976). Our findings across more than 30 participants provide evidence for the latter strategy: speakers are more likely to achieve the articulatory targets in fast speech by reducing the velum lowering amplitude instead of increasing the velocity.

Perception Chapter V explored how German listeners perceived vowel nasalization that either varied in its temporal extent within the vowel or was shifted as a constant interval along the vowel and nasal in the synthesized word /baːntə/. All listeners were previous participants of the MRI production study. Based on prior studies involving American English listeners (Beddor, 2007, 2009), the prediction was tested whether listeners were more sensitive to alterations of the overall duration of nasalization and less sensitive to a constant nasalized interval that was variably related to the vowel and nasal stop. The goal was to figure out whether listeners tended to treat the nasalized portions in the nasal and the vowel as perceptually equivalent. As a main finding, and contrary to the presumption, listeners showed better performances in discriminating stimuli with a variably distributed constant portion than stimuli with temporal variations of nasalization. It was considered that by shifting the constant interval more into the vowel, listeners might have perceived slight alterations in the silent intervals between the vowel and the oral stop. These silent intervals were then interpreted as variations of a more extended oral stop indicating a longer closure phase of the oral stop /t/, or, alternatively, as some weakening of the nasal stop. It was concluded that listeners probably rely more on modifications they are generally familiar with as native speakers when making a perceptual decision.

In general, this experiment provided some basic insights into the fine details listeners are able to detect when confronted with small modifications of the vowel and the following consonants. More precisely, the study explored how much temporal difference in vowel nasalization (or rather in the silent interval) was required for listeners to differentiate between the stimuli presented. This was achieved by determining listeners' individual perception thresholds, which is different from other methods that are based on relatively coarse-grained steps on a continua.

In addition to testing individual perceptual thresholds of vowel nasalization, it was further investigated whether a correlation was evident between the speakers' specific coarticulatory patterns in production and their sensitivity to fine acoustic details in perception. Prior research has proposed such a relationship (Beddor, 2015; Zellou, 2017) by suggesting that speakers who make use of extensive vowel nasalization during speech are also more sensitive to these acoustic cues in perception. Our data, however, do not support such an account. Although speakers showed individual extents of vowel nasalization, no correlation to their perceptual performances could be attested. It was proposed that the variation patterns of the German speakers might have been too inconspicuous to systematically affect the perceptual sensitivity. Moreover, listeners may generally show greater flexibility in perception than in production due to constant perceptual adjustments to the linguistic environment which they are exposed to (cf. Beddor, 2015). This may explain why a clear pattern was evident in the perception test across the listeners, whereas no relationship was found between individual production and perception performances.

Conclusion and implications The aim of this thesis was to explore the basic patterns of velar behaviour during fluent speech in a language with no contrastive nasal vowels. Some of the findings add to prior assumptions about the gestural amplitude of the soft palate and its timing of lowering during vowel-nasal sequences, whereas others indicate different results than suggested by previous studies.

The German data generally reflect the special role of the low vowel /a/ during nasalization, especially the role of tense /a/. Overall, findings are compatible with accounts suggesting that the low vowel becomes preferentially nasalized in the evolution of contrastive vowel nasalization, especially if it is followed by a nasal stop (Chen, 1973; Ohala, 1975; Ruhlen, 1973). Not all of these approaches consider articulatory mechanisms as the primary reason but instead propose that perceptual factors play a key role (Goddard, 1965, 1971; Matisoff, 1975; Busà and Ohala, 1995). These approaches, however, may be interrelated: if the speaker lowers the velum on purpose, presuming that the introduced nasality will not be perceived as unnatural, the palatoglossus may facilitate the lowering gesture. On the other hand, if the palatoglossus is primarily responsible for the low velum position during /a/, the acoustic result would still be acceptable for the speaker without the need to go against the lowering gesture. In oral contexts, in turn, the velum is raised due to sufficient activity of the levator muscle. However, data from other languages report some velum lowering for /a/ in the oral context as well (Amelot and Rossato, 2006, 2007; Henderson, 1984), which suggests that the articulatory and perceptual interplay proposed here is language-specific. Although the findings presented in this thesis give a respectable impression on the impact of vowel height on velum position, it must be pointed out that they are based on one single point in the vowel (the midpoint), displaying only a small part of information that is actually available in the original image data. For future work, it will thus be helpful to consider methods such as FPCA that are becoming increasingly popular for analyzing velum movement trajectories over the whole course of the vowel (e.g. Gubian et al., 2019; Cronenberg et al., 2020).

Besides the intrinsic vowel properties, the consonantal context is also considered a crucial factor for the extent of vowel nasalization. In our data, vowels were more nasalized when the post-vocalic nasal was followed by a voiceless stop rather than unstressed vowel, which is in agreement with numerous reports about the effect of a post-nasal voiceless stop on nasal shortening and vowel nasalization (Beddor, 2009; Beddor et al., 2013; Malécot, 1960; Busà and Ohala, 1995). At the same time, the interval of overall velum lowering was reduced if the nasal was followed by a voiceless stop compared to a post-nasal vowel, which is compatible with the assumption that the nasal is weakened due to incompatible articulatory and acoustic requirements of these two consecutive consonants. Thus, data support the view that vowel nasalization and nasal duration show systematical variations that depend on the post-nasal context. However, the clear differences in the overall lowering duration between the contexts also suggest that the concept of a trading relationship in the sense of a constant lowering gesture that is shifted along vowel-nasal sequences (c.f. Beddor, 2007, 2009) may not be transferred to all kinds of context comparisons by implication.

The impact of the post-nasal context also became visible in the experiments exploring the effect of contrastive focus on the velum and tongue tip behaviour. With broad focus, the tongue tip and velum were in a lower position when the nasal was followed by a voiced rather than voiceless stop, suggesting a higher degree of velum lowering but less enhancement of the tongue tip gesture in the voiced condition. With contrastive focus, this pattern was evident only for the velum. It is remarkable that even in a language without heavily nasalized vowels like Standard German, the voicing effect on the soft palate can be observed. Although the consonantal data merely considered the spatial amount of velum lowering across the nasal and oral stop, it is reasonable to assume that the temporal duration of the nasal is probably affected as well, especially if one considers the results on the overall lowering gesture which was generally reduced in sequences with a post-nasal stop.

The focus pattern itself seems to play a negligible role for velum lowering during vowels. Although our data showed tendencies for some position enhancements, i.e. a higher velum for high vowels and a lower velum for the low vowel, most of these differences were not reported as significant. This result is contrary to studies reporting enhancement or general lowering effects on the velum during vowels that are in a stressed position (e.g. Krakow, 1993). In contrast to the vowels, the velum showed a clearly lower position in consonantal sequences that were produced with contrastive focus. This is remarkable, as findings from prior research are inconsistent with respect to stress effects on the velum position during consonants (Fougeron, 2001; Krakow, 1993; Vaissière, 1988). Taking into account that our data are based on a relatively large number of participants, the finding that the velum exhibited a lower position with contrastive focus may have some implications for the constitution of the nasal itself: as the velum is lowered to a larger extent while the tongue tip induces stronger alveolar contact, it seems unlikely that the nasal stop is weakened with contrastive focus. Correspondingly, such a gestural pattern would also be unlikely to elicit nasal loss. Thus, our data suggest that the degree of vowel nasalization is probably more related to other factors than focus or stress.

Moreover, in contrast to previous studies reporting large speaker variability in fast speech with respect to velum movements, our data revealed that across speakers, the velum gesture was generally reduced when the speaking rate was increased, while the movement velocity was unaffected. This, in turn, provides evidence for a relatively consistent strategy across speakers for reaching the articulatory target in fast speech, which is the reduction of the velum lowering gesture. Such a reduction of velum lowering is usually accompanied by a narrower velopharyngeal port and thus by less nasal airflow, which may come along with less nasalization of the preceding vowel as well. However, reducing the lowering gesture does not exclude the presence of at least some nasal airflow. Given the fact that the tongue tip gesture is also reduced in its amplitude, initiating less contact with the alveolar ridge, this may have some implications for the articulatory and acoustical outcome: such a sound may resemble the articulatory and acoustic properties of a nasal flap, or, in the more extreme form, even a vowel-like nasal sound. Future research is necessary to investigate the acoustic and perceptual consequences of the interplay between the lingual (non-)contact with the alveolar ridge and the opening of the velopharyngeal port. In addition, instead of relying on one maximum or minimum signal value per participant, aggregated time trajectories may be helpful for a better understanding of the gestural interaction throughout the specific consonants when uttered with different speech rates.

Further investigation would also be required with respect to the perception of vowel nasalization in German compared to other languages with and without contrastive nasal vowels. The staircase method presented in this thesis allows for exploring the fine-detailed perception skills of individual listeners. The initial goal of the perception experiment was to test the hypotheses that listeners show perceived equivalence between stimuli with a constant but variably distributed nasal portion (Beddor, 2007, 2009; Beddor et al., 2013) and also that speakers' individual usage of vowel nasalization during articulation is correlated with their performance in perception (Beddor, 2015; Zellou, 2017). Our results, however, do not suggest a systematic relation between individual speakers' nasalization patterns in production and their sensitivity to this cue in perception. Moreover, data revealed that the German listeners appeared to be quite insensitive to nasalization variations in the vowel, but importantly, that they reliably noticed marginal differences in the post-vocalic contexts. This finding is in accordance with the general assumption that the perception of specific cues is highly language-specific: German listeners may not rely on cues that do not consistently occur in their linguistic environment but they are accurate perceivers of those cues they are familiar with based on their experience with this language. Future experiments may explore how much temporal nasalization is necessary for a vowel to be perceived as nasal in a larger group of participants and also in different language groups. It would also be interesting to explore how listeners' performances are affected by a post-nasal voiced or voiceless stop, given that a nasal stop shares more acoustic characteristics with a voiced rather than voiceless stop.

In summary, while the experiments in this thesis add to previous work on the characteristic behaviour of the velum lowering gesture during fluent speech, they also provide the basis for future research on comparisons of different languages with respect to articulatory patterns of velum lowering and perceptual sensitivity to vowel nasalization.

Zusammenfassung

Über ein Fünftel der attestierten Sprachen dieser Welt weisen kontrastive Nasalvokale auf, d.h. Vokale, bei denen das Hinzufügen von Nasalität die Bedeutung eines Wortes ändern kann (Maddieson, 1984, 2007). Für viele dieser Sprachen gilt jedoch, dass die Nasalvokale nicht schon in den frühesten Sprachstufen Bestandteil des Lautinventars waren, sondern, dass sie sich aus Lautsequenzen aus früheren Sprachstufen entwickelt haben, insbesondere aus solchen, in denen dem Vokal ein nasaler Konsonant folgte. Dabei, so die Annahme, war der Vokal zunächst lediglich antizipatorisch nasaliert, was sich mit der Zeit jedoch verstärkte, sodass die Nasalität schließlich nur noch mit dem Vokal assoziiert war und der Nasal vollständig wegfiel (Chen, 1972; Ferguson, 1963; Hajek, 1997; Ruhlen, 1973; Schourup, 1973). Einige Beispiele lassen sich anhand der romanischen Sprachen zeigen, wie etwa lat. tempus > frz. temps [tã] ‘Zeit’ oder lat. manus > port. maõ [mɐ̃w̃] ‘Hand’.

Um diese Art von Lautwandel besser verstehen zu können, wurde viel hinsichtlich der Frage geforscht, wie genau die beteiligten Artikulationsorgane in solchen Sequenzen miteinander zeitlich und auch räumlich interagieren. Obwohl einige Erkenntnisse dazu vorliegen, ist die Frage, welche artikulatorischen und perzeptiven Faktoren es sind, die einen solchen Lautwandel auslösen können, nach wie vor nicht eindeutig geklärt. Dabei scheint die koartikulatorische Vokalnasalierung eine zentrale Rolle zu spielen, bei der das Gaumensegel maßgeblich beteiligt ist. In Sprachen, die keine kontrastiven Nasalvokale aufweisen, werden Vokale in oralem Kontext normalerweise mit einem gehobenen Gaumensegel produziert. Somit fließt nahezu der gesamte Luftstrom, der für die Erzeugung von Lauten nötig ist, ausschließlich durch den Mund-Rachen-Raum. Bei nasalen Konsonanten jedoch bilden die Lippen oder die Zunge einen oralen Verschluss bei gleichzeitiger Senkung des Gaumensegels: Die Luft tritt ausschließlich über den Nasen-Rachen-Raum durch die Nase nach außen. Werden nun aneinanderhängende Laute artikuliert, wie es in natürlicher Sprache der Fall ist, beeinflussen sich die Gesten der einzelnen Laute stets gegenseitig: Es kommt zu koartikulatorischen Effekten. Somit ist in einem Wort wie *Bahn* eine gewisse Nasalierung des Vokals, ausgelöst durch ein antizipatorisch gesenktes Gaumensegel, zu erwarten, in *Bad* hingegen weniger.

Sprachen unterscheiden sich darin, in welchem zeitlichen Umfang sie von koartikulatorischer Vokalnasalierung Gebrauch machen: Während für manche Sprachen eine ausgeprägte Nasalierung attestiert ist, weisen andere nur schwach nasalierte Vokale auf

(Clumeck, 1976; Malécot, 1960; Solé, 1992). Allerdings gibt es auch linguistische Faktoren, die sprachübergreifend die Nasalierung von Vokalen zu begünstigen scheinen. Von einigen dieser Faktoren wird angenommen, dass sie bei der Entwicklung von koartikulatorischer Vokalnasalierung hin zu einem kontrastiven Nasalvokal von zentraler Bedeutung sind. Dazu zählen intrinsische Faktoren wie die Höhe und Länge des Vokals sowie der Lautkontext, in dem sich der Vokal befindet, wozu neben den Nasalen selbst auch post-nasale Obstruenten zählen. Hinzu kommen extrinsische Faktoren wie Intonation und Sprechgeschwindigkeit. Neben den artikulatorischen Voraussetzungen wurde in der Forschung darüber hinaus die Perzeption von Vokalnasalierung untersucht, d.h. wie genau der Vokal und dessen Kontext beschaffen sein müssen, damit der Vokal von Hörerinnen und Hörern als nasal wahrgenommen wird. Auch spielen für die Perzeption von Nasalierung die Vokalhöhe und -länge eine Rolle (House und Stevens, 1956; Maeda, 1993; Whalen und Beddor, 1989), sowie der den Vokal umgebende konsonantische Kontext (Ali et al., 1971; Beddor, 2009; Beddor und Krakow, 1999; Malécot, 1960; Ohala und Ohala, 1995).

Ziel der vorliegenden Arbeit war es, das komplexe Verhalten des Gaumensegels in einer Sprache zu erforschen, die weder über Nasalvokale noch über stark nasalierte Vokale verfügt, um somit weitere grundlegende Erkenntnisse zu dem aktuellen Stand der Forschung bezüglich der Funktionsweise des Gaumensegels und der linguistischen Faktoren, welche die Senkung und Hebung des Gaumensegels beeinflussen, beizutragen. Das besondere an dieser Studie ist zum einen die für eine phonetische Artikulationsstudie relativ große Anzahl an Probandinnen und Probanden und zum anderen das bildgebende Verfahren, das für die Datenerhebung eingesetzt wurde. Mithilfe modernster Echtzeit-MRT-Technologie wurden hochqualitative Aufnahmen mit einer räumlichen Auflösung von 1.4x1.4 mm bei einer Schichtdicke von 8 mm sowie mit einer zeitlichen Auflösung von 19.98 ms erstellt, was 50.05 Bildern pro Sekunde entspricht. Diese Messungen waren möglich dank einer Kooperation mit der Göttinger MRT-Forschungsgruppe am Max-Planck-Institut für Biophysikalische Chemie unter der Leitung von Jens Frahm. Insgesamt 36 Muttersprachlerinnen und Muttersprachler des Standarddeutschen nahmen an dieser Studie teil; allerdings enthalten die Analysen der vorliegenden Arbeit Daten von lediglich 33 Personen aufgrund technischer Schwierigkeiten bei der Registrierung der Bilder für die Bildanalyse. Dass das Standarddeutsche – anders als manche Dialekte des Deutschen – keine stark nasalierten Vokale aufweist, macht diese Untersuchung besonders interessant: Sollten sprachübergreifend artikulatorische Faktoren existieren, welche die koartikulatorische Nasalierung von Vokalen begünstigen, wären sie vermutlich auch in einer Sprache zu beobachten, in der es aktuell keinerlei Hinweise auf einen möglichen Lautwandel hin zu kontrastiven Nasalvokalen gibt.

Die MRT-Studie wurde ursprünglich in Hinsicht auf mehrere Fragestellungen konzipiert, weshalb das Korpus dementsprechend umfangreich ausfällt. Die Zielwörter wurden dabei in Kontextsätze eingebettet, die von den Probandinnen und Probanden während der MRT-

Messung vorgelesen wurden. Insgesamt umfasst das Korpus 152 Zielwörter, wobei diese mehrfach in unterschiedlichen prosodischen Kontexten wiederholt wurden. Pro Proband wurden somit jeweils ca. 350 Sätze geäußert. Zusätzlich wurden synchronisierte akustische Aufnahmen erstellt, die nach Satz, Wort, Vokal und post-vokalischem Konsonanten manuell segmentiert und für die Datenanalyse genutzt wurden. Die Bildanalyse erfolgte über die Software Matlab (Details in Carignan et al., 2020 und Carignan et al., 2021), wofür zunächst die Bilder jedes einzelnen Probanden registriert wurden. Um die Bewegung des Gaumensegels nachzuvollziehen, wurde um den Bereich des weichen Gaumens herum eine *Region of Interest* erstellt, deren enthaltene Pixel als Dimensionen für eine Hauptkomponentenanalyse dienten. Auf Grundlage der ersten Hauptkomponente konnte ein zeitabhängiges Signal erstellt werden, welches sich als Position des Gaumensegels über die Zeit hinweg interpretieren ließ. Zusätzlich wurden von dem zeitabhängigen Signal kinematische Parameter abgeleitet, mit denen gewisse Zeitpunkte während der Bewegung bestimmt werden konnten, wie etwa der Beginn der Senkung des Gaumensegels oder der Zeitpunkt der maximalen Geschwindigkeit während der Senkung und Hebung. Darüber hinaus wurde für den übrigen Vokaltrakt ein weiteres Analyseverfahren angewendet, das im Wesentlichen auf Veränderungen der Pixelintensitäten in vordefinierten Bereichen beruhte. Hohe Pixelintensitäten in einer bestimmten Region wiesen somit auf die Präsenz eines bestimmten Artikulators hin.

Neben den grundsätzlichen Analysen der Position des Gaumensegels in Abhängigkeit von der Beschaffenheit des Vokals und des konsonantischen Kontexts wurden darüber hinaus auch solche Hypothesen untersucht, die ein Wechselspiel zwischen der Dauer der Vokalnasalierung und des nachfolgenden Nasals in Betracht zogen. Im Mittelpunkt stand dabei das Lautwandelmodell von Beddor (2007, 2009), in welchem sie postuliert, dass Hörer aufmerksame Perzipienten der akustischen Konsequenzen von Koartikulation seien, aber dennoch zu einer anderen Interpretation und Repräsentation kämen als ursprünglich vom Sprecher beabsichtigt. Dies begründe sich darin, dass der zeitliche und räumliche Umfang von Koartikulation stark variiere, da der koartikulatorische Effekt und dessen Quelle unterschiedlich miteinander interagierten. Diese Variation wiederum führe auf Seiten des Hörers zu unterschiedlichen perzeptiven Gewichtungen der akustischen Eigenschaften, sodass das perzipierte akustische Signal mit verschiedenen phonologischen Analysen kompatibel sei, was dazu führe, dass die Grammatiken von Hörern hörerspezifisch seien. Die Grundannahmen dieses Modells veranschaulicht Beddor anhand mehrerer Produktions- und Perzeptionsexperimente, welche insbesondere Vokalnasalierung in unterschiedlichen Kontexten im amerikanischen Englischen untersuchen. Dabei untermauern die Daten des Artikulationsexperiments ihre Idee, dass der koartikulatorische Effekt eine Art Handelsbeziehung mit seiner Quelle eingeht: Einem Vokal, der zeitlich stärker nasaliert ist, folgt ein kürzerer Nasal; umgekehrt wird ein nur leicht nasalierter Vokal von einem längeren Nasal gefolgt. Darüber hinaus ist auch ein Zusammenhang mit der Stimmhaftigkeit des post-nasalen Plosivs erkennbar: Der Vokal ist länger nasaliert und der Nasal kürzer, wenn der folgende Plosiv stimmlos ist. Diese Beobachtung wurde bereits in vorangegangener Forschung mehrfach erwähnt (Malécot, 1960; Ohala und Ohala, 1991; 1993; Sampson, 1999; Sefton und Beddor, 2005). Als Begründung lassen sich aerodynamische und auditive Aspekte heranziehen: Während ein gewisser Grad an Stimmhaftigkeit durchaus mit einer leichten Senkung des Gaumensegels vereinbar ist, wird für einen stimmlosen Plosiv ein möglichst dichter Verschluss des velopharyngalen Durchgangs benötigt, um einen ausreichend großen intra-oralen Luftdruck für die Artikulation des Plosivs zu gewährleisten (Ohala und Ohala, 1991; 1993). Dieser Mechanismus wiederum begünstigt ein frühes Absenken des Gaumensegels bereits während des Vokals sowie eine zeitliche Kürzung des Nasals. Basierend auf der Beobachtung, dass die Dauer der Vokalnasalierung und die Nasaldauer miteinander korrelieren, formuliert Beddor (2007, 2009) das Konzept, dass die Senkungsgeste des Gaumensegels in beiden Stimmhaftigkeitskontexten zeitlich etwa konstant bleibt, dass sie jedoch in Abhängigkeit vom Folgekontext variabel "verschoben" werden kann. Diese koartikulatorische Variation werde von Hörern wahrgenommen, jedoch unterschiedlich interpretiert. Eine Möglichkeit dieser Interpretation bestehe darin, dass zumindest einige Hörer vor allem die kovariierende Information nutzen, um das Signal zu interpretieren. Demnach würden sich diese Hörer hauptsächlich an der Information "nasal" orientieren, weniger jedoch an der spezifischen Verteilung der Nasalität. Diese Annahme wurde von Ergebnissen aus diversen Perzeptionsexperimenten gestützt, in denen amerikanische Hörerinnen und Hörer Wortpaare mit unterschiedlichen Nasalierungsmustern weniger gut differenzieren konnten, wenn sich diese zwar in der Verteilung der Nasalität auf den Vokal und Nasal unterschieden, nicht jedoch in der Gesamtdauer der Nasalierung. Des Weiteren wurde gezeigt, dass manchen Hörern bereits ein nasalierter Vokal in einem Wort ausreichte, um es als eines zu identifizieren, das einen nasalen Konsonanten enthielt, wohingegen andere Hörer denselben akustischen Stimulus als ein Wort ohne einen solchen Nasal identifizierten. Diese und ähnliche Ergebnisse untermauern Beddors Konzept, in welchem unterschiedliche Hörer bei gleichem akustischen Input zu unterschiedlichen Interpretationen gelangen. Allgemeiner formuliert können Hörer, die sich an der koartikulatorischen Quelle orientieren, als "konservativ" klassifiziert werden und solche, die sich eher an dem koartikulatorischen Effekt orientieren, als "innovativ" (Beddor, 2012, 2015). Demnach könnten solche Innovationen mit der Zeit in der eigenen Produktion oder auch durch Erwartungen an auftretende koartikulatorische Muster manifestiert werden, wodurch bestimmte Lautwandelprozesse ermöglicht werden könnten.

Die Kernaussagen dieses Modells wurden - neben weiteren grundlegenden Fragestellungen - in dieser Arbeit anhand der erhobenen MRT-Daten in Kombination mit den akustischen Daten überprüft. Dabei wurde untersucht, ob sich auch im Standarddeutschen Hinweise auf eine systematische Handelsbeziehung zwischen der Nasalierung des Vokals und der Nasaldauer feststellen ließen, möglicherweise auch in Abhängigkeit vom post-nasalen Kontext.

Der vorliegenden Arbeit liegt die folgende Struktur zugrunde: Kapitel II untersuchte zunächst den Einfluß der Zungenlage in unterschiedlichen Vokalen auf die Position des Gaumensegels. Dabei wurden die Vokale sowohl von nasalem als auch oralem Kontext gefolgt. Kapitel III beleuchtete den Einfluss der Vokallänge auf das Ausmaß der Nasalierung sowie den Effekt des post-nasalen Kontexts auf die Gesamtdauer der Senkung des Gaumensegels. Die Rolle von Sprechgeschwindigkeit und kontrastivem Fokus auf die Zungen- und Gaumensegelposition wurde in Kapitel IV thematisiert. Kapitel V beschäftigte sich mit der Perzeption von Vokalnasalierung bei jenen Hörerinnen und Hörern, die zuvor an der MRT-Sprechstudie teilgenommen hatten und widmete sich der Frage, ob eine Korrelation erkennbar war zwischen der eigenen Nasalierung in der Produktion und der Sensitivität bezüglich Vokalnasalierung in der Perzeption. Das vorliegende Kapitel enthält eine Zusammenfassung der Ergebnisse.

Der Einfluss der Vokalhöhe auf die Position des Gaumensegels

Dass bestimmte Vokale prädestiniert dafür sind, stärker nasaliert zu werden, wurde oft vermutet, wobei für die Entwicklung von kontrastiven Nasalvokalen insbesondere der tiefe Vokal /a/ eine zentrale Rolle zu spielen scheint (Chen, 1972; Hajek, 1992; 1997; Hombert et al., 1979; Whalen und Beddor, 1989). Bereits in den Anfängen der phonetischen Forschung wurde beobachtet, dass das Gaumensegel während der Artikulation verschiedener Vokale unterschiedliche Positionen aufweist (Brucke, 1876; Passavant, 1869). Im Laufe der Zeit kamen mehrere Studien mit unterschiedlichen Messmethoden zu recht ähnlichen Ergebnissen: Das Gaumensegel befindet sich in einer höheren Position in hohen Vokalen wie etwa /i/, /e/ oder /u/, während es bei der Artikulation von /a/ auffallend gesenkt ist. Dies war teilweise auch dann der Fall, wenn der Vokal von oralem Kontext umgeben war (Bell-Berti et al., 1979; Clumeck, 1976; Kuehn, 1976; Lubker, 1968; Moll, 1962; Rossato et al., 2003). Allgemein wird dieses Muster zweierlei Erklärungsansätzen zugeschrieben: Zum einen besteht eine muskuläre Verbindung zwischen dem Rand der Zunge und der Gaumenaponeurose, einem starken Bindegewebe, welches die Grundlage für das Gaumensegel darstellt. Diese Verbindung unter dem Namen *Palatoglossus* dient hauptsächlich dazu, während des Schluckaktes den Zungengrund zu heben und die Schlundenge (*isthmus faucium*) zu verkleinern. Es wird vermutet, dass die Zungenlage als Nebeneffekt jedoch auch eine mechanische Wirkung auf das Gaumensegel ausübt: Aufgrund einer zurückgezogenen und eher tief liegenden Zunge während des Vokals /a/ wird das Gaumensegel über die Palatoglossus-Verbindung ein Stück weit hinuntergezogen (Dixit et al., 1987; Moll und Shriner 1967). Dies würde die so oft berichteten Positionsunterschiede in den einzelnen Vokalen erklären. Alternativ zu dem artikulatorischen Ansatz ließe sich die unterschiedliche Position des Gaumensegels auch perzeptiv erklären: In einigen Experimenten nahmen Hörerinnen und Hörer hohe Vokale bereits als nasal wahr, wenn diese mit einem geringen Grad an Nasalierung präsentiert wurden. Im Gegensatz dazu war für den tiefen Vokal /a/ ein wesentlich höherer Nasalierungsgrad erforderlich, um als nasal bewertet zu werden (House und Stevens, 1956; Lubker, 1968; Maeda, 1989; 1993; Ohala, 1975). Vereinzelt kamen andere Studien jedoch auch zu anderen Ergebnissen (Ali et al., 1971; Lintz und Sherman, 1961). Die Perzeptionsdaten lassen die Überlegung zu, dass sich Sprecher der akustischen Konsequenzen eines gesenkten Gaumensegels bei hohen Vokalen bewusst sein könnten und diesen Konsequenzen durch eine hohe Position vorbeugen. Ein Entgegenwirken wäre im Falle des tiefen Vokals hingegen nicht erforderlich, da dieser akustisch gesehen mehr Nasalierung toleriert, bevor er als nasal wahrgenommen wird.

Das erste Experiment dieser Arbeit (Abschnitt 2.1.2) diente dazu, mithilfe der MRT-basierten Daten festzustellen, ob und in welchem Umfang die bisherige Studienlage auch auf das Deutsche übertragen werden kann. Dafür wurden Zielwörter mit den Lautsequenzen CVNV und CVCV analysiert, wobei der erste Vokal gespannt oder ungespannt war und der zweite stets ungespannt. Die entsprechenden Daten bezogen sich dabei auf die Position des Gaumensegels zum zeitlichen Mittelpunkt des ersten Vokals. Wie erwartet war ein deutlicher Unterschied der Position des Gaumensegels zwischen den beiden Kontexten zu erkennen: Alle CVNV Strukturen wiesen ein leicht gesenktes Gaumensegel in den Vokalen auf, was bei den CVCV Sequenzen nicht in vergleichbarer Form der Fall war. In den CVNV Kontexten mit gespannten Vokalen war zudem ersichtlich, was bereits für andere Sprachen beschrieben worden war: Während der Artikulation des Vokals /a/ war das Gaumensegel deutlich gesenkt im Vergleich zu den anderen Vokalen /i/, /o/, /ø/, /u/ und /y/. Für die ungespannten Vokale war dies jedoch nicht der Fall. Des Weiteren konnte auch im oralen Kontext kein genereller Unterschied in der Position des Gaumensegels zwischen den einzelnen Vokalen festgestellt werden, was im Widerspruch zu einigen Befunden aus anderen Sprachen steht (Bell-Berti et al., 1979; Clumeck, 1976; Kuehn, 1976; Lubker, 1968; Moll, 1962; Rossato et al., 2003). Darüber hinaus war auch kein Positionsunterschied zwischen den vorderen und hinteren Vokalen erkennbar, was eher gegen die Annahme spricht, dass räumliche Beschränkungen ein Absenken des Gaumensegels verhindern, wie es bei einer hohen und zurückgezogenen Zunge in /u/ der Fall wäre. Da aber kein Unterschied zu /i/ ersichtlich war, spielt die horizontale Lage der Zunge vermutlich eine untergeordnete Rolle für die Senkung des Gaumensegels. Etwas unerwartet mag das Resultat sein, dass /a/ im oralen Kontext nicht mit einem weiter gesenkten Gaumensegel artikuliert wurde als die übrigen Vokale und auch, dass kein klarer Unterschied in den ungespannten Vokalen ersichtlich war. Als Überlegung wurde eine antagonistische Wirkungsweise des Palatoglossus und des

Levator Palatini herangezogen (vgl. Kuehn und Azzam, 1978, S. 356), einem Muskel, der hauptsächlich für die Hebung des Gaumensegels zuständig ist: In oralen Kontexten ist der Levator Palatini aktiv und hebt das Gaumensegel an, sodass seine Zugkraft größer ist als die des Palatoglossus. In nasalen Kontexten hingegen ist der Levator Palatini – in Antizipation des Nasals – etwas weniger angespannt, sodass die Zugkraft des Palatoglossus zum Tragen kommt: Das Gaumensegel wird mechanisch gesenkt. In welcher Form Sprecher diese Muskeln miteinander interagieren lassen, ist vermutlich sprachspezifisch, wie die Ergebnisse aus anderen Sprachen zeigen, in denen das Gaumensegel auch in oralen Kontexten gesenkt ist.

Dass sich in den hier präsentierten Daten die Position des Gaumensegels in den ungespannten Vokalen kaum unterschied, könnte sowohl mit der kürzeren Vokaldauer als auch mit der spezifischen Art der Artikulation zusammenhängen: Da im Deutschen ungespannte Vokale generell kürzer als gespannte Vokale sind, bleibt möglicherweise nicht genügend Zeit, um ausreichende Unterschiede in der Position des Gaumensegels auszubilden. Andererseits sollte ein rein mechanischer Effekt auch in kürzeren Vokalen zu erkennen sein. Ungespannte Vokale jedoch werden allgemein etwas zentraler gebildet als gespannte Vokale (vgl. Hoole und Mooshammer, 2002). Somit könnte eine etwas zentralere Zungenposition ausreichen, um den mechanischen Effekt der Palatoglossus-Verbindung zu unterbinden, anders als es bei einem peripher gelegenen gespannten Vokal der Fall wäre.

Darüber hinaus sind die Daten jedoch ebenso mit dem perzeptiven Erklärungsansatz kompatibel: Ausgehend von der Annahme, dass sich Sprecher der akustischen Konsequenzen ihrer Lautproduktion bewusst sind, wäre es denkbar, dass das Gaumensegel in besonderem Maße im gespannten Vokal /a/ gesenkt wird, weil dieser Vokal akustisch gesehen einen hohen Grad an Nasalierung toleriert, bevor er als nasal perzipiert wird. Somit könnte der Sprecher die für den Nasal erforderliche Senkungsgeste des Gaumensegels frühzeitig im Vokal antizipieren, ohne diesen mit einer unnatürlich wirkenden Nasalierung zu versehen.

Zusammengefasst lässt sich festhalten, dass auch die in dieser Arbeit diskutierten Daten auf die besondere Rolle des tiefen Vokals /a/ schließen lassen, wenn auch nicht in demselben Maß wie es für andere Sprachen beschrieben ist.

Der Einfluss der Vokallänge und des post-nasalen Lautkontexts auf die Dauer von Vokalnasalierung

Ein weiterer Faktor, der vermutlich im Zusammenhang mit der Entwicklung von kontrastiven Nasalvokalen steht, ist die Länge eines Vokals. In allen Sprachen, in denen sich lange und kurze Nasalvokale aus ehemals oralen Vokalen entwickelt haben, waren stets zuerst die langen Vokale von Nasalierung betroffen (Hajek, 1997; Hajek und Maeda, 2000). Der zeitliche Umfang von Nasalierung in Vokalen unterschiedlicher Länge wurde in einigen Studien erforscht, mit dem Ergebnis, dass verschiedene Sprachen spezifische Muster bezüglich der Senkung des Gaumensegels aufweisen (Clumeck, 1976; Solé, 1992). So werden beispielsweise im amerikanischen Englischen die Vokale bereits sehr früh nasaliert, während die Senkungsgeste im Spanischen erst kurz vor dem nasalen Konsonanten beginnt, unabhängig von der Vokallänge (Solé, 1992). Dies wird mit der Annahme verbunden, dass in der einen Sprache Nasalierung bereits zu einem gewissen Teil phonologisiert, d.h. fester Bestandteil des Vokals ist, in anderen Sprachen dagegen die Senkung lediglich als obligatorische Geste vor dem Nasal erfolgt (Solé, 1992).

Abgesehen von den Artikulationsexperimenten wurden vermehrt Perzeptionsstudien durchgeführt, welche die Rolle der Vokallänge bei der Beurteilung des Nasalitätsgrades untersuchten. Viele von ihnen stellten einen klaren Zusammenhang zwischen der Vokallänge und der wahrgenommenen Nasalität fest: je länger ein Vokal, desto eher wird er als nasal perzipiert (Delattre und Monnot, 1968; Hajek und Watson, 1998; Whalen und Beddor, 1989). Darüber hinaus gibt es Anhaltspunkte dafür, dass der nasale Konsonant sowie der post-nasale Kontext einen Einfluss auf die Dauer der Vokalnasalierung haben: je kürzer der Nasal, desto länger ist der Vokal nasaliert. Eine Schwächung des Nasals wiederum wird insbesondere dann begünstigt, wenn ein stimmloser Obstruent folgt (Beddor, 2007; 2009; Clumeck, 1976; Malécot, 1960; Sefton und Beddor, 2005).

Die Fragestellungen nach der Rolle der Vokallänge und des post-nasalen Kontexts für das zeitliche Ausmaß der Nasalierung wurden für das Deutsche in den Experimenten II und III anhand der erhobenen MRT-Daten näher untersucht. Dafür wurden Zielwörter analysiert, die gespannte und ungespannte prä-nasale Vokale in CVNV und CVNCV Sequenzen enthielten. Mittels kinematischer und akustischer Analysen wurden die Zeitpunkte der maximalen Geschwindigkeit in der Gaumensegelsenkung und -hebung sowie die akustischen Grenzen des Vokalbeginns und -endes bestimmt. Als Vokalnasalierung galt demnach das Intervall zwischen dem Zeitpunkt der maximalen Geschwindigkeit während der Senkung und dem Vokalende. Die Ergebnisse zeigten, dass sich ungespannte und gespannte Vokale zwar tendenziell in der Nasalierungsdauer unterschieden, jedoch war dieser Unterschied in CVNV Kontexten nur für /a/ und /o/ signifikant. In CVNCV Sequenzen hingegen waren alle gespannten Vokale signifikant länger nasaliert. Proportional zur Vokallänge gesehen fielen die Unterschiede allerdings recht gering aus. Sowohl in den CVNV als auch CVNCV Sequenzen wies lediglich der Vokal /a/ einen proportional deutlich längeren Anteil an Nasalierung in der gespannten Variante auf. Demnach öffnete das Gaumensegel während /a/ wesentlich früher als im ungespannten Vokal, aber auch früher als in den anderen gespannten Vokalen, was dazu passt, dass es zum zeitlichen Mittelpunkt des Vokals räumlich weiter gesenkt war als bei den anderen Vokalen (wie in Experiment I gezeigt wurde). Mit Ausnahme von /a/ sprechen die Ergebnisse dafür, dass die Senkung des Gaumensegels im Deutschen relativ zur Vokallänge erfolgt, dass also ungespannte und gespannte Vokale zu ähnlichen Anteilen nasaliert sind. Des Weiteren war hinsichtlich der Dauer von Vokalnasalierung ein deutlicher Unterschied zwischen den post-nasalen

Kontexten zu erkennen: Unabhängig von der Gespanntheit der Vokale wiesen diese eine signifikant längere Nasalierung in den CVNCV Sequenzen auf als in den CVNV Sequenzen. Eine Ausnahme bildete hier der gespannte Vokal /a/, der keinen Nasalierungsunterschied zwischen den Kontexten erkennen ließ. Da der konsonantische Kontext keinen Einfluss auf die Vokaldauer hatte (gespannte und ungespannte Vokale für sich betrachtet), wohl aber auf die Dauer der Vokalnasalierung, liegt die Vermutung nahe, dass für die Dauer der Vokalnasalierung neben der Vokallänge vor allem der post-nasalen Kontext von zentraler Bedeutung ist: Hätte ausschließlich die Vokallänge Einfluss auf die Dauer der Nasalierung, dürfte kaum ein Nasalierungsunterschied zwischen den Kontexten zu sehen sein, da die Vokallänge zwischen den Kontexten vergleichbar war.

Bezüglich einer möglichen Korrelation zwischen der Dauer der Vokalnasalierung und des Nasals wurde zunächst diskutiert, dass diese beiden zeitlichen Intervalle in dem vorliegenden Fall nicht korreliert werden sollten, da sie eine gemeinsame manuell definierte akustische Grenze teilten: Die Dauer der Vokalnasalierung war als Intervall zwischen dem Zeitpunkt der maximalen Geschwindigkeit während der Senkung des Gaumensegels und der akustischen Grenze des Vokalendes definiert, während der Nasal dem Intervall zwischen dem Vokalende und dem Zeitpunkt der maximalen Geschwindigkeit während der Hebung des Gaumensegels entsprach. Eine Korrelation zweier Segmente, die eine gemeinsame manuell definierte Grenze teilen, führt jedoch aufgrund unvermeidlicher geringer manueller Messfehler zwangsläufig zu einer negativen Korrelation (Ohala und Lyberg, 1976). Aus diesem Grund wurde die Dauer der Vokalnasalierung nicht in Relation zum Nasal gesetzt, sondern zur Gesamtdauer der Senkung des Gaumensegels, also dem Intervall zwischen den beiden Zeitpunkten der maximalen Geschwindigkeit in der Senkung und Hebung. Zunächst wurde festgehalten, dass sich die Gesamtdauer der Senkung in CVNV Kontexten nur geringfügig hinsichtlich der Gespanntheit der Vokale unterschied: Abgesehen von /a/ war die Gesamtdauer für Kontexte mit gespannten und ungespannten Vokalen vergleichbar, was mit den Ergebnissen hinsichtlich der reinen Vokalnasalierung kompatibel ist. In CVNCV Kontexten hingegen war die Gesamtdauer der Senkungsgeste signifikant länger, wenn die Sequenzen gespannte Vokale enthielten. Des Weiteren hatte der post-nasale Kontext selbst einen signifikanten Einfluss auf die Dauer der Senkungsgeste, welche in CVNV Kontexten wesentlich ausgeprägter war als in CVNCV Sequenzen. Dies galt für fast alle Vokale unabhängig von ihrer Gespanntheit; einzige Ausnahme war hier der gespannte Vokal /y/. Dass die Gesamtdauer der Senkung in CVNV Sequenzen erhöht war, wurde damit begründet, dass in diesem Kontext grundsätzlich keine Notwendigkeit bestand, das Gaumensegel nach dem Nasal wieder zu heben, zumindest nicht so lange bis ein nachfolgender Obstruent (zu Beginn des Folgewortes) dies erforderte. Dementsprechend erstreckte sich die Senkung des Gaumensegels in solchen Kontexten vermutlich auch über den post-nasalen unbetonten Vokal.

Zusammenhängend betrachtet lässt sich aus diesen Daten schließen, dass Vokale in CVNCV Sequenzen zwar länger nasaliert sind als in CVNV Kontexten, die Gesamtdauer der Senkung des Gaumensegels in diesen Sequenzen jedoch gleichzeitig reduziert ist, was zur Folge hat, dass der Vokal in CVNCV Sequenzen proportional zur Gesamtdauer der Senkungsgeste länger nasaliert und der Nasal demnach reduzierter ist als in CVNV Kontexten. Zu einem ähnlichen Ergebnis kamen Carignan et al. (2021), die, basierend auf demselben Gesamtkorpus der vorliegenden Arbeit, das Verhalten des Gaumensegels in /Vnd/ und /Vnt/ Sequenzen untersuchten. Dabei zeigte sich, dass der Vokal in /Vnt/ nur marginal länger nasaliert war als in /Vnd/, die Gesamtgeste der Gaumensegelsenkung in /Vnt/ jedoch generell räumlich und zeitlich reduzierter war. Somit nahm der nasalierte Teil des Vokals in /Vnt/ einen proportional höheren Anteil an der Gesamtgeste ein. Hinsichtlich möglicher Konsequenzen für einen Lautwandelprozess vermuten die Autoren, dass solche Änderungen in der Proportionalität auf der Seite des Perzipienten dazu führen könnten, dass dieser die akustischen Details eines Signals neu gewichtet, sodass die Nasalierung im Vokal an Bedeutung gewinnt.

Des Weiteren lassen die vorliegenden Ergebnisse erkennen, dass gespannte Vokale absolut gesehen tendenziell länger nasaliert sind und die Gesamtdauer der Senkungsgeste in Kontexten mit gespannten Vokalen gleichzeitig zunimmt. Beide Resultate zeigen somit deutliche zeitliche Variationen der Gesamtdauer abhängig vom konsonantischen und vokalischen Kontext und unterscheiden sich von Daten aus anderen Sprachen, für die eine konstante Senkungsgeste postuliert wird, die kontextabhängig über den Vokal und Nasal hinweg verschoben wird (vgl. Beddor, 2007, 2009). Nichtsdestotrotz weisen die Vergleiche zwischen den Kontexten die Tendenz zu einer Art Handelsbeziehung auf: während die velare Geste in CVNCV Sequenzen, verglichen mit CVNV Kontexten, deutlich früher endet, ist der vorangehende Vokal wesentlich länger nasaliert. Hinsichtlich dieses Aspekts sind die Daten mit denen aus vorangegangener Forschung vergleichbar: kurzen Nasalen – deren geringe Dauer mit der Beschaffenheit des Folgekontextes zusammenhängt – gehen zeitlich stark nasalierte Vokale voran und längeren Nasalen gehen Vokale mit zeitlich geringerer Nasalierung voran.

Der Einfluss von Fokus und Sprechgeschwindigkeit auf die Position der Zunge und des Gaumensegels

Dass prosodische Faktoren die artikulatorischen Gesten in ihren zeitlichen und räumlichen Bewegungen beeinflussen, ist gut erforscht. Gegenstand vieler Studien ist insbesondere die Bewegung der Zunge in solchen Lautsequenzen, die mit erhöhtem Sprechtempo artikuliert werden (Adams et al., 1993; Flege, 1988; Gay et al., 1974; Kuehn, 1976; Lindblom, 1963) oder unterschiedliche Betonungsmuster aufweisen (Cho und Keating, 2009; Farnetani und Vayra, 1996; Giot, 1977; Kent und Netsell, 1971; Meynadier et al. 1998; Mooshammer et al., 1999). Dabei ist meist eine Tendenz hin zu einer Reduktion der Zungengeste in unbetonten wie auch in schnell gesprochenen Sequenzen erkennbar. Inwiefern das Gaumensegel von solchen Faktoren beeinflusst wird, ist jedoch weniger gut untersucht. Offenbar zeigen dabei unterschiedliche Sprecher verschiedene Strategien. Während das Gaumensegel bei manchen Sprechern grundsätzlich eine niedrigere Position in betonten verglichen mit unbetonten

Vokalen aufweist, wird bei anderen die vokalintrinsische Position verstärkt: Betonte hohe Vokale werden mit einem noch höheren Gaumensegel artikuliert, während tiefe Vokale ein noch weiter gesenktes Gaumensegel aufweisen (Krakow, 1993). Ähnliches gilt für den Einfluss der Sprechgeschwindigkeit auf das Verhalten des Gaumensegels: Um in schnell gesprochenen Sequenzen das Artikulationsziel zu erreichen, erhöhen manche Sprecher die Geschwindigkeit der Geste, während andere die Amplitude der Bewegung verringern (Kent et al., 1974; Kuehn, 1976).

Die Rolle der prosodischen Faktoren bezüglich des Verhaltens der artikulatorischen Gesten wurde in den Experimenten IV, V und VI (Abschnitte 4.1.2, 4.1.3 und 4.2.2) untersucht. Die Zielwörter wurden in zwei unterschiedlichen Betonungsweisen vorgelesen: In der ersten galt es, das Wort im Rahmen des Kontextsatzes mit allgemeinem Fokus zu artikulieren, ohne besondere Kontrastivität. In der zweiten Betonungsweise sollte das Zielwort stark kontrastiv betont werden. Da jedes Zielwort in beiden Betonungsarten vorkam, war ein Vergleich der Position und der zeitlichen Bewegung des Gaumensegels sowie der Zungenspitze zwischen den Betonungsarten möglich. Experiment IV untersuchte die Auswirkung von Fokus und der Stimmhaftigkeit des post-nasalen Plosivs auf die Positionen der Zungenspitze und des Gaumensegels in den Lautsequenzen /nd/ und /nt/. Hierfür wurden Zielwörter mit der Struktur CVNCV analysiert, wobei der erste Vokal ausschließlich ungespannt war. Als Referenzwerte wurden die maximale Werte im alveolaren Bereich beziehungsweise der maximale Wert der Gaumensenkung innerhalb dieser Lautsequenzen herangezogen. Die Ergebnisse zeigten ein konstant niedrigeres Gaumensegel in der stimmhaften /nd/ Sequenz in beiden Betonunsgarten mit Ausnahme von solchen Sequenzen, in denen /u/ mit allgemeinem Fokus dem Nasal voranging. Zwar war die Zungenspitze ebenfalls während /nd/ niedriger als bei /nt/, jedoch nur mit allgemeinem Fokus. In der kontrastiven Bedingung hingegen nahm die Zungenspitze eine vergleichbare Position zu /nt/ an. Hinsichtlich der eigentlichen Intonationsbedingung waren eindeutige Tendenzen erkennbar: Sowohl die Zungenspitze als auch das Gaumensegel wiesen mit kontrastivem Fokus eine signifikante Verstärkung ihrer Positionen auf, d.h. eine höhere Zungenspitzenposition und ein niedrigeres Gaumensegel. Einzige Ausnahme waren solche /nt/-Sequenzen, denen der Vokal /o/ voranging.

Experiment V beleuchtete die räumliche und zeitliche Ausdehnung der Senkung des Gaumensegels in Vokalen. Dafür wurden CVNCV, CVNC und CVNCC Sequenzen analysiert, in denen der erste Vokal ungespannt oder gespannt war. Die Ergebnisse zeigten allerdings keine nennenswerten Unterschiede bezüglich der Position des Gaumensegels, abgesehen von den ungespannten Vokalen /i/ und /u/, die ein höheres Gaumensegel in der betonten Variante anzeigten. Ein signifikanter Unterschied in der Dauer der Vokalnasalierung wurde zudem für den tiefen Vokal /a/ angegeben, welcher in der betonten Variante länger nasaliert war. Insgesamt war die leichte Tendenz zu erkennen, dass die vokalintrinsische Position des Gaumensegels mit kontrastivem Fokus verstärkt wurde. Allerdings fielen diese Tendenzen gering aus und waren nicht signifikant.

Zusammengefasst zeigen die Ergebnisse, dass die kontrastive Bedingung keinen systematischen Effekt auf die Position des Gaumensegels in den Vokalen hatte und nur bedingt auf die Dauer von Vokalnasalierung. Dies spricht tendenziell gegen die Annahme, dass Vokale allein aufgrund ihrer Betonung prädestiniert für Vokalnasalierung sind (Schourup, 1973). Im Gegensatz dazu waren jedoch deutliche Unterschiede in den konsonantischen Sequenzen erkennbar: Das Gaumensegel war weiter gesenkt und die Zungenspitze weiter gehoben, wenn das Zielwort mit kontrastivem Fokus artikuliert wurde. Da diese Konfiguration einem ausgeprägten nasalen Konsonanten entspricht, wäre ein Verlust des Nasals bei starker Betonung eher unwahrscheinlich. Zudem unterstreichen die Daten erneut die Rolle des post-nasalen Kontextes: Sowohl die Zungenspitze als auch das Gaumensegel wiesen abhängig von der post-nasalen Stimmhaftigkeit Unterschiede in ihrer Position auf. Dies wäre mit der Annahme vereinbar, dass ein stimmhafter Obstruent in akustischer und auch zu einem gewissen Grad in artikulatorischer Hinsicht kompatibel mit einem nasalen Laut ist, während ein stimmloser Obstruent eine schnelle und hinreichende Schließung des velopharyngalen Durchgangs erfordert (Ohala und Ohala 1991; 1993).

Experiment VI widmete sich der Frage, welche Rolle die Sprechgeschwindigkeit für die Artikulation des nasalen Konsonanten spielt und welche Unterschiede zum oralen Konsonanten erkennbar sind. Als Zielwörter wurden CVNV und CVCV Strukturen untersucht, in denen der erste Vokal entweder gespannt oder ungespannt war. Die Zielkonsonanten waren /n/ und /t/. Die Zielwörter wurden im Rahmen ihres Kontextsatzes in zwei unterschiedlichen Sprechgeschwindigkeiten vorgelesen. Durch einen Vergleich dieser beiden Gruppen sollte einerseits der Effekt des Sprechtempos auf die Zungenspitze und das Gaumensegel in jeweils /n/ und /t/ untersucht werden. Andererseits galt es auch zu ergründen, ob sich die Differenzen zwischen den Positionen, die in hohem gegenüber moderatem Sprechtempo erkennbar waren, für die einzelnen Konsonanten unterschieden und umgekehrt, ob die Differenz der Zungen- und Gaumensegelposition zwischen /n/ und /t/ in einem der Sprechstile größer war als in dem anderen. Darüber hinaus wurde für die CVNV Sequenzen die maximale Velozität des Gaumensegels während der Senkungsgeste evaluiert, um festzustellen, ob das Sprechtempo die Geschwindigkeit der Bewegung beeinflusste.

Die Ergebnisse zeigten in beiden Sprechgeschwindigkeiten eine signifikant schwächer ausgeprägte Zungengeste in /n/ als in /t/. Entgegen der Vorhersage, dass eine größere Distanz zwischen /n/ und /t/ mit höherem Sprechtempo zu erkennen sein sollte, waren die Differenzen in beiden Sprecharten ähnlich. Hinsichtlich des Gaumensegels erschien ein direkter Vergleich zwischen /n/ und /t/ als wenig sinnvoll, da dieses von Natur aus unterschiedliche Positionen in nasalen und oralen Konsonanten aufweist.

Das Sprechtempo selbst hatte einen signifikanten Effekt auf die Zungenspitzengeste: in beiden Konsonanten war diese bei höherem Sprechtempo weniger stark ausgeprägt. Die Ausprägung der Reduktion war jedoch für beide Konsonanten vergleichbar. Das Gaumensegel zeigte bei erhöhtem Sprechtempo ebenfalls eine signifikant geringere Amplitude als bei moderater Sprechgeschwindigkeit, jedoch nur in den /n/-Kontexten. Somit war der

Positionsabstand auch signifikant größer bei /n/ in der schnellen gegenüber der moderaten Bedingung im Vergleich zu /t/. Die Positionsabstände zwischen /n/ und /t/ hingegen waren in beiden Sprechgeschwindigkeiten ähnlich. Darüber hinaus deutete eine Analyse der Velozität der Gaumensegelbewegung darauf hin, dass die Geschwindigkeit während der Ausführung der Geste nicht systematisch verändert war.

Zusammengefasst stimmen die Ergebnisse teilweise mit denen aus vorheriger Forschung überein: Die Zungenspitze zeigte eine reduzierte Geste in /n/ gegenüber /t/ (Jaeger und Hoole, 2011); auch die Sprechgeschwindigkeit hatte einen signifikanten Einfluss auf die Position der Zungenspitze unabhängig vom Konsonanten (Adams et al., 1993; Flege, 1988; Gay et al., 1974; Kuehn, 1976; Lindblom, 1963). Allerdings blieb ein Unterschied in der Amplitude zwischen den Konsonanten aus. Da /t/ einen ausgeprägten alveolaren Verschluss in beiden Sprechgeschwindigkeiten erfordert, was bei /n/ nicht in demselben Maße der Fall ist, wäre ein deutlicher Unterschied zwischen schnellem zu moderatem /n/ verglichen mit schnellem zu moderatem /t/ zu erwarten gewesen. Dies war jedoch nicht der Fall. Des Weiteren wurde zuvor angenommen, dass hinsichtlich des Gaumensegels eine geringere Distanz zwischen /n/ und /t/ in der schnellen Bedingung erkennbar sein sollte: Wenn eine Geste mit erhöhtem Sprechtempo in ihrer Amplitude reduziert wird, sollte das Gaumensegel verglichen mit der moderaten Geschwindigkeit eine etwas niedrigere Position in /t/ und eine etwas höhere Position in /n/ aufweisen. Die Ergebnisse deuten jedoch nicht auf ein solches Verhalten hin. Darüber hinaus gab es keinerlei Hinweise auf eine erhöhte Velozität der Gaumenbewegung während der Öffnungsphase in /n/-Kontexten. Zusammen mit der Beobachtung, dass das Gaumensegel in diesen Sequenzen eine geringe Amplitude bei erhöhter Sprechgeschwindigkeit aufwies, spricht dies dafür, dass Sprecher eher die Amplitude der Gaumenbewegung verringern anstatt deren Velozität erhöhen, um das Artikulationsziel bei erhöhter Sprechgeschwindigkeit zu erreichen. Diskutiert wurde auch, ob eine reduzierte Zungenspitzengeste bei einem gleichzeitig gesenkten Gaumensegel akustische Änderungen bewirken könnte, auch wenn das Gaumensegel nicht so stark gesenkt ist verglichen mit moderatem Sprechtempo, sodass tendenziell weniger nasaler Luftstrom den Nasen-Rachen-Raum passiert. Möglicherweise aber ist diese Menge dennoch ausreichend, damit ein von der Zungenspitze lediglich angedeuteter Nasal artikulatorische und akustische Eigenschaften eines vokalähnlichen Lautes aufweist. Wie genau solche Konstellationen von Hörerinnen und Hörern interpretiert werden, bleibt allerdings Gegenstand zukünftiger Forschung.

Vokalnasalierung in der Perzeption

Kapitel V untersuchte den perzeptiven Aspekt von Vokalnasalierung anhand von Experiment VII (Abschnitt 5.1.2), an dem ausschließlich ehemalige Probandinnen und Probanden aus der MRT-Sprechstudie teilnahmen. Ziel war es – in Anlehnung an die Perzeptionsexperimente mit amerikanischen Probanden bei Beddor (2009)– herauszufinden, ob es Hörern leichter fällt, Stimuli voneinander zu unterschieden, die eine unterschiedliche Dauer an Vokalnasalierung aufweisen, als solche, die mit einem konstanten Anteil an Nasalierung ausgestattet sind, der sich jedoch unterschiedlich auf den Nasal und den Vokal verteilt. Darüber hinaus wurde der Frage nachgegangen, ob Sprecher, die in ihrer Produktion von ausgeprägter Nasalierung Gebrauch machen, ein bestimmtes Muster als Hörer in der Perzeption zeigen. Als Basiswort diente die synthetisch erzeugte Sequenz /baːntə/ *bahnte*. Die Fragestellungen wurden in Form eines sogenannten Staircase-Experiments untersucht, in welchem die Teilnehmerinnen und Teilnehmer jeweils vier Stimuli hörten, von denen der zweite oder dritte eine leichte Änderung aufwies. Die Probandinnen und Probanden sollten per Tastendruck entscheiden, welcher der beiden mittleren Stimuli sich von den anderen drei Stimuli unterschied. Dabei waren die Hörerinnen und Hörer mit zwei Versionen des Tests konfrontiert: In der ersten wurde eine konstant gehaltene Senkungsgeste des Gaumensegels variabel in den Vokal oder Nasal verschoben; in der zweiten wurde ausschließlich die Nasalierung im Vokal geändert. Das besondere an dem Staircase-Aufbau ist, dass sich die teilnehmenden Personen an ihre individuelle perzeptive Grenze herantasten können, da mit jeder gegebenen Antwort der Unterschied zwischen den Stimuli in vorgegebenen Schritten angepasst wird, abhängig davon, ob die vorangegangene Antwort korrekt oder falsch war. In unserem Experiment begannen alle teilnehmenden Personen mit dem maximalen Unterschied zwischen den Stimuli, der anschließend mit jeder korrekten Antwort schrittweise reduziert oder bei einer falschen Antwort erneut vergrößert wurde. Auf diese Weise erreichten die Hörerinnen und Hörer ihre individuelle Perzeptionsschwelle, um die sie sich im Verlauf des Experiments herumbewegten.

Im Gegensatz zu der Vorhersage, dass Hörer eine Änderung der Gesamtdauer an Nasalierung besser wahrnehmen als die unterschiedliche Verteilung einer konstanten Geste, zeigten die deutschsprachigen Probandinnen und Probanden eine insgesamt deutlich niedrigere Perzeptionsschwelle, wenn die konstante Geste verschoben wurde. Somit war eine wesentlich geringere Differenz zwischen diesen Stimuli nötig, um Unterschiede wahrzunehmen. Im Gegensatz dazu wiesen viele der teilnehmenden Personen größere Schwierigkeiten beim Erkennen von Nasalierungsunterschieden im Vokal auf: Hier war eine große Differenz zwischen den Stimuli erforderlich, damit die Stimuli als unterschiedlich wahrgenommen wurden. Der Erklärungsansatz wurde diskutiert, dass sich die Probandinnen und Probanden in der konstanten Bedingung möglicherweise weniger an der Nasalierungsgeste orientierten, sondern viel mehr an den Änderungen der Pause, die sich zwischen dem Vokalende und dem oralen Plosiv ergab, sobald die Geste verschoben wurde. Es wurde argumentiert, dass Muttersprachler des Deutschen viel eher mit Alternationen in der Pause vor einem oralen Plosiv vertraut sind, da sich beispielsweise /t/ und /d/ vor allem durch die Länge der Verschlussphase unterscheiden. Somit könnten die Änderungen der konstanten Geste als Variation des Plosivs interpretiert worden sein. Alternativ ließe sich das Ergebnis auch so deuten, dass sich die Hörerinnen und Hörer an der Schwächung des Nasals orientierten, da sie auch mit diesem Muster vertraut sind, insbesondere, wenn der Nasal von einem stimmlosen Plosiv gefolgt wird. Insgesamt betrachtet jedoch scheint Vokalnasalierung selbst kaum eine Rolle in der Perzeption deutscher Muttersprachler zu spielen, wohl aber solche akustische Feinheiten, mit denen sie vertraut sind.

Unabhängig von diesem Ergebnis stellt sich das Experiment als gute Möglichkeit dar, zu erforschen, wie viel zeitliche Nasalierung einzelne Sprecher einer Gruppe generell benötigen, um einen Vokal als nasaliert wahrzunehmen und auch, wie sich Muttersprachler verschiedener Sprachen darin unterscheiden.

Zusätzlich zu den Unterschieden zwischen den beiden Testtypen wurde außerdem überprüft, ob ein Zusammenhang zwischen den individuellen Nasalierungsmustern in der Produktion und Perzeption einzelner Sprecherinnen und Sprecher bestand, wie es in vorangegangenen Studien vermutet wurde (Beddor, 2015; Zellou, 2017). Hierfür wurde für jede teilnehmende Person auf Grundlage der MRT-Daten das Verhältnis von Vokalnasalierung zur Gesamtdauer der Senkungsgeste des Gaumensegels errechnet und als Funktion des Perzeptionswertes definiert. Die Daten deuteten jedoch nicht auf einen Zusammenhang zwischen dem Gebrauch von Nasalierung in der Produktion und besonderer perzeptiver Sensibilität hin. Obwohl sich durchaus Unterschiede zwischen dem Ausmaß an Nasalierung unter den Sprechern erkennen ließen, waren diese Muster womöglich zu gering, um die eigene perzeptive Sensibilität systematisch zu beeinflussen. Darüber hinaus wurde vermutet, dass Sprachnutzer flexibler in der Perzeption als in der Produktion sein könnten, da sie sich perzeptiv ständig an ihre wechselnde linguistische Umgebung anpassen, etwa aufgrund von unterschiedlichen Sprechern, Akzenten oder Sprachstilen (Beddor, 2015). Dies könnte eine Rolle dabei spielen, warum die Probandinnen und Probanden in dem hier vorgestellten Experiment zwar perzeptiv sensibel für feine Änderungen in dem konsonantischen Teil waren, jedoch kein Zusammenhang zwischen der Perzeption und Produktion festgestellt werden konnte.

Schlussfolgerung

Bei einem Lautwandel von koartikulatorisch nasalierten Vokalen hin zu Nasalvokalen sind mehrere Faktoren beteiligt, welche das Auslösen eines solchen Prozesses begünstigen. Voraussetzung dafür ist, dass der Vokal durch die gegenseitige Beeinflussung der artikulatorischen Gesten nasaliert wird oder aufgrund bestimmter akustischer Eigenschaften zumindest als nasaliert wahrgenommen wird.

Die Ergebnisse dieser Arbeit tragen zu Erkenntnissen über das Verhalten des Gaumensegels in einer Sprache bei, die weder kontrastive Nasalvokale noch stark nasalierte Vokale aufweist. Dem zugrunde liegt der Gedanke, dass viele Sprachen mit Nasalvokalen in früheren Sprachstufen diesen Status einmal aufwiesen und es bestimmte Änderungen in der Artikulation und Perzeption gegeben haben muss, damit der Wandel hin zu Nasalvokalen erfolgen konnte, was zudem oft mit dem Verlust des nasalen Konsonanten einherging. Die elementaren Bedingungen dafür sollten somit auch in synchronen Sprachen ohne Nasalvokale zu finden sein. Tatsächlich konnten unsere Daten bestätigen, dass der tiefe Vokal /a/ eine besondere Rolle einnimmt, der als gespannter Vokal in nasaler Umgebung mit einem deutlich niedrigeren Gaumensegel artikuliert wurde als andere Vokale und zudem auch länger nasaliert war. Auffällig war dabei, dass dieses Schema nicht in den oralen Kontexten erkennbar war. Als Erklärungsansatz wurde ein antagonistisches Wechselspiel der beteiligten Muskeln während der Gaumensegelsenkung angeführt, in der je nach konsonantischem Kontext der eine oder andere muskuläre Gegenspieler überwiegt. Gleichzeitig können auch perzeptive Fakoren beteiligt sein, sodass Sprecher den tiefen Vokal womöglich bewusst mit einem gewissen Grad an antizipatorischer Nasalität artikulieren, da dieser bei Hinzufügen von akustischen nasalen Anteilen toleranter gegenüber Änderungen des Frequenzspektrums ist. Sollten jedoch Hörer diese Nasalität, womöglich aufgrund einer längeren Vokaldauer, dennoch perzipieren, könnte dies eine Erklärung dafür sein, warum kontrastive Nasalierung oft zuerst den Vokal /a/ betrifft. Diese beiden Erklärungsansätze müssen sich nicht ausschließen: Sollten primär perzeptive Gründe für die Senkung des Gaumensegels verantwortlich sein, könnte die muskuläre Verbindung zwischen Zunge und Gaumensegel dabei unterstützend wirken. Wenn die muskuläre Verbindung hingegen die Hauptursache für die tiefe Position des Velums darstellt, müsste der Sprecher dem nicht entegenwirken, da der tiefe Vokal perzeptiv toleranter gegenüber Nasalität ist.

Eine weitere Erkenntnis bezieht sich auf das Resultat, dass der post-nasale Lautkontext eine tragende Rolle für den Grad der zeitlichen Nasalierung des Vokals und auch der Gesamtdauer der Senkung des Gaumensegels spielt. Für andere Sprachen wurde gezeigt, dass der Vokal stärker nasaliert ist und der folgende Nasal eine kürzere Dauer aufweist, wenn dieser von einem stimmlosen Obstruenten gefolgt wird. Unsere Daten konnten dies sowohl in zeitlicher als auch positioneller Hinsicht bestätigen: Ein dem Nasal vorangehender Vokal war länger nasaliert, wenn der Nasal von einem stimmlosen Plosiv gefolgt wurde, verglichen mit einem unbetonten post-nasalen Vokal. Gleichzeitig war die Gesamtdauer der Geste in den Plosiv-Kontexten reduziert. Darüber hinaus zeigten sich in rein konsonantischen Sequenzen Unterschiede in der Position des Gaumensegels in Abhängigkeit von der Stimmhaftigkeit des post-nasalen Konsonanten: Wurde der Nasal von einem stimmhaften Plosiv gefolgt, wies das Gaumensegel in diesen Sequenzen eine niedrigere Position auf. Dies ist mit dem aerodynamischen und akustischen Erklärungsansatz vereinbar, dass stimmhafte Plosive zu einem gewissen Grad mit den akustischen und artikulatorischen Eigenschaften eines Nasals kompatibel sind. Für die Bildung eines stimmlosen Plosivs hingegen ist ein möglichst dichter Verschluss des velopharyngalen Durchgangs erforderlich, um einen ausreichend großen Luftdruck für die Artikulation des Plosivs zu gewährleisten. Unsere Daten unterstützen somit die Annahme, dass Nasale in stimmlosen Kontexten prädestiniert sind für deren Schwächung, was langfristig gesehen in einem Verlust des Nasals resultieren kann.

Ob eine erhöhte Sprechgeschwindigkeit einen wesentlichen Faktor für die Entwicklung von Vokalnasalierung darstellt, ist fraglich. Dennoch konnten unsere Daten zu weiteren Erkenntnissen bezüglich der artikulatorischen Strategien in schnell gesprochenen Sätzen beitragen.

In vorangegangener Forschung wurde berichtet, dass manche Sprecher die Geschwindigkeit der Gaumensegelgeste erhöhen, während andere deren Amplitude reduzieren. Unsere Daten hingegen weisen auf eine recht einheitliche Strategie hin, die in der Abschwächung der Senkungsgeste besteht.

Ein weiterer Punkt betrifft die Perzeption von Nasalität: Offenbar sind Muttersprachlerinnen und Muttersprachler des Standarddeutschen recht unsensibel bezüglich moderater Schwankungen der Nasalierungdauer in Vokalen. Wesentlich präziser jedoch nehmen sie Modifikationen wahr, wenn diese den post-nasalen konsonantischen Kontext betreffen. Insofern können Hörerinnen und Hörer durchaus als aufmerksame Perzipienten gesehen werden, die feine akustische Änderungen wahrnehmen, die ihnen vertraut sind.

Zusammengefasst tragen die Ergebnisse dieser Arbeit zu einem tieferen Verständnis des charakteristischen Verhaltens des Gaumensegels in gesprochener Sprache sowie dessen Auswirkung auf die Nasalierung von Vokalen bei. Darüber hinaus bieten die Experimente eine Grundlage für zukünftige Forschung, die sich mit dem direkten Vergleich mit anderen Sprachen befasst hinsichtlich spezifischer Artikulationsmuster des Gaumensegels sowie der sprachspezifischen perzeptiven Sensibilität für die Nasalierung von Vokalen.

Appendices

Appendix A

Speech materials

A.1 Overall speech corpus

Table A.1: Overall speech corpus

Spelling	Gloss	IPA transcription
ahnde	‘avenge’ (1.sg.)	[aːndə]
Ahne	‘ancestor’	[aːnə]
ahnte	‘guessed’	[aːntə]
Ende	‘end’	[ɛndə]
Ente	‘duck’	[ɛntə]
bahne	‘channel’ (1.sg.)	[baːnə]
Bande	‘gang’	[bande]
bangst	‘tremble’ (2.sg.)	[baŋst]
Banner	‘banner’	[banɐ]
bannst	‘banish’ (2.sg.)	[banst]
bannte	‘banished’	[bantə]
Bast	‘bast’	[bast]
bat	‘asked for’	[baːt]
Biene	‘bee’	[biːnə]
biete	‘offer’ (1.sg.)	[biːtə]
Bitte	‘request’	[bɪtə]
Bohne	‘bean’	[boːnə]
Bote	‘carrier’	[boːtə]
Brunst	‘rutting season’	[bʁʊnst]
Brust	‘breast’	[bʁʊst]
buhte	‘booed’	[buːtə]
Bunde	‘league’	[bʊndə]

Table A.1: Overall speech corpus

Spelling	Gloss	IPA transcription
bunte	‘colorful’	[bʊntə]
Butte	‘butts’ (fish)	[bʊtə]
Diener	‘servant’	[diːnɐ]
diente	‘served’	[diːntə]
Dieter	‘Dieter’ (name)	[diːtɐ]
Düne	‘dune’	[dyːnə]
dünne	‘thin’	[dʏnə]
finde	‘find’ (1.sg.)	[fɪndə]
Finne	‘Fin’	[fɪnə]
Finte	‘feint’	[fɪntə]
fitter	‘fitter’	[fɪtɐ]
Gönner	‘patron’	[gœnɐ]
gönnte	‘indulged’	[gœntə]
Götter	‘gods’	[gœtɐ]
gähnst	‘yawn’ (2.sg.)	[gɛːnst]
gehst	‘go’ (2.sg.)	[geːst]
Kieme	‘gill’ (fish)	[kiːmə]
Kimme	‘notch’	[kɪmə]
kommst	‘come’ (2.sg.)	[kɔmst]
Kost	‘food’	[kɔst]
Künste	‘arts’	[kʏnstə]
Küste	‘coast’	[kʏstə]
lahm	‘lame’	[laːm]
lahmst	‘founder’ (2.sg.)	[laːmst]
lahmt	‘founders’ (3.sg.)	[laːmt]
Lamm	‘lamb’	[lam]
lehnst	‘lean’ (2.sg.)	[leːnst]
lehnt	‘leans’ (3.sg.)	[leːnt]
Linde	‘linden’	[lɪndə]
Linse	‘lense’	[lɪnzə]
lohne	‘(I am) worth it’	[loːnə]
lohnst	‘(you are) worth it’	[loːnst]
lohnt	‘(is) worth it’	[loːnt]
lohnte	‘(was) worth it’	[loːntə]
lost	‘draws lots’ (3.sg.)	[loːst]
Lote	‘perpendicular’	[loːtə]

Table A.1: Overall speech corpus

Spelling	Gloss	IPA transcription
Panda	‘panda’	[panda]
Panne	‘breakdown’	[panə]
Panther	‘panther’	[pantɐ]
Pate	‘godfather’	[paːtə]
Patte	‘patch’	[patə]
pennst	‘sleep’ (2.sg.)	[pɛnst]
pennt	‘sleeps’ (3.sg.)	[pɛnt]
Pest	‘pestilence’	[pɛst]
Peter	‘Peter’	[peːtɐ]
piep	‘peep’	[piːp]
Pute	‘turkey hen’	[puːtə]
Rahm	‘cream’	[ʁaːm]
rahmst	‘frame’ (2.sg.)	[ʁaːmst]
rahmt	‘frames’ (3.sg.)	[ʁaːmt]
rahmte	‘framed’	[ʁaːmtə]
rannte	‘ran’	[ʁantə]
rast	‘rushes’ (3.sg.)	[ʁaːst]
Rate	‘rate’	[ʁaːtə]
Ratte	‘rat’	[ʁatə]
Ränder	‘rims’	[ʁɛndɐ]
Renner	‘racer’	[ʁɛnɐ]
Rente	‘pension’	[ʁɛntə]
Retter	‘saviour’	[ʁɛtɐ]
Riem	‘lace’	[ʁiːm]
rinnst	‘trickle’ (2.sg.)	[ʁɪnst]
rinnt	‘trickles’ (3.sg.)	[ʁɪnt]
Rist	‘instep’	[ʁɪst]
Rita	‘Rita’	[ʁiːta]
Ruhm	‘glory’	[ʁuːm]
Rum	‘rum’	[ʁʊm]
Rute	‘rod’	[ʁuːtə]
Saate	‘seed’	[zaːtə]
Saft	‘juice’	[zaft]
Sahne	‘cream’	[zaːnə]
sahnst	‘cream’ (2.sg.)	[zaːnst]
sahnt	‘creams’ (3.sg.)	[zaːnt]

Table A.1: Overall speech corpus

Spelling	Gloss	IPA transcription
sahnte	‘creamed’	[zaːntə]
sahst	‘saw’ (2.sg.)	[zaːst]
Sande	‘sand’	[zande]
sandte	‘sent’	[zantə]
sanft	‘soft’	[zanft]
satte	‘saturated’	[zatə]
Sänfte	‘palanquin’	[zɛnftə]
Säfte	‘juices’	[zɛftə]
schienst	‘splint’ (2.sg.)	[ʃiːnst]
schient	‘splints’ (3.sg.)	[ʃiːnt]
schießt	‘shoots’ (3.sg.)	[ʃiːst]
schone	‘rest’ (1.sg.)	[ʃoːnə]
schonte	‘rested’	[ʃoːntə]
Schote	‘pod’ (vanilla)	[ʃoːtə]
Schotte	‘Scot’	[ʃɔtə]
Sehne	‘tendon’	[zeːnə]
sehnte	‘longed for’	[zeːntə]
sende	‘send’ (1.sg.)	[zɛndə]
Senne	‘Senne’ (landscape)	[zɛnə]
Sense	‘scythe’	[zɛnzə]
Senta	‘Senta’	[zɛnta]
Sonde	‘probe head’	[zɔndə]
Sonne	‘sun’	[zɔnə]
sonnte	‘sunned’	[zɔntə]
staunst	‘(you are) astonished’	[ʃta͜ʊnst]
staunt	‘(she is) astonished’	[ʃta͜ʊnt]
staust	‘impound’ (2.sg.)	[ʃta͜ʊst]
stöhnst	‘groan’ (2.sg.)	[ʃtøːnst]
stöhnt	‘groans’ (3.sg.)	[ʃtøːnt]
stößt	‘pokes’ (3.sg.)	[ʃtøːst]
Sühne	‘expiation’	[zyːnə]
sühnst	‘expiate’ (2.sg.)	[zyːnst]
sühnt	‘expiates’ (3.sg.)	[zyːnt]
süßt	‘sweetens’ (3.sg.)	[zyːst]
Tat	‘act’	[taːt]
Tina	‘Tina’	[tiːna]

Table A.1: Overall speech corpus

Spelling	Gloss	IPA transcription
Toner	‘toner’	[toːnɐ]
Tonne	‘ton’	[tɔnə]
Toter	‘dead man’	[toːtɐ]
Töne	‘tones’	[tøːnə]
tönte	‘tinted’	[tøːntə]
töte	‘kill’ (1.sg.)	[tøːtə]
thronst	‘(you are) enthroned’	[tʁoːnst]
thront	‘(she is) enthroned’	[tʁoːnt]
Trost	‘comfort’	[tʁoːst]
tun	‘do’	[tuːn]
tut	‘does’	[tuːt]
wandte	‘turned’	[vantə]
Wanne	‘tub’	[vanə]
wate	‘wade’ (1.sg.)	[vaːtə]
Watte	‘wadding’	[vatə]
weine	‘cry’ (1.sg.)	[va͜ɪnə]
weinte	‘cried’	[va͜ɪntə]
Weite	‘width’	[va͜ɪtə]
winde	‘twist’ (1.sg.)	[vɪndə]
Windeln	‘diapers’	[vɪndeln]
winseln	‘whimper’ (inf.)	[vɪnzeln]
Winter	‘winter’	[vɪntɐ]

A.2 Chapter II: Vowel height and velum position

Table A.2: Speech materials: Vowel height and velum position

Spelling	Gloss	IPA transcription	Number of items
bahne	‘channel’ (1.sg.)	[baːnə]	33
Banner	‘banner’	[banɐ]	32
bat	‘asked for’	[baːt]	32
Biene	‘bee’	[biːnə]	34
biete	‘offer’	[biːtə]	33
Bitte	‘request’	[bɪtə]	34
Bohne	‘bean’	[boːnə]	33
Bote	‘carrier’	[boːtə]	33
buhte	‘booed’	[buːtə]	34
Butte	‘butts’ (fish)	[bʊtə]	34
Diener	‘servant’	[diːnɐ]	33
Dieter	‘Dieter’ (name)	[diːtɐ]	33
Düne	‘dune’	[dyːnə]	32
dünne	‘thin’	[dʏnə]	36
Finne	‘Finn’	[fɪnə]	33
fitter	‘fitter’	[fɪtɐ]	34
Gönner	‘sponsor’	[gœnɐ]	33
Götter	‘gods’	[gœtɐ]	32
Panne	‘breakdown’	[panə]	31
Patte	‘flap’ (clothes)	[patə]	32
schone	‘rest’ (1.sg.)	[ʃoːnə]	35
Schote	‘pod’ (vanilla)	[ʃoːtə]	34
Schotte	‘Scot’	[ʃɔtə]	33
Töne	‘tones’	[tøːnə]	35
töte	‘kill’ (1.sg.)	[tøːtə]	34
Toner	‘toner’	[toːnɐ]	33
Toter	‘dead man’	[toːtɐ]	35
Tonne	‘ton’	[tɔnə]	32
tun	‘do’	[tuːn]	33
tut	‘does’	[tuːt]	32
Wanne	‘tub’	[vanə]	33
Watte	‘cotton batting’	[vatə]	33
Sahne	‘cream’	[zaːnə]	33
Saate	‘seed’	[zaːtə]	33

Table A.2: Speech materials: Vowel height and velum position

Spelling	Gloss	IPA transcription	Number of items
satte	‘saturated’	[zatə]	32
Sonne	‘sun’	[zɔnə]	33
Sühne	‘expiation’	[zyːnə]	34
süßt	‘sweetens’ (3.sg.)	[zyːst]	36

A.3 Chapter III: Vowel length, vowel nasalization and nasal duration

Table A.3: Speech materials: vowel length, vowel nasalization and nasal duration

Spelling	Gloss	IPA transcription	Number of items
bahne	‘channel’ (1.sg.)	[baːnə]	33
Banner	‘banner’	[banɐ]	32
bannte	‘banished’	[bantə]	33
Diener	‘servant’	[diːnɐ]	33
diente	‘served’	[diːntə]	30
Düne	‘dune’	[dyːnə]	32
dünne	‘thin’	[dʏnə]	35
Finne	‘Fin’	[fɪnə]	33
Finte	‘trick’	[fɪntə]	34
Gönner	‘sponsor’	[gœnɐ]	33
gönnte	‘indulged’	[gœntə]	34
Kieme	‘gill’ (fish)	[kiːmə]	33
Kimme	‘notch’	[kɪmə]	31
lahm	‘slow’	[laːm]	31
Lamm	‘lamb’	[lam]	34
Panne	‘breakdown’	[panə]	31
Panther	‘panther’	[pantɐ]	31
Sahne	‘cream’	[zaːnə]	33
sahnte	‘creamed’	[zaːntə]	32
sandte	‘sent’	[zantə]	35
schone	‘rest’ (1.sg.)	[ʃoːnə]	34
schonte	‘rested’	[ʃoːntə]	34
Sehne	‘tendon’	[zeːnə]	34
sehnte	‘longed for’	[zeːntə]	32

Table A.3: Speech materials: vowel length, vowel nasalization and nasal duration

Spelling	Gloss	IPA transcription	Number of items
Senne	'Senne' (landscape)	[zɛnə]	34
Senta	'Senta'	[zɛnta]	33
Sonne	'sun'	[zɔnə]	33
sonnte	'sunned'	[zɔntə]	33
Sühne	'expiation'	[zyːnə]	34
sühnt	'expiates'	[zyːnt]	31
Töne	'tones'	[tøːnə]	35
tönte	'sounded'	[tøːntə]	33
Toner	'toner'	[toːnɐ]	33
Tonne	'ton'	[tɔnə]	32
Wanne	'tub'	[vanə]	33
wandte	'turned'	[vantə]	32

Table A.4: Subset: CVNV

Spelling	Gloss	IPA transcription	Number of items
bahne	'channel' (1.sg.)	[baːnə]	33
Banner	'banner'	[banɐ]	32
lahm	'slow'	[laːm]	31
Lamm	'lamb'	[lam]	34
Sahne	'cream'	[zaːnə]	33
Wanne	'tub'	[vanə]	33
Panne	'breakdown'	[panə]	31
Sehne	'tendon'	[zeːnə]	34
Senne	'Senne' (landscape)	[zɛnə]	34
Kieme	'gill' (fish)	[kiːmə]	33
Kimme	'notch'	[kɪmə]	31
Diener	'servant'	[diːnɐ]	33
Finne	'Fin'	[fɪnə]	33
schone	'rest' (1.sg.)	[ʃoːnə]	34
Sonne	'sun'	[zɔnə]	33
Toner	'toner'	[toːnɐ]	33
Tonne	'ton'	[tɔnə]	32
Sühne	'expiation'	[zyːnə]	33
Düne	'dune'	[dyːnə]	32

Table A.4: Subset: CVNV

Spelling	Gloss	IPA transcription	Number of items
dünne	'thin'	[dʏnə]	35
Töne	'tones'	[tøːnə]	35
Gönner	'sponsor'	[gœnɐ]	33

Table A.5: Subset: CVNCV

Spelling	Gloss	IPA transcription	Number of items
sahnte	'creamed'	[zaːntə]	32
sandte	'sent'	[zantə]	35
bannte	'banished'	[bantə]	33
Panther	'panther'	[pantɐ]	31
wandte	'turned'	[vantə]	32
diente	'served'	[diːntə]	30
Finte	'trick'	[fɪntə]	34
tönte	'sounded'	[tøːntə]	33
gönnte	'indulged'	[gœntə]	34
schonte	'rested'	[ʃoːntə]	34
sonnte	'sunned'	[zɔntə]	33
sehnte	'longed for'	[zeːntə]	32
Senta	'Senta'	[zɛnta]	33
sühnt	'expiates' (3.sg.)	[zyːnt]	31

Table A.6: Subset: tense vowels

Spelling	Gloss	IPA transcription	Number of items
bahne	'channel' (1.sg.)	[baːnə]	33
lahm	'slow'	[laːm]	31
Sahne	'cream'	[zaːnə]	33
sahnte	'creamed'	[zaːntə]	32
Sehne	'tendon'	[zeːnə]	34
sehnte	'longed for'	[zeːntə]	32
Diener	'servant'	[diːnɐ]	33
diente	'served'	[diːntə]	30
Kieme	'gill' (fish)	[kiːmə]	33

Table A.6: Subset: tense vowels

Spelling	Gloss	IPA transcription	Number of items
schone	‘rest’ (1.sg.)	[ʃoːnə]	34
schonte	‘rested’	[ʃoːntə]	34
Toner	‘toner’	[toːnɐ]	33
Düne	‘dune’	[dyːnə]	32
Sühne	‘expiation’	[zyːnə]	34
sühnt	‘expiates’ (3.sg.)	[zyːnt]	31
Töne	‘tones’	[tøːnə]	35
tönte	‘sounded’	[tøːntə]	33

Table A.7: Subset: lax vowels

Spelling	Gloss	IPA transcription	Number of items
Banner	‘banner’	[banɐ]	32
bannte	‘banished’	[bantə]	33
Lamm	‘lamb’	[lam]	34
Panne	‘breakdown’	[panə]	31
Panther	‘panther’	[pantɐ]	31
sandte	‘sent’	[zantə]	35
Wanne	‘tub’	[vanə]	33
wandte	‘turned’	[vantə]	32
Senne	‘Senne’ (landscape)	[zɛnə]	34
Senta	‘Senta’	[zɛnta]	33
Finne	‘Fin’	[fɪnə]	33
Finte	‘trick’	[fɪntə]	34
Kimme	‘notch’	[kɪmə]	31
Sonne	‘sun’	[zɔnə]	33
sonnte	‘sunned’	[zɔntə]	33
Tonne	‘ton’	[tɔnə]	32
Gönner	‘sponsor’	[gœnɐ]	33
gönnte	‘indulged’	[gœntə]	34

A.4 Chapter IV: Stress and speaking rate

Table A.8: Speech materials: Focus effects on the velum during the vowel

Spelling	Gloss	IPA transcription	Number of items contr. focus	Number of items broad focus
bannst	‘banish’ (2.sg.)	[banst]	35	34
bannte	‘banished’	[bantə]	34	33
bunte	‘colorful’	[bʊntə]	35	31
Finte	‘trick’	[fɪntə]	35	34
gähnst	‘yawn’ (2.sg.)	[gɛːnst]	33	35
Künste	‘arts’	[kʏnstə]	36	33
lehnst	‘lean’ (2.sg.)	[leːnst]	33	35
lehnt	‘leans’ (3.sg.)	[leːnt]	33	32
lohnst	‘worth it’ (2.sg.)	[loːnst]	34	32
lohnt	‘worth it’ (3.sg.)	[loːnt]	32	33
Panther	‘panther’	[pantɐ]	34	31
pennst	‘sleep’ (2.sg.)	[pɛnst]	34	32
pennt	‘sleeps’ (3.sg.)	[pɛnt]	32	35
schienst	‘splint’ (2.sg.)	[ʃiːnst]	35	33
schient	‘splints’ (3.sg.)	[ʃiːnt]	31	33
stöhnst	‘groan’ (2.sg.)	[ʃtøːnst]	34	31
stöhnt	‘groans’ (3.sg.)	[ʃtøːnt]	35	33
Winter	‘winter’	[vɪntɐ]	35	33
sahnst	‘cream’ (2.sg.)	[zaːnst]	34	33
sahnt	‘creams’ (3.sg.)	[zaːnt]	38	34
Senta	‘Senta (name)’	[zɛntɐ]	39	33
sonnte	‘sunned’	[zɔntə]	35	33
sühnst	‘expiate’ (2.sg.)	[zyːnst]	34	33
sühnt	‘expiates’ (3.sg.)	[zyːnt]	37	31

Table A.9: Speech materials: Focus effects on the tongue and velum during /nd/ and /nt/

Spelling	Gloss	IPA transcription	Number of items contr. focus	Number of items broad focus
Ende	‘end’	[ɛndə]	37	35
Ente	‘duck’	[ɛntə]	34	33
Bande	‘gang’	[bandə]	35	34
bannte	‘banished’	[bantə]	34	33
Bunde	‘bunch’	[bʊndə]	35	35
bunte	‘colorful’	[buntə]	35	31
finde	‘find’ (1.sg.)	[fɪndə]	34	33
Finte	‘trick’	[fɪntə]	35	34
Panda	‘panda it’	[panda]	34	32
Panther	‘panther’	[pantɐ]	34	31
pennt	‘sleeps’ (3.sg.)	[pɛnt]	32	35
winde	‘twist’ (1.sg.)	[vɪndə]	35	33
Winter	‘winter’	[vɪntɐ]	35	33
sende	‘send’ (1.sg.)	[zɛndə]	34	33
Senta	‘Senta’	[zɛnta]	39	33
Sonde	‘probe’	[zɔndə]	35	37
sonnte	‘sunned’	[zɔntə]	35	33

Table A.10: Speech materials: Speaking rate effects on the tongue and velum during /n/ and /t/

Spelling	Gloss	IPA transcription	Number of items fast sp.rate	Number of items normal sp.rate
bahne	‘channel’ (1.sg.)	[baːnə]	35	33
Banner	‘banner’	[banɐ]	35	32
bat	‘asked for’	[baːt]	34	32
Biene	‘bee’	[biːnə]	32	34
biete	‘offer’	[biːtə]	34	33
Bitte	‘request’	[bɪtə]	35	34
buhte	‘booed’	[buːtə]	33	34
Butte	‘butts’ (fish)	[bʊtə]	32	34
Diener	‘servant’	[diːnɐ]	32	33
Dieter	‘Dieter’ (name)	[diːtɐ]	34	33
Düne	‘dune’	[dyːnə]	34	32
dünne	‘thin’	[dʏnə]	34	36
Finne	‘Finn’	[fɪnə]	34	33

Table A.10: Speech materials: Speaking rate effects on the tongue and velum during /n/ and /t/

Spelling	Gloss	IPA transcription	Number of items fast sp.rate	Number of items normal sp.rate
fitter	‘fitter’	[fɪtɐ]	37	34
Gönner	‘sponsor’	[gœnɐ]	29	33
Götter	‘gods’	[gœtɐ]	35	32
Panne	‘break down’	[panə]	31	31
Pate	‘godfather’	[paːtə]	31	34
Patte	‘patch’	[patə]	33	32
Sahne	‘cream’	[zaːnə]	32	33
Saate	‘seed’	[zaːtə]	35	33
satte	‘saturated’	[zatə]	34	32
schone	‘rest’ (1.sg.)	[ʃoːnə]	34	35
Schote	‘pod’ (vanilla)	[ʃoːtə]	36	34
Schotte	‘Scot’	[ʃɔtə]	34	33
Sehne	‘tendon’	[zeːnə]	35	34
Senne	*landscape*	[zɛnə]	35	34
Toner	‘toner’	[toːnɐ]	34	33
Tonne	‘ton’	[tɔnə]	35	32
toter	‘dead’	[toːtɐ]	34	35
Töne	‘tones’	[tøːnə]	34	35
töte	‘kill’ (1.sg.)	[tøːtə]	33	34
tun	‘do’	[tuːn]	33	33
tut	‘does’	[tuːt]	34	32
Wanne	‘tub’	[vanə]	32	33
wate	‘wade’ (1.sg.)	[vaːtə]	35	33
Watte	‘wadding’	[vatə]	33	33

A.5 Chapter V: Vowel nasalization in production and perception

Table A.11: Speech materials for calculating the vowel nasalization ratio

Spelling	Gloss	IPA transcription	Number of items
bahne	‘channel’ (1.sg.)	[baːnə]	33
Biene	‘bee’	[biːnə]	32
Bohne	‘bean for’	[boːnə]	32
Diener	‘servant’	[diːnɐ]	34
diente	‘served’	[diːntə]	33
Düne	‘dune’	[dyːnə]	34
Kieme	‘gill’ (fish)	[kiːmə]	33
lahm	‘lame’	[laːm]	33
lahmt	‘founders’ (3.sg.)	[laːmt]	34
lehnt	‘leans’ (3.sg.)	[leːnt]	34
lohne	‘(I am) worth it’	[loːnə]	33
lohnt	‘is worth it’ (3.sg.)	[loːnt]	33
schone	‘rest’ (1.sg.)	[ʃoːnə]	32
schonte	‘rested’	[ʃoːntə]	36
stöhnt	‘groans’ (3.sg.)	[ʃtøːnt]	33
Töne	‘tones’	[tøːnə]	34
tönte	‘sounded’	[tøːntə]	33
Toner	‘toner’	[toːnɐ]	32
tun	‘do’	[tuːn]	31
Sahne	‘cream’	[zaːnə]	32
sahnte	‘creamed’	[zaːntə]	35
sahnt	‘creams’	[zaːnt]	34
Sehne	‘tendon’	[zeːnə]	33
sehnte	‘longed for’	[zeːntə]	35
Sühne	‘expiation’	[zyːnə]	34
sühnt	‘expiates’ (3.sg.)	[zyːnt]	33

Appendix B

Further statistical analyses

B.1 Chapter IV: Focus and speaking rate effects

B.1.1 Focus effects

Differences between vowel duration in BF vs. CF and PF conditions

Analyses were performed by applying a linear mixed model in the programming environment *RStudio* using the *lmerTest* package. For all conditions tested, the following parameters were defined: dependent variable: vowel duration; fixed effects: focus condition and vowel; random effects: speaker and word onset.

BF vs. CF, tense vowels Two main effects were found for the condition (F[1,32]=103.96, $p<0.001$) and the vowel (F[5,2]=149.10, $p<0.001$). Vowels in the focal condition were rated as significantly longer than in the broad condition.

BF vs. CF, lax vowels Two main effects were found for the condition and the vowel (F[1,36]=32.41, $p<0.001$ and F[5,5]=11.06, $p<0.05$). A significant interaction F[5,578]=7.12, $p<0.001$) revealed clear duration differences for the vowels in the two focus conditions except for /u/ and /o/.

BF vs. PF, tense vowels No main effect was found for the condition, but for the vowel (F[5,2]=138.36, $p<0.01$). A marginal interaction (F[5,681]=2.32, $p<0.05$) between the condition and the vowel showed no significant differences for any vowel.

BF vs. PF, lax vowels A main effect was found for the vowel (F[5,4]=15.38, $p<0.01$) as well as an interaction between the vowel and the condition (F[5,561]=3.28, $p<0.01$). However, no significant differences were found for any vowel.

Differences between the temporal extent of vowel nasalization in vowels followed by /nt/ vs. /nst/

Analyses were performed by applying a linear mixed model in the programming environment *RStudio* using the *lmerTest* package. For both the tense and lax vowel items, the following parameters were defined: dependent variable: duration of vowel nasalization; fixed effects: focus condition, coda and vowel; random effects: speaker and word onset.

Temporal extent: tense vowels Two main effects were found for the vowel ($F[5,2]=52.19$, $p<0.05$) and the focus condition ($F[1,32.32]=13.78$, $p<0.001$), but not for the coda. Two interactions were reported for the vowel and condition ($F[5,696]=8.94$, $p<0.001$) and for the coda and vowel ($F[5,694]=2.39$, $p<0.05$). The latter, which is of interest here, referred to differences between the vowels dependent on the condition (e.g. /e–o/ was different in /nst/ but not in /nt/), whereas no significant differences were found between the /nt/ and /nst/ contexts for any vowel.

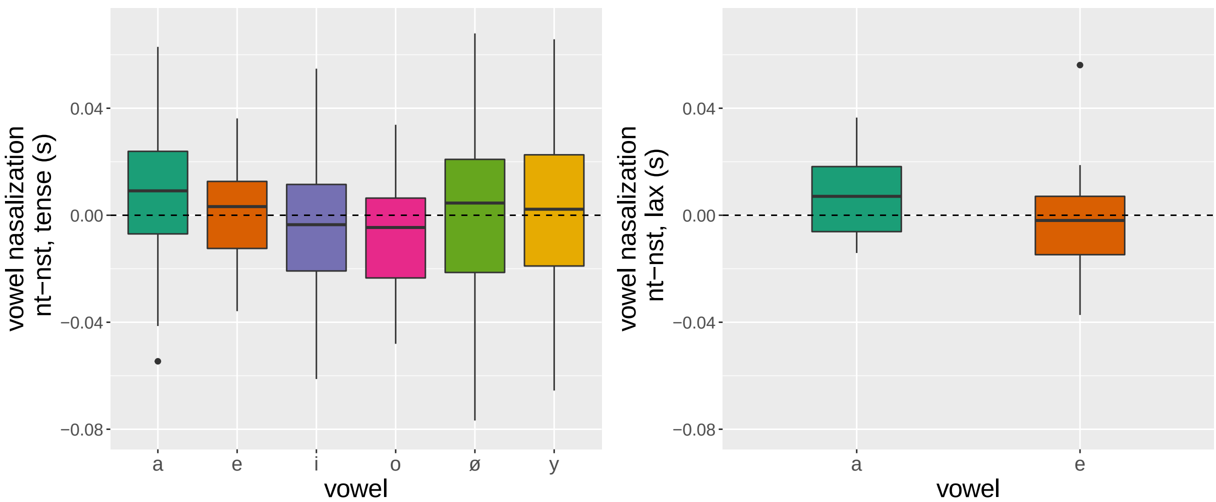


Figure B.1: Temporal differences of vowel nasalization in tense (left) and lax (right) vowels when followed by different consonantal contexts (/nt/ vs. /nst/). Mean values are given for each vowel per speaker.

Temporal extent: lax vowels Two main effects were found for the vowel ($F[5,6]=8.67$, $p<0.05$) and the condition ($F[1,39]=5.14$, $p<0.05$), but not for the coda. In addition, an interaction was found between the vowel and the condition ($F[5,574]=3.37$, $p<0.01$). As no main or interaction effect was found for the coda, results suggest that the coda context had no impact on the temporal extent of vowel nasalization.

Differences between the spatial amount of velum lowering in vowels followed by /nt/ vs. /nst/

Analyses were performed by applying a linear mixed model in the programming environment *RStudio* using the *lmerTest* package. For both the tense and lax vowel items, the following parameters were defined: dependent variable: velum lowering degree at the vowel midpoint; fixed effects: condition, coda and vowel; random effects: speaker and word onset.

Spatial amount: tense vowels A main effect was found for the vowel (F[5,48]=34.40, p<0.001) as well as an interaction between the condition and the vowel (F[5,704]=3.12, p<0.01), but not for the coda. Results thus do not suggest a significant impact of the coda on the degree of velum lowering in words with /nt/ vs. /nst/ sequences.

Spatial amount: lax vowels Two main effects were found, one for the vowel (F[5,6]=7.07, p<0.05) and one for the condition (F[1,36]=9.12, p<0.01), but not for the coda. In addition, an interaction was found between the vowel and the condition (F[5,578]=3.17, p<0.01). As no main or interaction effect was found for the coda, results suggest that the coda context had no impact on the spatial amount of velum lowering at the vowel midpoint.

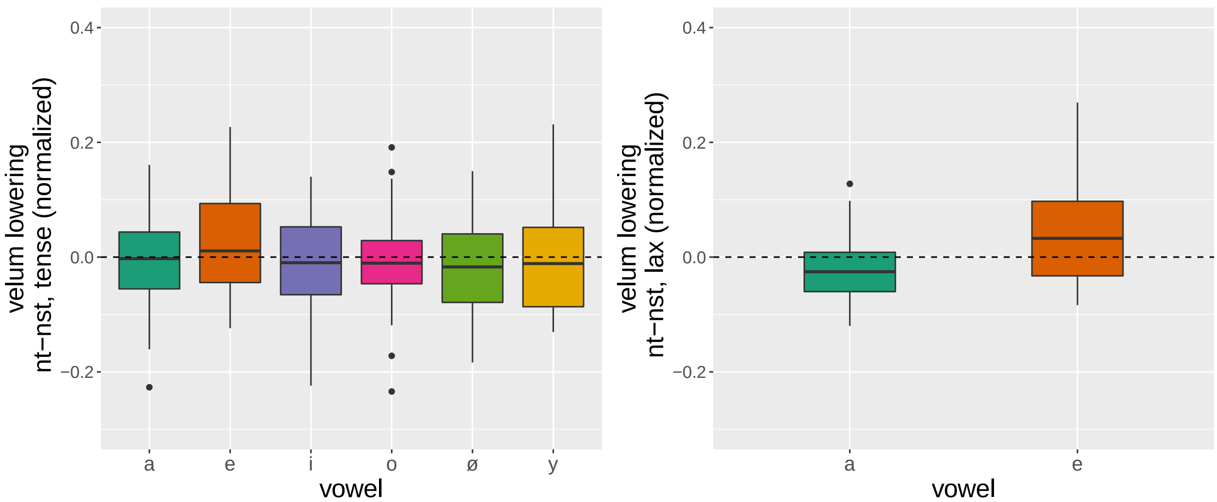


Figure B.2: Differences of the spatial amount of velum lowering in tense (left) and lax (right) vowels when followed by different consonantal contexts (/nt/ vs. /nst/). Mean values are given for each vowel per speaker.

B.1.2 Speaking rate

Differences of word and vowel duration in target words produced in fast vs. moderate speech

The plots in figs. B.3 and B.4 illustrate word and vowel duration differences when produced with fast speech compared to a moderate speaking rate. The analyses differentiate between CVNV and CVCV words with tense vs. lax vowels. The following parameters were used: dependent variable: word and vowel duration; fixed effects: speaking rate and vowel; random effects: speaker and word onset.

CVNV: tense vowels For word duration, two main effects were found for the rate condition ($F[1,33]=72.14$, $p<0.001$) and the vowel category ($F[6,7]=83.36$, $p<0.001$). An interaction of these effects ($F[6,590]=3.57$, $p<0.01$) referred to differences between specific vowel pairs dependent on the speaking rate. For all vowels involved, however, significant word duration differences were found between the fast and the moderate speaking rate condition. Similarly, with respect to vowel duration, two main effects were found for the rate condition and the vowel category ($F[1,34]=72.33$, $p<0.001$; $F[6,16]=100.76$, $p<0.001$) as well as an interaction ($F[6,442]=8.66$, $p<0.001$) referring to differences considering individual vowel pairs in the two rate conditions. However, for each vowel considered for the two conditions, vowel duration was significantly affected, with longer vowels in normal speech.

CVNV: lax vowels Word duration was found to be significantly affected by the condition ($F[1,36]=45.82$, $p<0.001$), but not the vowel. An interaction between the rate and the vowel ($F[5,454]=2.24$, $p<0.05$) showed significant differences for all vowels except for /ø/ ($p=0.8037$). Considering vowel duration, a main effect was found for the rate condition ($F[1,36]=23.88$, $p<0.001$) as well as an interaction between the condition and the vowel ($F[5,362]=4.27$, $p<0.001$). Significant duration differences were found for /a/, /e/, /o/, but not for /i/, /ø/, /y/.

CVCV: tense vowels For word duration, two main effects were found for the condition ($F[1,33]=76.85$, $p<0.001$) and the vowel category ($F[4,668]=186.45$, $p<0.001$). With respect to vowel duration, a main effect was found for both the condition ($F[1,35]=51.73$, $p<0.001$) and the vowel ($F[4,76]=80.33$, $p<0.001$). An interaction ($F[4,565]=3.10$, $p<0.05$) revealed that vowel duration was significantly affected for all vowels except for /i/ ($p=0.0560$).

CVCV: lax vowels Considering CVCV Words with lax vowels involved, word duration was found to be significantly affected by the condition ($F[1,32]=79.64$, $p<0.001$), with no interaction reported. Considering vowel duration, a main effect was found for the rate

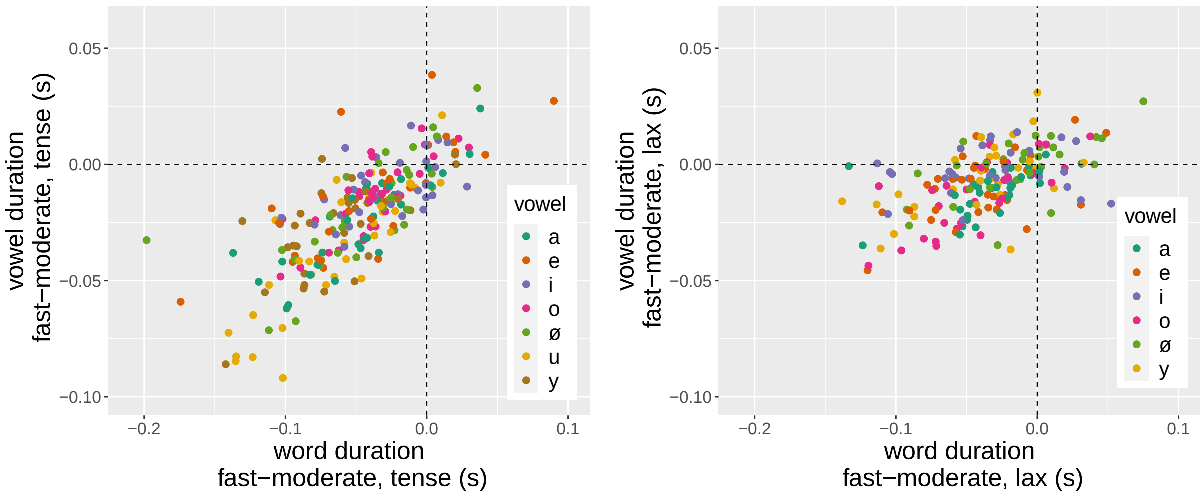


Figure B.3: Differences of the word (x-axis) and vowel duration (y-axis) between CVNV words in fast vs. moderate speech. Left: CVNV with pre-nasal tense vowels; right: CVNV with pre-nasal lax vowels. Mean difference values are given, separated by vowel and speaker.

condition ($F[1,33]=25.75$, $p<0.001$) as well as for the vowel category ($F[4,4]=21.76$, $p<0.01$). No interaction was reported.

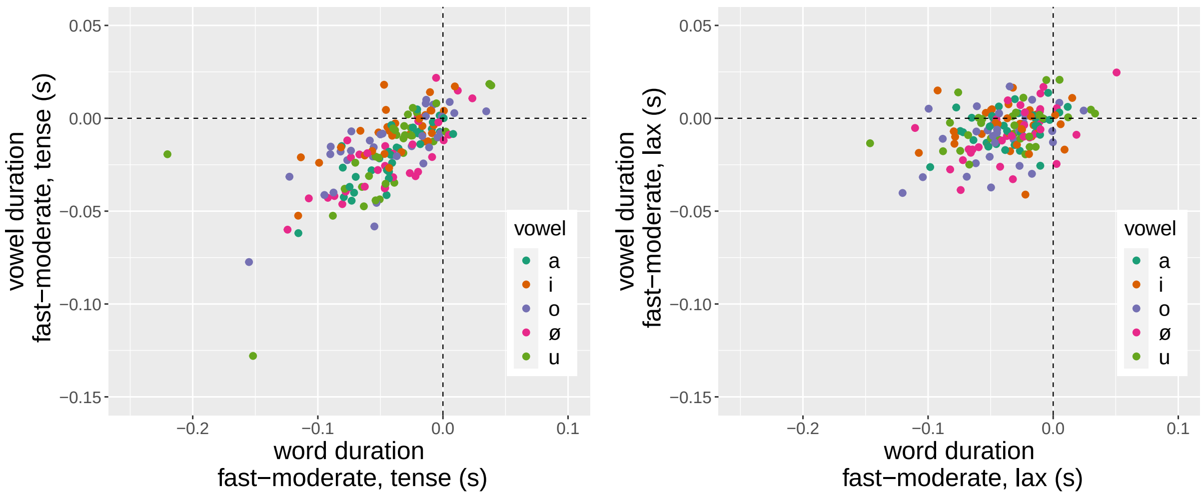


Figure B.4: Differences of the word (x-axis) and vowel duration (y-axis) between CVCV words in fast vs. moderate speech. Left: CVCV with pre-nasal tense vowels; right: CVCV with pre-nasal lax vowels. Mean difference values are given, separated by vowel and speaker.

Differences in velum position during /t/ and /n/ in fast and moderate speech

Figure B.5 (x-axis) depicts differences in the velum position between /t/ vs. /n/ for which the maximum velum signal was considered for both stops (in contrast to the maxmin variable considered in chapter IV). The maximum velum signal was determined as the dependent variable, the consonant and speaking rate were defined as the fixed effects and the speaker and word stem as random effects. With tense vowels involved, a main effect was found for both the consonant ($F[1,33]=759$, $p<0.001$) and the speaking rate ($F[1,32]=4.70$,

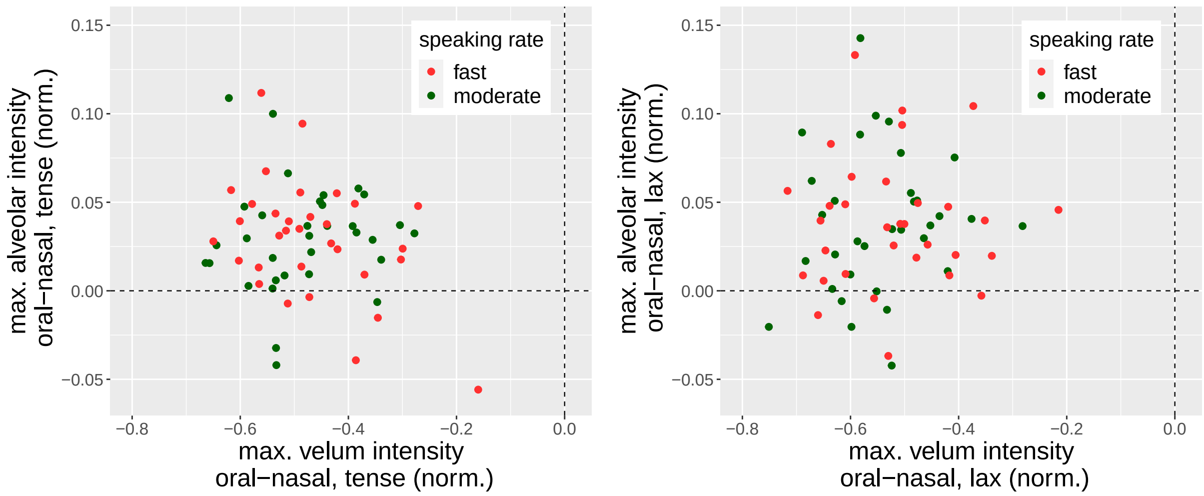


Figure B.5: Differences of the maximum signal intensity of /n/ and /t/ uttered with fast vs. moderate speech. Results are given for the velar region (x-axis) and the alveolar region (y-axis). Left: stops preceded by tense vowels; right: stops preceded by lax vowels. Normalized mean values are indicated for the oral–nasal difference, separated by the speaker.

$p<0.05$), indicating a lower velum in moderate speech and a generally lower position in /n/. For lax vowels, a main effect for the consonant was reported ($F[1,35]=771.08$, $p<0.001$) as well as an interaction between the consonant and the speaking rate ($F[1,955]=9.17$, $p<0.01$), suggesting a significant effect of the speaking rate in /n/ ($p<0.05$), but not in /t/.

B.2 Chapter V: Vowel nasalization in production and perception

Example for the Staircase procedure

Table B.1: Fictitious example of the staircase procedure: the nasalized portions are temporally adjusted depending on the responses.

trials	delta const/ext	condition: constant	condition: extend
$1_{correct}$	80ms/300ms	bã$_{0ms}$n$_{80ms}$te – bã$_{80ms}$n$_{0ms}$te	bã$_{300ms}$nte – bã$_{0ms}$nte
$1a_{correct}$	80ms/300ms	bã$_{80ms}$n$_{0ms}$te – bã$_{0ms}$n$_{80ms}$te	bã$_{0ms}$nte – bã$_{300ms}$nte
$2_{correct}$	40ms/150ms	bã$_{0ms}$n$_{80ms}$te – bã$_{40ms}$n$_{40ms}$te	bã$_{150ms}$nte – bã$_{0ms}$nte
$2a_{correct}$	40ms/150ms	bã$_{80ms}$n$_{0ms}$te – bã$_{40ms}$n$_{40ms}$te	bã$_{0ms}$nte – bã$_{150ms}$nte
3_{wrong}	20ms/75ms	bã$_{0ms}$n$_{80ms}$te – bã$_{20ms}$n$_{60ms}$te	bã$_{75ms}$nte – bã$_{0ms}$nte
$4_{correct}$	30ms/107.5ms	bã$_{0ms}$n$_{80ms}$te – bã$_{30ms}$n$_{50ms}$te	bã$_{107.5ms}$nte – bã$_{0ms}$nte
$4a_{correct}$	30ms/107.5ms	bã$_{80ms}$n$_{0ms}$te – bã$_{50ms}$n$_{30ms}$te	bã$_{0ms}$nte – bã$_{107.5ms}$nte
5	15ms/53.75ms	bã$_{0ms}$n$_{80ms}$te – bã$_{15ms}$n$_{65ms}$te	bã$_{53.75ms}$nte – bã$_{0ms}$nte

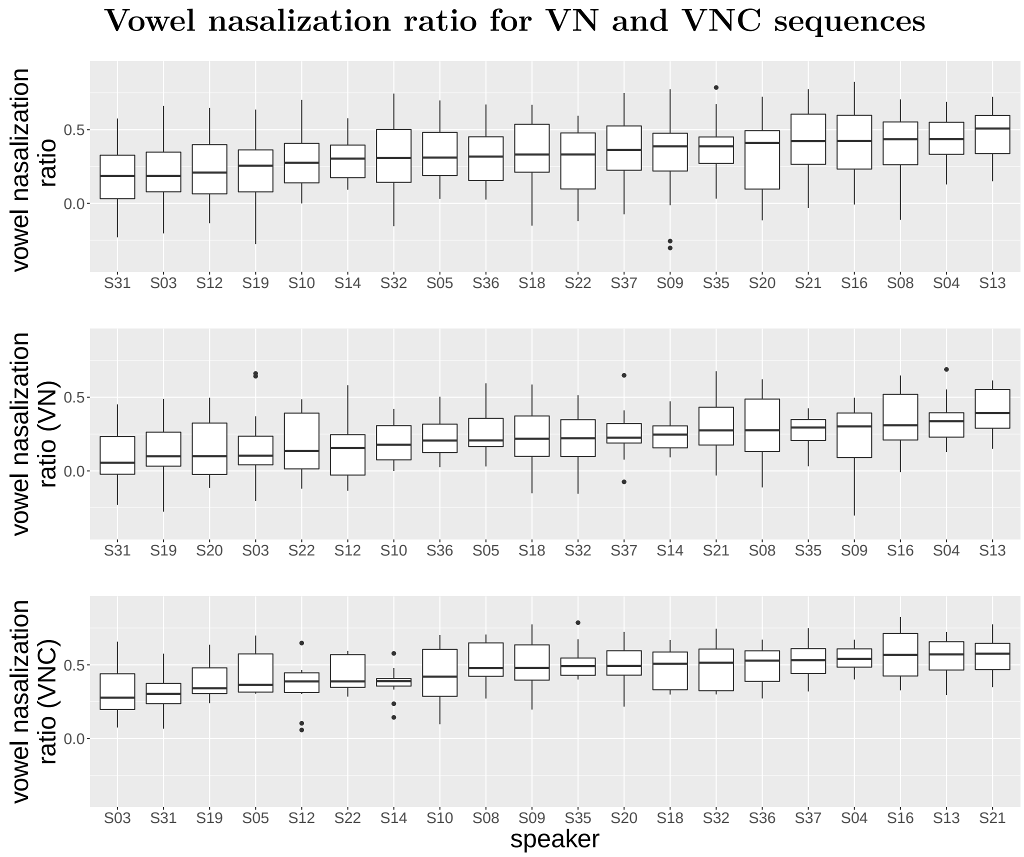


Figure B.6: Ratio of vowel nasality to OVL (vow.nas/OVL) for 20 speakers. Data include words with tense vowels only. Top: ratio for VN + VNC sequences; middle: ratio for VN sequences; bottom: ratio for VNC sequences.

Ratio referring to overall data: constant gesture

Figure B.7: Ratio of vowel nasality to OVL (y-axis: vow.nas/OVL) and perception thresholds (x-axis: means of the lowest six consecutive reversals in ms) for 20 speakers in the constant condition.

Ratio referring to overall data: extended gesture

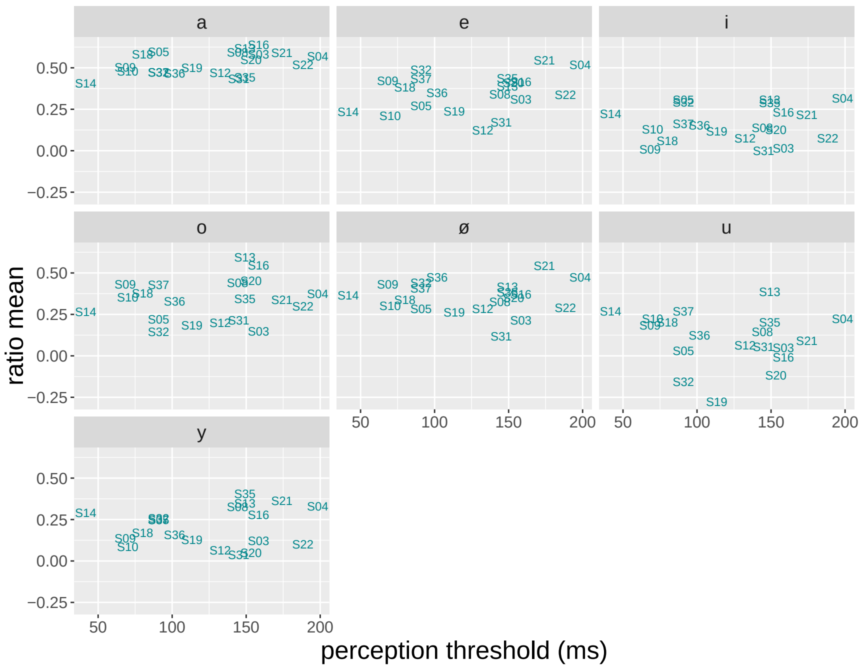


Figure B.8: Ratio of vowel nasality to OVL (y-axis: vow.nas/OVL) and perception thresholds (x-axis: means of the lowest six consecutive reversals in ms) for 20 speakers in the extended condition.

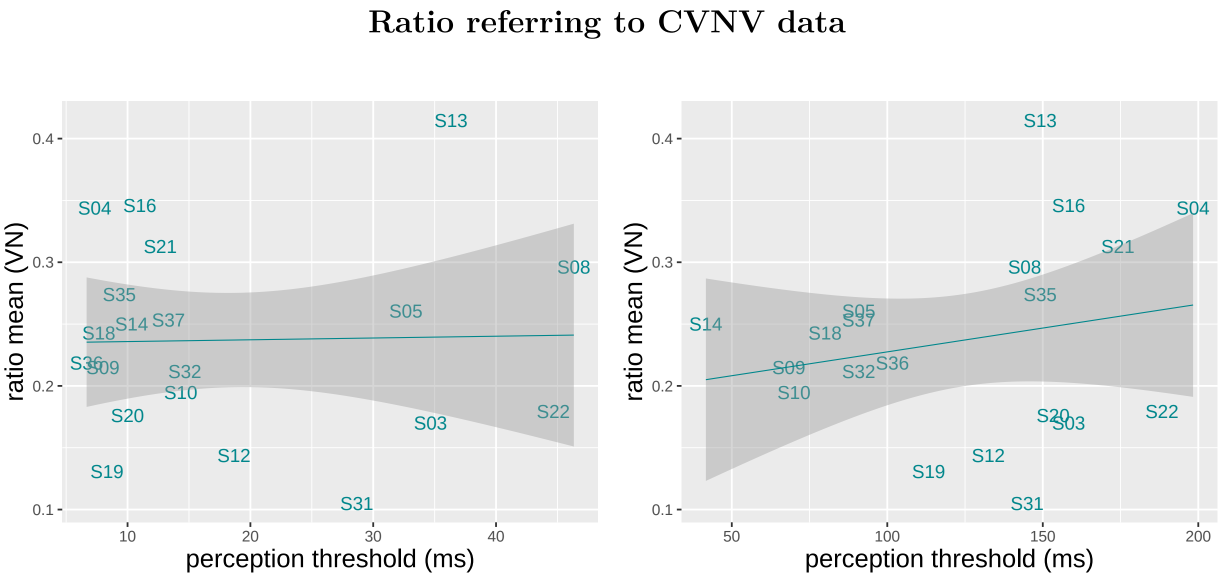


Figure B.9: Means of the nasality ratio (vow.nas/OVL) for words with VN sequences and the perception thresholds of the lowest six consecutive turning points for 20 participants. Results are shown for the constant (left) and extended (right) condition. A regression line indicates tendencies in correlation.

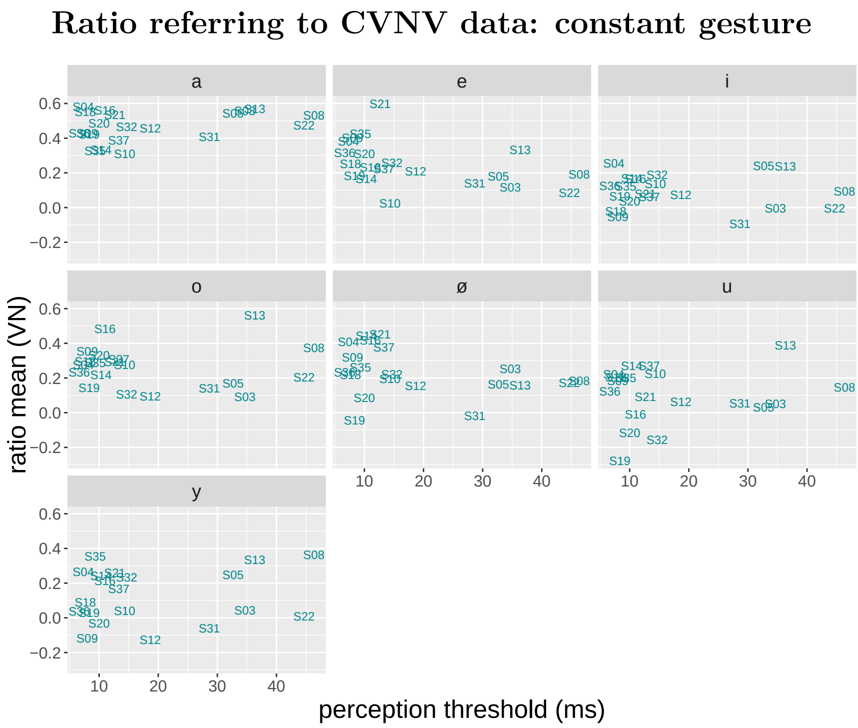


Figure B.10: Ratio of vowel nasality to OVL (y-axis: vow.nas/OVL) of target words with VN sequences and perception thresholds (x-axis: means of the lowest six consecutive reversals in ms) for 20 speakers. Results are shown for the constant condition.

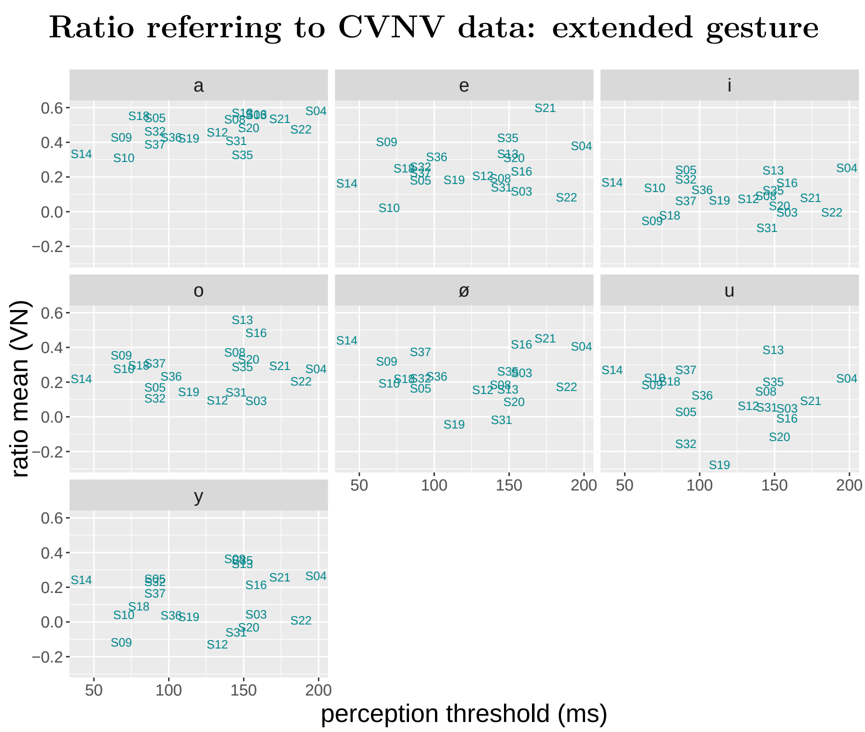


Figure B.11: Ratio of vowel nasality to OVL (y-axis: vow.nas/OVL) of target words with VN sequences and perception thresholds (x-axis: means of the lowest six consecutive reversals in ms) for 20 speakers. Results are shown for the extended condition.

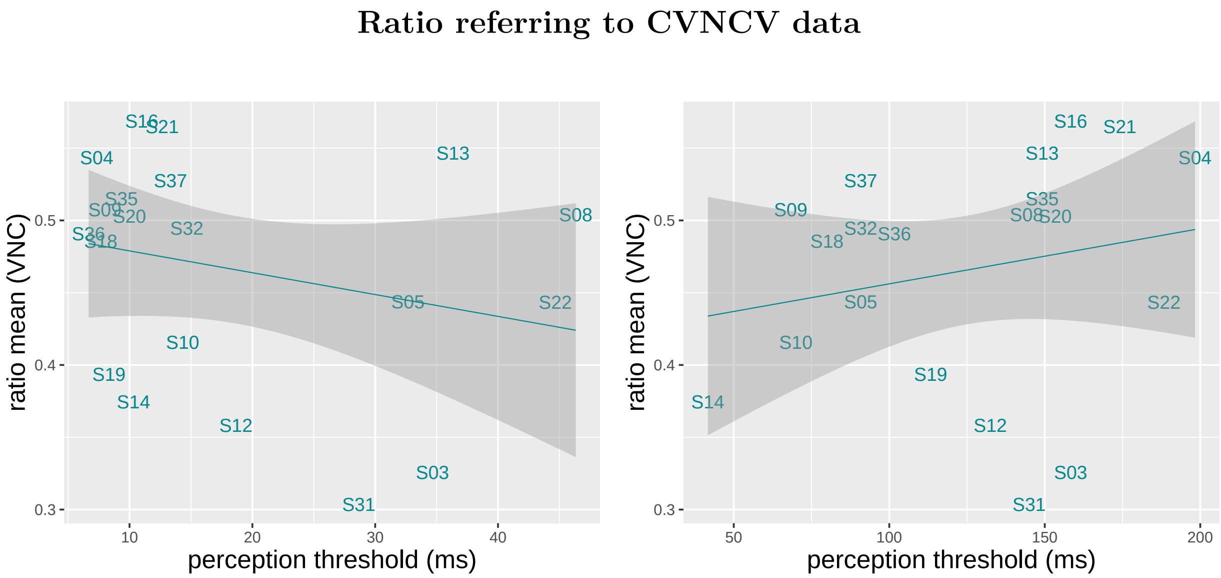


Figure B.12: Means of the nasality ratios (vow.nas/OVL) for words with VNC sequences and the perception thresholds of the lowest six consecutive turning points for 20 participants. Results are shown for the constant (left) and extended (right) condition. A regression line indicates tendencies in correlation.

Ratio referring to CVNCV data: constant gesture

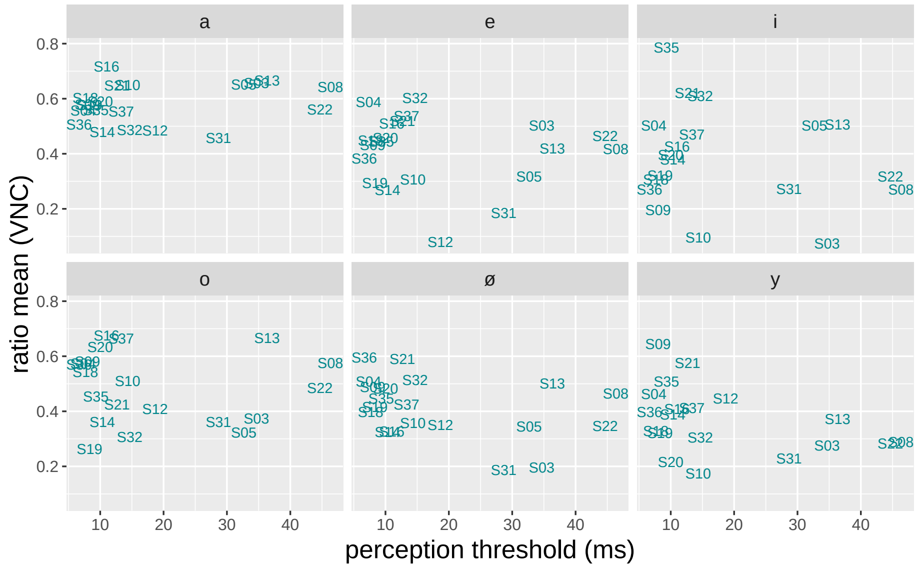


Figure B.13: Ratio of vowel nasality to OVL (y-axis: vow.nas/OVL) of target words with VNC sequences and perception thresholds (x-axis: means of the lowest six consecutive reversals in ms) for 20 speakers. Results are shown for the constant condition.

Ratio referring to CVNCV data: extended gesture

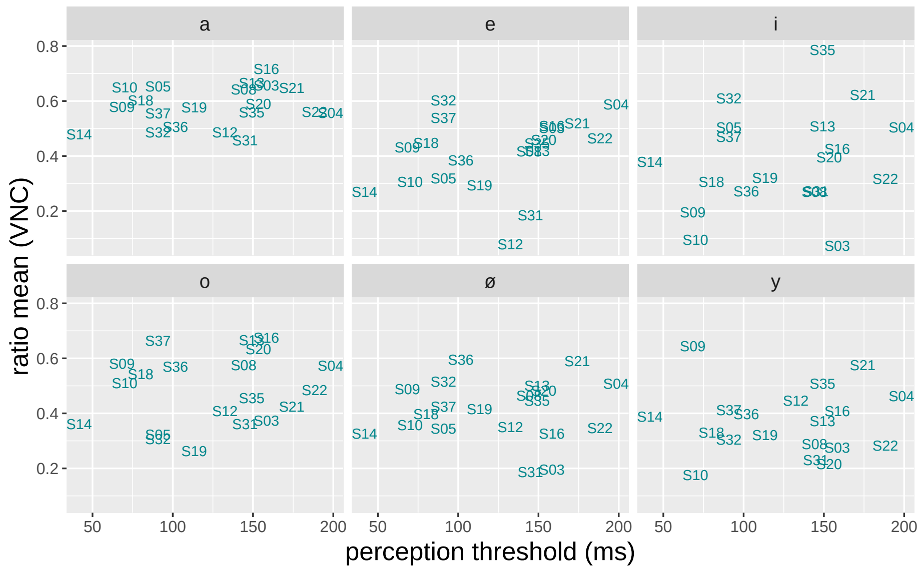


Figure B.14: Ratio of vowel nasality to OVL (y-axis: vow.nas/OVL) of target words with VNC sequences and perception thresholds (x-axis: means of the lowest six consecutive reversals in ms) for 20 speakers. Results are shown for the extended condition.

Appendix C

Demographics

Table C.1: Demographic information of the participants. M = Male, F = Female, NRW = North Rhine-Westphalia

Age	Sex	Hometown
22	M	Bremen, Bremen
25	F	Bad Vilbel, Hessen
31	M	Bernburg, Sachsen-Anhalt
26	F	Neuenbeken, NRW
25	F	Lintig, Niedersachsen
22	M	Göttingen, Niedersachsen
20	F	Bad Bodenteich, Niedersachsen
35	M	Halle, Sachsen-Anhalt
20	F	Gütersloh, NRW
28	F	Schwerte, NRW
25	M	Beverungen, NRW
21	F	Hameln, Niedersachsen
22	F	Göttingen, Niedersachsen
33	M	Wettenberg, Hessen
22	F	Leer, Niedersachsen
23	F	Fürth, Bayern
19	M	Stade, Niedersachsen
22	M	Glückstadt, Schleswig-Holstein
28	F	Lennestadt, NRW
30	F	Brilon, NRW
22	F	Rheinland
19	F	Sankt Augustin, NRW
22	F	Hann. Münden, Niedersachsen
22	F	Miehlen, Reinland-Pfalz

Table C.1: Demographic information of the participants. M = Male, F = Female, NRW = North Rhine-Westphalia

Age	Sex	Hometown
25	F	Oldenburg, Niedersachsen
23	M	Salzgitter, Niedersachsen
23	F	Hohenkirchen, Niedersachsen
24	M	Göttingen, Niedersachsen
35	M	Bad Lauterberg, Niedersachsen
22	F	Kiel, Schleswig-Holstein
25	F	Otterberg, Reinland-Pfalz
19	F	Dransfeld, Niedersachsen
25	F	Clausthal, Niedersachsen
22	M	Eschwege, Hessen
22	M	Seesen, Niedersachsen
28	M	Bregenstedt, Sachsen-Anhalt

Appendix D

Publications

The author of the thesis at hand has been among the authors of the following publications:

Carignan, C., Hoole, P., Kunay, E., Joseph, A., Voit, D., Frahm, J., and Harrington, J. (2019). The phonetic basis of phonological vowel nasality: Evidence from real-time MRI velum movement in German. In *Proceedings of the 19th International Congress of Phonetic Sciences*, pages 413–417. Melbourne, Australia.

Carignan, C., Hoole, P., Kunay, E., Pouplier, M., Joseph, A., Voit, D., Frahm, J., and Harrington, J. (2020). Analyzing speech in both time and space: Generalized additive mixed models can uncover systematic patterns of variation in vocal tract shape in real-time MRI. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 11(1). doi: `https://doi.org/10.5334/labphon.214`.

Carignan, C., Coretta, S., Frahm, J., Harrington, J., Hoole, P., Joseph, A., Kunay, E., and Voit, D. (2021). Planting the seed for sound change: Evidence from real-time mri of velum kinematics in german. *Language*, 97(2):333–364.

Niebergall, A., Zhang, S., Kunay, E., Keydana, G., Job, M., Uecker, M., and Frahm, J. (2013). Real-time MRI of speaking at a resolution of 33 ms: undersampled radial FLASH with nonlinear inverse reconstruction. *Magnetic Resonance in Medicine*, 69(2):477–485.

Bibliography

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

Ali, L., Gallagher, T., Goldstein, J., and Daniloff, R. (1971). Perception of coarticulated nasality. *The Journal of the Acoustical Society of America*, 49(2B):538–540.

Amelot, A., Honda, K., Maeda, S., Crevier-Buchman, L., and Basset, P. (2006). Evaluation of a photonasography. *The Journal of the Acoustical Society of America*, 120(5):3353–3353.

Amelot, A. and Rossato, S. (2006). Velar movements for the feature [±nasal] for two French speakers. In *International Seminar on Speech Production*, page 459–467, Ubatuba, Brasil.

Amelot, A. and Rossato, S. (2007). Velar movements for two French speakers. In *Proceedings of the 16th International Congress of Phonetic Sciences*, pages 6–10. Saarbrücken, Germany.

Basset, P., Amelot, A., Vaissière, J., and Roubeau, B. (2001). Nasal airflow in French spontaneous speech. *Journal of the International Phonetic Association*, pages 87–99.

Beddor, P. S. (2007). Nasals and nasalization: The relation between segmental and coarticulatory timing. In *Proceedings of the 16th International Congress of Phonetic Sciences*, pages 249–254. Saarbrücken, Germany.

Beddor, P. S. (2009). A coarticulatory path to sound change. *Language*, 85(4):785–821.

Beddor, P. S. (2012). Perception grammars and sound change. In Solé, M.-J. and Recasens, D., editors, *The Initiation of Sound Change: Perception, Production, and Social Factors*, pages 37–55. Benjamins Amsterdam, the Netherlands.

Beddor, P. S. (2015). The relation between language users' perception and production repertoires. In *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, U.K.

Beddor, P. S., Coetzee, A. W., Styler, W., McGowan, K. B., and Boland, J. E. (2018). The time course of individuals' perception of coarticulatory information is linked to their production: Implications for sound change. *Language*, 94(4):931–968.

Beddor, P. S. and Krakow, R. A. (1999). Perception of coarticulatory nasalization by speakers of English and Thai: Evidence for partial compensation. *The Journal of the Acoustical Society of America*, 106(5):2868–2887.

Beddor, P. S., McGowan, K. B., Boland, J. E., Coetzee, A. W., and Brasher, A. (2013). The time course of perception of coarticulation. *The Journal of the Acoustical Society of America*, 133(4):2350–2366.

Beddor, P. S. and Strange, W. (1982). Cross-language study of perception of the oral–nasal distinction. *The Journal of the Acoustical Society of America*, 71(6):1551–1561.

Bell-Berti, F. (1973). The velopharyngeal mechanism: An electromyographic study. *Status Report on Speech Research*. New York : Haskins Labs.

Bell-Berti, F. (1975). Control of pharyngeal cavity size for English voiced and voiceless stops. *The Journal of the Acoustical Society of America*, 57(2):456–461.

Bell-Berti, F. (1976). An electromyographic study of velopharyngeal function in speech. *Journal of Speech and Hearing Research*, 19(2):225–240.

Bell-Berti, F. (1993). Understanding velic motor control: studies of segmental context. In Huffman, M. K. and Krakow, R. A., editors, *Nasals, Nasalization, and the Velum*, pages 63–85. Elsevier.

Bell-Berti, F., Baer, T., Harris, K. S., and Niimi, S. (1979). Coarticulatory effects of vowel quality on velar function. *Phonetica*, 36(3):187–193.

Bell-Berti, F. and Hirose, H. (1973). Patterns of palatoglossus activity and their implications for speech organization. *The Journal of the Acoustical Society of America*, 54(1):320–320.

Bell-Berti, F. and Hirose, H. (1975). Palatal activity in voicing distinctions: a simultaneous fiberoptic and electromyographic study. *Journal of Phonetics*, 3(1):69–74.

Bell-Berti, F. and Krakow, R. A. (1991). Anticipatory velar lowering: A coproduction account. *The Journal of the Acoustical Society of America*, 90(1):112–123.

Birch, P., Gümoes, B., Stavad, H., Prytz, S., Björkner, E., and Sundberg, J. (2002). Velum behavior in professional classic operatic singing. *Journal of Voice*, 16(1):61–71.

Birkholz, P. (2013). Modeling consonant-vowel coarticulation for articulatory speech synthesis. *PloS one*, 8(4):e60603. Download: `https://www.vocaltractlab.de/index.php?page=vocaltractlab-download`.

Blumstein, S. E. (1986). On acoustic invariance in speech. *Invariance and variability in speech processes*, pages 178–193.

Boersma, Paul & Weenink, D. (2017). *Praat: doing phonetics by computer [Computer program]. Version 6.0.37, retrieved 3. April 2018 from http://www.praat.org/.*

Brooks, M. Z. (1968). *Nasal vowels in contemporary Standard Polish.* The Hague: Mouton.

Brucke, E. (1876). *Grundzüge der Physiologie und Systematik der Sprachlaute für Linguisten und Traub-Stummenlehrer.* Carl Gerold.

Brunelle, M. and Kirby, J. (2016). Tone and phonation in Southeast Asian languages. *Language and Linguistics Compass*, 10(4):191–207.

Busà, M. G. (2003). Vowel nasalization and nasal loss in Italian. In *Proceedings of the 15th International Congress of Phonetic Sciences*, pages 711–714. Barcelona, Spain.

Busà, M. G. (2007). Coarticulatory nasalization and phonological developments: Data from Italian and English nasal-fricative sequences. In Sole, M.-J., Beddor, P., and Ohala, M., editors, *Experimental Approaches to Phonology*, pages 155–174. Oxford University Press Oxford.

Busà, M. G. and Ohala, J. (1995). Nasal loss before voiceless fricatives: a perceptually-based sound change. *Rivista di Linguistica*, 7:125–144.

Byrd, D., Lee, S., Riggs, D., and Adams, J. (2005). Interacting effects of syllable and phrase position on consonant articulation. *The Journal of the Acoustical Society of America*, 118(6):3860–3873.

Byrd, D., Tobin, S., Bresch, E., and Narayanan, S. (2009). Timing effects of syllable structure and stress on nasals: a real-time MRI examination. *Journal of Phonetics*, 37(1):97–110.

Carignan, C. (2014). An acoustic and articulatory examination of the "oral" in "nasal": The oral articulations of French nasal vowels are not arbitrary. *Journal of Phonetics*, 46:23–33.

Carignan, C. (2018). An examination of oral articulation of vowel nasality in the light of the independent effects of nasalization on vowel quality. In *Associazione Italiana Scienze della Voce*, pages 19–40.

Carignan, C., Coretta, S., Frahm, J., Harrington, J., Hoole, P., Joseph, A., Kunay, E., and Voit, D. (2021). Planting the seed for sound change: Evidence from real-time mri of velum kinematics in german. *Language*, 97(2):333–364.

Carignan, C., Hoole, P., Kunay, E., Joseph, A., Voit, D., Frahm, J., and Harrington, J. (2019). The phonetic basis of phonological vowel nasality: Evidence from real-time MRI velum movement in German. In *Proceedings of the 19th International Congress of Phonetic Sciences*, pages 413–417. Melbourne, Australia.

Carignan, C., Hoole, P., Kunay, E., Pouplier, M., Joseph, A., Voit, D., Frahm, J., and Harrington, J. (2020). Analyzing speech in both time and space: Generalized additive mixed models can uncover systematic patterns of variation in vocal tract shape in real-time MRI. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 11(1). doi: `https://doi.org/10.5334/labphon.214`.

Carignan, C., Shosted, R., Fu, M., Liang, Z.-P., and Sutton, B. P. (2013). The role of the pharynx and tongue in enhancement of vowel nasalization: a real-time MRI investigation of French nasal vowels. In *INTERSPEECH*, pages 3042–3046.

Carignan, C., Shosted, R. K., Fu, M., Liang, Z.-P., and Sutton, B. P. (2015). A real-time MRI investigation of the role of lingual and pharyngeal articulation in the production of the nasal vowel system of French. *Journal of Phonetics*, 50:34–51.

Chen, M. (1972). *Nasals and nasalization in Chinese: explorations in phonological universals.* University of California, Berkeley.

Chen, M. (1973). *Nasals and nasalization in the history of Chinese.* PhD thesis.

Chen, N. F., Slifka, J. L., and Stevens, K. N. (2007). Vowel nasalization in American English: acoustic variability due to phonetic context. *Speech Communication*, pages 905–918.

Chi, Y., Honda, K., Wei, J., Feng, H., and Dang, J. (2015). Measuring oral and nasal airflow in production of Chinese plosive. In *Sixteenth Annual Conference of the International Speech Communication Association*, pages 2167–2171.

Cho, T. and Keating, P. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics*, 37(4):466–485.

Clumeck, H. (1976). Patterns of soft palate movements in six languages. *Journal of Phonetics*, 4(4):337–351.

Coetzee, A. W., Beddor, P. S., and Wissing, D. P. (2014). Emergent tonogenesis in Afrikaans. *The Journal of the Acoustical Society of America*, 135(4):2421–2422.

Cohn, A. C. (1990). *Phonetic and phonological rules of nasalization.* PhD thesis, University of California, Los Angeles.

Cronenberg, J., Gubian, M., Harrington, J., and Ruch, H. (2020). A dynamic model of the change from pre-to post-aspiration in Andalusian Spanish. *Journal of Phonetics*, 83:101016.

Dalston, R. M. (1982). Photodetector assessment of velopharyngeal activity. *The Cleft Palate Journal*, 19(1):1–8.

De Jong, K. J. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *The Journal of the Acoustical Society of America*, 97(1):491–504.

Delattre, P. (1946). Stages of Old French phonetic changes observed in Modern Spanish. *Publications of the Modern Language Association of America*, pages 7–41.

Delattre, P. and Monnot, M. (1968). The role of duration in the identification of French nasal vowels. *International Review of Applied Linguistics in Language Teaching*, 6(3):267.

Delvaux, V., Metens, T., and Soquet, A. (2002). French nasal vowels: acoustic and articulatory properties. In *Seventh International Conference on Spoken Language Processing*, pages 53–56.

Demolin, D., Delvaux, V., Metens, T., and Soquet, A. (2003). Determination of velum opening for French nasal vowels by magnetic resonance imaging. *Journal of Voice*, 17(4):454–467.

Demolin, D., George, M., Lecuit, V., Metens, T., Soquet, A., and Raeymaekers, H. (1997). Coarticulation and articulatory compensations studied by dynamic MRI. In *Fifth European Conference on Speech Communication and Technology*.

Demolin, D., Lecuit, V., Metens, T., Nazarian, B., and Soquet, A. (1998). Magnetic resonance measurements of the velum port opening. In *Fifth International Conference on Spoken Language Processing*.

Dixit, P., Bell-Berti, F., and Harris, K. S. (1987). Palatoglossus activity during nasal/nonnasal vowels of Hindi. *Phonetica*, 44(4):210–226.

Eefting, W. (1991). The effect of "information value"and "accentuation"on the duration of Dutch words, syllables, and segments. *The Journal of the Acoustical Society of America*, 89(1):412–424.

Eefting, W. (1992). The effect of accentuation and word duration on the naturalness of speech. *The Journal of the Acoustical Society of America*, 91(1):411–420.

Elster, A. D. (2021). Questions and answers in MRI. `https://www.mriquestions.com/index.html`. Accessed: 2019–2021.

Engwall, O. (2013). Assessing magnetic resonance imaging measurements. *Speech Production: Models, Phonetic Processes, and Techniques*, page 300.

Engwall, O., Delvaux, V., and Metens, T. (2006). Interspeaker variation in the articulation of nasal vowels. *Proceedings of the 7th ISSP*, pages 3–10.

Farnetani, E. and Vayra, M. (1996). The role of prosody in the shaping of articulation in Italian CV syllables. In *1st ETRW on Speech Production Modeling: From Control Strategies to Acoustics; 4th Speech Production Seminar: Models and Data*, pages 9–12.

Feng, G. and Castelli, E. (1996). Some acoustic features of nasal and nasalized vowels: A target for vowel nasalization. *The Journal of the Acoustical Society of America*, 99(6):3694–3706.

Ferguson, C. A. (1963). Assumptions about nasals: A sample study in phonological universals. *Universals of Language*, 53(60):337–339.

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, J. and Vatikiotis-Bateson, E. (1994). Prosody and intrasyllabic timing in French. *Working Papers in Linguistics*, (43):41–46.

Fougeron, C. (1999). Prosodically conditioned articulatory variations: A review. *UCLA Working*.

Fougeron, C. (2001). Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics*, 29(2):109–135.

Fowler, C. A. (1996). Listeners do hear sounds, not tongues. *The Journal of the Acoustical Society of America*, 99(3):1730–1741.

Fowler, C. A. and Brown, J. M. (2000). Perceptual parsing of acoustic consequences of velum lowering from information for vowels. *Perception & Psychophysics*, 62(1):21–32.

Fowler, C. A. and Housum, J. (1987). Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of memory and language*, 26(5):489–504.

Fowler, C. A. and Smith, M. R. (1986). Speech perception as "vector analysis": An approach to the problems of invariance and segmentation. In Perkell, J. S. and Klatt, D. H., editors, *Invariance and Variability in Speech Processes*, page 123–139. Lawrence Erlbaum Associates, Inc.

Frahm, J., Haase, A., and Matthaei, D. (1986). Rapid NMR imaging of dynamic processes using the FLASH technique. *Magnetic Resonance in Medicine*, 3(2):321–327.

Fritzell, B. (1969). The velopharyngeal muscles in speech. An electromyographic and cineradiographic study. *Acta Oto Larygol*, 250:1–78.

Fuchs, S., Perrier, P., Geng, C., and Mooshammer, C. (2005). What role does the palate play in speech motor control? Insights from tongue kinematics for German alveolar obstruents. In Harrington, J. and Tabain, M., editors, *Speech Production: Models, Phonetic Processes, and Techniques*, pages 149–164.

Fujimura, O. and Lindqvist, J. (1971). Sweep-tone measurements of vocal-tract characteristics. *The Journal of the Acoustical Society of America*, 49(2B):541–558.

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

Gay, T., Ushijima, T., Hiroset, H., and Cooper, F. S. (1974). Effect of speaking rate on labial consonant-vowel articulation. *Journal of Phonetics*, 2(1):47–63.

Gerrits, E. and Schouten, M. E. (2004). Categorical perception depends on the discrimination task. *Perception & Psychophysics*, 66(3):363–376.

Giot, J. (1977). Etude comparative de syllabes accentuées et prétoniques du français sur les plans articulatoire et acoustique. *Travaux de l'Institut de Phonétique de Strasbourg*, 9:89–169.

Goddard, I. (1965). The Eastern Algonquian intrusive nasal. *International Journal of American Linguistics*, 31(3):206–220.

Goddard, I. (1971). More on the nasalization of PA* a · in Eastern Algonquian. *International Journal of American Linguistics*, 37(3):139–145.

Grosvald, M. and Corina, D. (2012). The production and perception of sub-phonemic vowel contrasts and the role of the listener in sound change. In Solé, M.-J. and Recasens, D., editors, *The Initiation of Sound Change: Perception, Production, and Social Factors*, volume 323, pages 77–100. Benjamins Amsterdam, the Netherlands.

Gubian, M., Harrington, J., Stevens, M., Schiel, F., and Warren, P. (2019). Tracking the New Zealand English NEAR/SQUARE merger using functional principal components analysis.

Hajek, J. (1992). *The interrelationship between vowels and nasal consonants: A case study in Northern Italian.* PhD thesis, University of Oxford.

Hajek, J. (1997). *Universals of Sound Change in Nasalization*, volume 31. Blackwell Oxford.

Hajek, J. and Maeda, S. (2000). Investigating universals of sound change: The effect of vowel height and duration on the development of distinctive nasalization. In Broe, M. and J., P., editors, *Papers in Laboratory Phonology V*, pages 52–69. Cambridge University Press.

Hajek, J. and Watson, I. (1998). More evidence for the perceptual basis of sound change? Suprasegmental effects in the development of distinctive nasalization. In *Fifth International Conference on Spoken Language Processing*.

Harrington, J., Fletcher, J., and Beckman, M. (2000). Manner and place conflicts in the articulation of accent in Australian English. In *Papers in Laboratory Phonology V*, pages 40–55. Cambridge University Press.

Hattori, S., Yamamoto, K., and Fujimura, O. (1958). Nasalization of vowels in relation to nasals. *The Journal of the Acoustical Society of America*, 30(4):267–274.

Hawkins, S. and Stevens, K. N. (1985). Acoustic and perceptual correlates of the non-nasal–nasal distinction for vowels. *The Journal of the Acoustical Society of America*, 77(4):1560–1575.

Henderson, J. B. (1984). *Velopharyngeal function in oral and nasal vowels: A cross-language study*. PhD thesis, University of Connecticut.

Hertrich, I. and 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.

Hombert, J.-M., Ohala, J. J., and Ewan, W. G. (1979). Phonetic explanations for the development of tones. *Language*, 55(1):37–58.

Hoole, P. and Mooshammer, C. (2002). Articulatory analysis of the German vowel system. *Silbenschnitt und Tonakzente*, 1:129–152.

Hoole, P., Zierdt, A., and Geng, C. (2003). Beyond 2D in articulatory data acquisition and analysis. In *Proceedings of the 15th International Congress of Phonetic Sciences*, pages 265–268. Barcelona, Spain.

Horiguchi, S. and Bell-Berti, F. (1987). The velotrace: A device for monitoring velar position. *Cleft Palate Journal*, 24(2):104–111.

House, A. S. (1957). Analog studies of nasal consonants. *Journal of Speech and Hearing Disorders*, 22(2):190–204.

House, A. S. and Stevens, K. N. (1956). Analog studies of the nasalization of vowels. *Journal of Speech and Hearing Disorders*, 21(2):218–232.

Iltis, P. W., Schoonderwaldt, E., Zhang, S., Frahm, J., and Altenmüller, E. (2015). Real-time MRI comparisons of brass players: A methodological pilot study. *Human Movement Science*, 42:132–145.

Jaeger, M. and Hoole, P. (2011). Articulatory factors influencing regressive place assimilation across word boundaries in German. *Journal of Phonetics*, 39(3):413–428.

Kaernbach, C. (1991). Simple adaptive testing with the weighted up-down method. *Perception & psychophysics*, 49(3):227–229.

Katz, W., Machetanz, J., Orth, U., and Schönle, P. (1990). A kinematic analysis of anticipatory coarticulation in the speech of anterior aphasic subjects using electromagnetic articulography. *Brain and Language*, 38(4):555–575.

Kawasaki, H. (1986). Phonetic explanation for phonological universals: The case of distinctive vowel nasalization. In Ohala, J. J. and Jaeger, J. J., editors, *Experimental Phonology*, pages 239–252. Orlando, FL: Academic Press.

Kent, R. D., Carney, P. J., and Severeid, L. R. (1974). Velar movement and timing: Evaluation of a model for binary control. *Journal of Speech and Hearing Research*, 17(3):470–488.

Kent, R. D. and Moll, K. L. (1969). Vocal-tract characteristics of the stop cognates. *The Journal of the Acoustical Society of America*, 46(6B):1549–1555.

Kent, R. D. and Netsell, R. (1971). Effects of stress contrasts on certain articulatory parameters. *Phonetica*, 24(1):23–44.

Kingston, J. and Macmillan, N. A. (1995). Integrality of nasalization and F1 in vowels in isolation and before oral and nasal consonants: A detection-theoretic application of the Garner paradigm. *The Journal of the Acoustical Society of America*, 97(2):1261–1285.

Kirby, J. P. and Ladd, D. R. (2016). Effects of obstruent voicing on vowel f0: Evidence from "true voicing" languages. *The Journal of the Acoustical Society of America*, 140(4):2400–2411.

Kitamura, T., Takemoto, H., Honda, K., Shimada, Y., Fujimoto, I., Syakudo, Y., Masaki, S., Kuroda, K., Oku-Uchi, N., and Senda, M. (2005). Difference in vocal tract shape between upright and supine postures: Observations by an open-type MRI scanner. *Acoustical Science and Technology*, 26(5):465–468.

Kollmeier, J. M. (2021). *Multi-Directional Phase-Contrast Flow MRI in Real Time*. PhD thesis, Georg-August-Universität Göttingen.

Krakow, R. A. (1989). *The articulatory organization of syllables: A kinematic analysis of labial and velar gestures*. PhD thesis, Yale University.

Krakow, R. A. (1993). Nonsegmental influences on velum movement patterns: Syllables, sentences, stress, and speaking rate. In Huffman, M. K. and Krakow, R. A., editors, *Nasals, Nasalization, and the Velum*, pages 87–116. Elsevier.

Krakow, R. A. (1999). Physiological organization of syllables: a review. *Journal of Phonetics*, 27(1):23–54.

Krakow, R. A., Beddor, P. S., Goldstein, L. M., and Fowler, C. A. (1988). Coarticulatory influences on the perceived height of nasal vowels. *The Journal of the Acoustical Society of America*, 83(3):1146–1158.

Krakow, R. A. and Huffman, M. K. (1993). Instruments and techniques for investigating nasalization and velopharyngeal function in the laboratory: An introduction. In Huffman, M. K. and Krakow, R. A., editors, *Nasals, Nasalization, and the Velum*, pages 3–59. Elsevier.

Kroos, C., Hoole, P., Kühnert, B., and Tillmann, H. G. (1997). Phonetic evidence for the phonological status of the tense-lax distinction in German. *Forschungsberichte-Institut Für Phonetik Und Sprachliche Kommunikation Der Universität München*, (35):17–25.

Kuehn, D. P. (1976). A cineradiographic investigations of velar movement variables in two normals. *The Cleft Palate Journal*, 13(2):88–103.

Kuehn, D. P. and Azzam, N. A. (1978). Anatomical characteristics of palatoglossus and the anterior faucial pillar. *The Cleft Palate Journal*, 15(4):349–359.

Kuenzel, H. (1977). A photo-electric investigation of velar height during vowel articulation: First application of the velograph. *Arbeitsberichte*, 8:111–131.

Kühnert, B. and Hoole, P. (2004). Speaker-specific kinematic properties of alveolar reductions in English and German. *Clinical Linguistics & Phonetics*, 18(6-8):559–575.

Lahiri, A. and Marslen-Wilson, W. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition*, 38(3):245–294.

Lahiri, A. and Marslen-Wilson, W. D. (1992). Lexical processing and phonological representation. *Papers in Laboratory Phonology II*, pages 229–254.

Laver, J. (1994). *Principles of Phonetics*. Cambridge University Press.

Lee, Y.-J., Goldstein, L., and Narayanan, S. S. (2015). Systematic variation in the articulation of the Korean liquid across prosodic positions. In *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, U.K.

Lindblom, B. (1963). Spectrographic study of vowel reduction. *The Journal of the Acoustical Society of America*, 35(11):1773–1781.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In Hardcastle, W. J. and Marchal, A., editors, *Speech Production and Speech Modelling*, pages 403–439. Springer.

Lindblom, B., Guion, S., Hura, S., Moon, S.-J., and Willerman, R. (1995). Is sound change adaptive? *Rivista di Linguistica*, 7:5–36.

Lintz, L. B. and Sherman, D. (1961). Phonetic elements and perception of nasality. *Journal of Speech and Hearing Research*, 4(4):381–396.

Lubker, J. F. (1968). An electromygraphic-cineflourographic investigation of velar function during normal speech production. *The Cleft Palate Journal*, 5(1):1–18.

Lubker, J. F., Fritzell, B., and Lindqvist, J. (1970). Velopharyngeal function: An electromyographic study. *Speech Transmission Laboratory Quarterly Progress and Status Report*, 11:9–20.

Maddieson, I. (1984). *Patterns of sound*. Cambridge: Cambridge University Press.

Maddieson, I. (2007). Areal distribution of nasalized vowels. In *Proceedings of the 16th International Congress of Phonetic Sciences*, pages 1381–84. Saarbrücken, Germany.

Maeda, S. (1989). The distance between two nasal spectral peaks as an acoustic measure for vowel nasalization. In *Recueil de Publications et Communications 1988 en Traitement Automatique de la Parole [Collection of Publications and Communications in the Automatic Processing of Speech 1988]*, pages 275–304. Centre National d'Études des Télécommunications Lannion.

Maeda, S. (1993). Acoustics of vowel nasalization and articulatory shifts in French nasal vowels. In Huffman, M. K. and Krakow, R. A., editors, *Nasals, Nasalization, and the Velum*, pages 147–167. Elsevier.

Malécot, A. (1960). Vowel nasality as a distinctive feature in American English. *Language*, 36(2):222–229.

Matisoff, J. (1975). Rhinoglottophilia: The mysterious connection between nasality and glottality. In Ferguson, C. A., Hyman, L. M., and Ohala, J. J., editors, *Nasalfest: Papers from a Symposium on Nasals and Nasalization*, pages 289–316. Stanford: Language Universals Project.

Matsuya, T., Miyazaki, T., and Yamaoka, M. (1974). Fiberscopic examiniation of velopharyngeal closure in normal individuals. *The Cleft Palate Journal*, 11(3):286–291.

Matthaei, D., Frahm, J., Haase, A., and Hanicke, W. (1985). Regional physiological functions depicted by sequences of rapid magnetic resonance images. *The Lancet*, 326(8460):893.

McClean, M. (1973). Forward coarticulation of velar movement at marked junctural boundaries. *Journal of Speech and Hearing Research*, 16(2):286–296.

Meynadier, Y., Marchal, A., and Pitermann, M. (1998). Frontière syllabique, accent focal contrastif et coarticulation linguale dans le groupe consonantique/kskl/en position médiane de mot en français. In *XXIIè Journées d'Etudes sur la Parole*.

Mielke, J., Carignan, C., and Thomas, E. R. (2017). The articulatory dynamics of pre-velar and pre-nasal/æ/-raising in English: An ultrasound study. *The Journal of the Acoustical Society of America*, 142(1):332–349.

Moll, K. L. (1962). Velopharyngeal closure on vowels. *Journal of Speech and Hearing Research*, 5(1):30–37.

Moll, K. L. and Daniloff, R. G. (1971). Investigation of the timing of velar movements during speech. *The Journal of the Acoustical Society of America*, 50(2B):678–684.

Moll, K. L. and Shriner, T. H. (1967). Preliminary investigation of a new concept of velar activity during speech. *The Cleft Palate Journal*, 4(1):58–69.

Moon, J. B., Smith, A. E., Folkins, J. W., Lemke, J. H., and Gartlan, M. (1994). Coordination of velopharyngeal muscle activity during positioning of the soft palate. *The Cleft Palate-Craniofacial Journal*, 31(1):45–55.

Mooshammer, C., Fuchs, S., and Fischer, D. (1999). Effects of stress and tenseness on the production of CVC syllables in German. In *Proceedings of the 14th International Congress of Phonetic Sciences*, pages 409–412. San Francisco, CA, USA.

Niebergall, A., Zhang, S., Kunay, E., Keydana, G., Job, M., Uecker, M., and Frahm, J. (2013). Real-time MRI of speaking at a resolution of 33 ms: undersampled radial FLASH with nonlinear inverse reconstruction. *Magnetic Resonance in Medicine*, 69(2):477–485.

Ohala, J. (1975). Phonetic explanations for nasal sound patterns. In Ferguson, C. A., Hyman, L. M., and Ohala, J. J., editors, *Nasalfest: Papers from a Symposium on Nasals and Nasalization*, pages 289–316. Stanford: Language Universals Project.

Ohala, J. (2012). The listener as a source of change: An update. In Solé, M.-J. and Recasens, D., editors, *The Initiation of Sound Change: Perception, Production, and Social Factors*, pages 21–36. Benjamins Amsterdam, the Netherlands.

Ohala, J., Masek, C. S., Hendrick, R. A., and Miller, M. F. (1981). The listener as a source of sound change. In Masek, C. S., Hendrick, R. A., and Miller, M. F., editors, *Parasession on Language and Behavior*, pages 178–203. Chicago Linguistics Society, Chicago 1981.

Ohala, J. J. (1971). Monitoring soft palate movements in speech. *The Journal of the Acoustical Society of America*, 50(1A):140–140.

Ohala, J. J. (1993). Coarticulation and Phonology. *Language and Speech*, 36(2-3):155–170.

Ohala, J. J. (2005). Phonetic explanations for sound patterns. *A figure of speech: A festschrift for John Laver*, pages 23–38.

Ohala, J. J. and Lyberg, B. (1976). Comments on "temporal interactions within a phrase and sentence context". *The Journal of the Acoustical Society of America*, 59(4):990–992.

Ohala, J. J. and Ohala, M. (1993). The phonetics of nasal phonology: Theorems and data. In Huffman, M. K. and Krakow, R. A., editors, *Nasals, Nasalization, and the Velum*, pages 225–249. Elsevier.

Ohala, J. J. and Ohala, M. (1995). Speech perception and lexical representation: the role of vowel nasalization in Hindi and English. *Papers in Laboratory Phonology IV: Phonology and Phonetic Evidence*, pages 41–60.

Ohala, M. and Ohala, J. J. (1991). Nasal epenthesis in Hindi. *Phonetica*, 48(2-4):207–220.

Oliveira, C., Martins, P., and Teixeira, A. (2009). Speech rate effects on European Portuguese nasal vowels. In *Tenth Annual Conference of the International Speech Communication Association*, pages 480–483.

Onsuwan, C. (2005). *Temporal relations between consonants and vowels in Thai syllables.* PhD thesis, University of Michigan.

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

Otsuka, R., Ono, T., Ishiwata, Y., and Kuroda, T. (2000). Respiratory-related genioglossus electromyographic activity in response to head rotation and changes in body position. *The Angle Orthodontist*, 70(1):63–69.

Passavant, G. (1869). Ueber die Verschliessung des Schlundes beim Sprechen. *Archiv für pathologische Anatomie und Physiologie und für klinische Medicin*, 46(1):1–31.

Perkell, J. S., Cohen, M. H., Svirsky, M. A., Matthies, M. L., Garabieta, I., and Jackson, M. T. (1992). Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements. *The Journal of the Acoustical Society of America*, 92(6):3078–3096.

Perry, J. L. (2011a). Anatomy and physiology of the velopharyngeal mechanism. In *Seminars in Speech and Language*, volume 32, pages 83–92. © Thieme Medical Publishers.

Perry, J. L. (2011b). Variations in velopharyngeal structures between upright and supine positions using upright magnetic resonance imaging. *The Cleft Palate-Craniofacial Journal*, 48(2):123–133.

Proctor, M., Lo, C. Y., Narayanan, S., et al. (2015). Articulation of English vowels in running speech: A real-time MRI study. In *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, U.K., London: International Phonetic Association.

Proctor, M. and Walker, R. (2012). Articulatory bases of sonority in English liquids. In Parker, S., editor, *The Sonority Controversy*, pages 285–312. Mouton de Gruyter Berlin, Germany.

Proctor, M. I., Goldstein, L., Lammert, A. C., Byrd, D., Toutios, A., Narayanan, S. S., et al. (2013). Velic coordination in French nasals: a real-time magnetic resonance imaging study. In *INTERSPEECH*, pages 577–581.

Ramanarayanan, V., Tilsen, S., Proctor, M., Töger, J., Goldstein, L., Nayak, K. S., and Narayanan, S. (2018). Analysis of speech production real-time MRI. *Computer Speech & Language*, 52:1–22.

Raphael, L. J., Dorman, M. F., Freeman, F., and Tobin, C. (1975). Vowel and nasal duration as cues to voicing in word-final stop consonants: Spectrographic and perceptual studies. *Journal of Speech and Hearing Research*, 18(3):389–400.

Roon, K. D., Gafos, A. I., Hoole, P., and Zeroual, C. (2007). Influence of articulator and manner on stiffness. In *Proceedings of the 16th International Congress of Phonetic Sciences*, pages 409–412. Saarbrücken, Germany.

Rossato, S., Badin, P., and Bouaouni, F. (2003). Velar movements in French: an articulatory and acoustical analysis of coarticulation. In *Proceedings of the 15th International Congress of Phonetic Sciences*, pages 3141–3144. Barcelona, Spain.

Ruhlen, M. (1973). Nasal vowels. *Working Papers on Language Universals*, (12).

Sampson, R. (1999). *Nasal vowel evolution in Romance*. Oxford University Press on Demand.

Sauerland, E. and Mitchell, S. (1975). Electromyographic activity of intrinsic and extrinsic muscles of the human tongue. *Texas Reports on Biology and Medicine*, 33(3):444–455.

Schourup, L. C. (1973). A cross-language study of vowel nasalization. *Ohio State University Working Papers in Linguistics*, 15:190–221.

Seaver, I. and Kuehn, D. P. (1980). A cineradiographic and electromyographic investigation of velar positioning in non-nasal speech. *The Cleft Palate Journal*, 17(3):216–226.

Sefton, S. E. and Beddor, P. S. (2005). Nasals and nasalization: The interplay between segmental duration and coarticulation. *The Journal of the Acoustical Society of America*, 117(4):2620–2620.

Serrurier, A. and Badin, P. (2005). A three-dimensional linear articulatory model of velum based on MRI data. In *Ninth European Conference on Speech Communication and Technology*, pages 2161–2164.

Serrurier, A. and Badin, P. (2008). A three-dimensional articulatory model of the velum and nasopharyngeal wall based on MRI and CT data. *The Journal of the Acoustical Society of America*, 123(4):2335–2355.

Shosted, R., Carignan, C., and Rong, P. (2012). Managing the distinctiveness of phonemic nasal vowels: Articulatory evidence from Hindi. *The Journal of the Acoustical Society of America*, 131(1):455–465.

Shosted, R., Smith, J., and Ihsane, T. (2015). Nasal vowels are not [+ nasal] oral vowels. In *Romance Linguistics 2012: Selected papers from the 42nd Linguistic Symposium on Romance Languages (LSRL)*, pages 63–76. Jon Benjamins Amsterdam.

Silva, S. and Teixeira, A. (2015). Unsupervised segmentation of the vocal tract from real-time MRI sequences. *Computer Speech & Language*, 33(1):25–46.

Silva, S., Teixeira, A., Oliveira, C., and Martins, P. (2013). Segmentation and analysis of vocal tract from midsagittal real-time MRI. In *International Conference Image Analysis and Recognition*, pages 459–466. Springer.

Slis, I. H. (1971). Articulatory effort and its durational and electromyographic correlates. *Phonetica*, 23(3):171–188.

Solé, M.-J. (1992). Phonetic and phonological processes: The case of nasalization. *Language and Speech*, 35(1-2):29–43.

Solé, M.-J. (2007). Controlled and mechanical properties in speech. *Experimental approaches to phonology*, pages 302–321.

Steriade, D. (2000). Alternatives to syllable-based accounts of consonantal phonotactics. *Proceedings of the 1998 Linguists and Phonetic Conference.*

Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics*, 17(1-2):3–45.

Stevens, K. N. (2000). *Acoustic Phonetics*, volume 30. MIT press.

Stevens, K. N., Andrede, A., and Viana, M. C. (1987). Perception of vowel nasalization in VC contexts: A cross-language study. *The Journal of the Acoustical Society of America*, 82(S1):S119–S119.

Stone, M., Stock, G., Bunin, K., Kumar, K., Epstein, M., Kambhamettu, C., Li, M., Parthasarathy, V., and Prince, J. (2007). Comparison of speech production in upright and supine position. *The Journal of the Acoustical Society of America*, 122(1):532–541.

Straka, G. (1963). *La division des sons du langage en voyelles et consonnes peut-elle être justifiée?* C. Klincksieck.

Tobin, S., Byrd, D., Bresch, E., and Narayanan, S. (2006). Syllable structure effects on velum-oral coordination evaluated with real-time MRI. *The Journal of the Acoustical Society of America*, 119(5):3302–3302.

Toivonen, I., Blumenfeld, L., Gormley, A., Hoiting, L., Logan, J., Ramlakhan, N., and Stone, A. (2015). Vowel height and duration. In *Proceedings of the 32nd West Coast Conference on formal Linguistics*, volume 32, pages 64–71. Cascadilla Proceedings Project Somerville, MA.

Uecker, M., Zhang, S., Voit, D., Karaus, A., Merboldt, K.-D., and Frahm, J. (2010). Real-time MRI at a resolution of 20 ms. *NMR in Biomedicine*, 23(8):986–994.

Uecker, M., Zhang, S., Voit, D., Merboldt, K.-D., and Frahm, J. (2012). Real-time MRI: recent advances using radial FLASH. *Imaging in Medicine*, 4(4):461.

Ushijima, T. and Hirose, H. (1974). Electromyographic study of the velum during speech. *Journal of Phonetics*, 2(4):315–326.

Vaissière, J. (1988). Prediction of velum movement from phonological specifications. *Phonetica*, 45(2-4):122–139.

Weishaupt, Dominik; Köchli, V. D. M. B. (2014). *Wie funktioniert MRI? Eine Einführung in Physik und Funktionsweise der Magnetresonanzbildgebung*. Springer, 7 edition.

Whalen, D. H. (1990). Intrinsic velar height in supine vowels. *The Journal of the Acoustical Society of America*, 88(S1):S54–S54.

Whalen, D. H. and Beddor, P. S. (1989). Connections between nasality and vowel duration and height: Elucidation of the Eastern Algonquian intrusive nasal. *Language*, 65(3):457–486.

Wright, J. T. (1975). Effects of vowel nasalization on the perception of vowel height. In Ferguson, C. A., Hyman, L. M., and Ohala, J. J., editors, *Nasalfest: Papers from a Symposium on Nasals and Nasalization*, pages 373–388. Stanford: Language Universals Project.

Zellou, G. (2017). Individual differences in the production of nasal coarticulation and perceptual compensation. *Journal of Phonetics*, 61:13–29.

Zhang, S., Olthoff, A., and Frahm, J. (2012). Real-time magnetic resonance imaging of normal swallowing. *Journal of Magnetic Resonance Imaging*, 35(6):1372–1379.