Syllabic /l/ in Slovak and the effect of prosodic emphasis

Lia Saki Bučar Shigemori

SIGILL·UNIVERSITAT·LUDOVIC·MAXIM

München 2017

Syllabic /l/ in Slovak and the effect of prosodic emphasis

Inaugural-Dissertation
zur Erlangung des Doktorgrades der Philosophie
an der Ludwig–Maximilians–Universität
München

vorgelegt von
Lia Saki Bučar Shigemori
aus Ljubljana

2017

Erstgutachter: Prof. Dr. Marianne Pouplier

Zweitgutachter: Prof. Dr. Philip Hoole

Datum der mündlichen Prüfung: 12.02.2018

Contents

List of Figures

List of Tables

Zusammenfassung

Der Unterschied zwischen Vokalen und Konsonanten ist nicht allein anhand ihrer akustischen und artikulatorischen Merkmale festzulegen. Während insbesondere offene Vokale mit ihrem weit offenen Mundraum und Plosive mit einem vollständigen Verschluss zwei Extrempositionen einer Skala belegen, ist es schwieriger zu begründen, warum Approximanten zur Klasse der Konsonanten gehören, werden sie doch gelegentlich auch Halbvokale genannt. Der Unterschied zwischen Vokalen und Konsonanten ist zusätzlich mit ihrer Position innerhalb der Silbe und damit verbundenen Funktion gekoppelt. Da in den meisten Fällen Konsonanten um einen vokalischen Nukleus gegliedert sind, ist es oft nicht möglich die beiden direkt miteinander zu vergleichen. Nicht selten wird behauptet, dass Vokale die Prosodie tragen, während Konsonanten eher zur lexikalischen Verarbeitung beitragen. In dieser Hinsicht ist Slowakisch eine sehr untersuchungswerte Sprache, da die Konsonanten /l/ und /r/ auch im Nukleus einer betonten Silbe stehen können. In dieser Arbeit wurde mithilfe akustischer und artikulatorischer Aufnahmen untersucht, wie Satzakzent auf Silben mit konsonantischen Nuklei realisiert wird. Dabei beschränkten wir uns auf /l/ als konsonantischen Nukleus, mit Ausnahme in Kapitel 2, in dem beide silbischen Konsonanten in Betracht gezogen wurden.

Silbische Konsonanten sind typologisch keine Seltenheit. Jedoch ist in den meisten Sprachen der Kontext in dem sie auftreten können phonotaktisch eingeschränkter, verglichen mit vokalischen Nuklei. Im Deutschen sind silbische Konsonanten postlexikalisch und entstehen durch Vokalreduktion in unbetonten Silben und stellen eine Aussprachevariation dar. So kann das Verb *reden* sowohl als [redən], also mit einer Schwa+Konsonant-Endung als auch als [redn̩], also einem silbischen Konsonanten realisiert werden. Wie schon oben erwähnt sind silbische Konsonanten des Slowakischen keine Folge von Vokalreduktion in flüssiger Sprache sondern Bestandteil des Lexikons. Wie Vokale können sie sowohl in betonten als

auch unbetonten Silben auftreten, weisen einen phonologischen Längenkontrast auf wenn sie im Silbennukleus stehen und unterliegen den selben morpho-phonologischen Regeln, die die Veränderung der Nukleuslänge auslösen. Artikulatorisch weisen /r/, als auch /l/, das im Slowakischen dunkel ist, zusätzlich zur primären Zungenspitzengeste eine ausgeprägte Zugenrückenbewegung auf. Dadurch, dass diese Geste, ähnlich wie Vokale, keine enge Konstriktion bildet und sich langsamer bewegt, wird sie in der Literatur auch als 'vokalische' Geste bezeichnet. Es stellt sich also die Frage, ob genau diese Geste eine wichtige Rolle in der prosodischen Modulation spielt.

Im Slowakischen wird der Satzakzent mit erhöhter F0 auf der betonten Silbe markiert. Zusätzlich können Segmente einer akzentuierten Silbe oder eines akzentuierten Wortes durch Längung verstärkt werden. Es wurde zunächst im Kapitel 2 überprüft, ob diese zum gleichen Maße sowohl auf Silben mit vokalischen als auch konsonantischen Nuklei beobachtet werden können. Dies konnte bestätigt werden.

In akzentuierten Silben können an Vokalen auch Qualitätsunterschiede beobachtet werden. Die Sonoritätshyphothese (Beckman, Edwards, & Fletcher, 1992) besagt, dass unter Akzentuierung Kiefer und Zungenposition gesenkt werden, damit der Mundraum weiter geöffnet ist und eine höhere Sonorität entstehen kann. Die Hyperartikulationshypothese (De Jong, 1995) besagt, dass unter Akzentuierung die Merkmale des jeweiligen Lautes bestärkt werden. Für offene Vokale bedeutet das in beiden Fällen eine tiefere Zunge und ein tieferer Kiefer und damit verbunden ein offenerer Mundraum unter Akzentuierung. Für die apikale Geste des /l/ hingegen geraten die Vorhersagen der beiden Hypothesen in Konflikt. Einerseits sagt die Sonoritätshypothese, ungeachtet des Lautes der im Nukleus steht, einen weiten Vokaltrakt unter Akzentuierung voraus. Für /l/ bedeutet das eine Schwächung der apikalen Verengung. Die Hyperartikulationshypothese andererseits sagt eine deutlichere Verengung im vorderen Mundbereich durch die apikale Geste voraus. Eine ausgeprägtere dorsale Geste führt hingegen nicht notwendigerweise zur Verengung des vorderen Vokaltrakts wodurch die Vorhersagen der beiden Hypothesen im Einklang sind. Im Kapitel 3 wurden die Zungenbewegungen, die mittels Ultraschall aufgenommen wurden, sowie die Formantdaten untersucht. Da die apikale Verengung in vielen Fällen in den Ultraschallaufnahmen nicht sichtbar war, wurde der Bereich dahinter untersucht. Zum Vergleich wurde auch der vokalische Nukleus /e/ untersucht. Für /l/ wurde entgegen der Erwartungen kein Effekt von Satzakzent auf die Zungenrückengeste festgestellt. Der vordere Zungenbereich wurde während des Anlautplosivs gesenkt und im Laufe des akustisch ermittelten /l/ konnte unter Akzentuierung eine Aufwärtsbewegung der Zunge beobachtet werden. Wir interpretierten diese dynamische Bewegung als eine Umpositionierung der Zunge. Die so erlangte Zungenform erlaubt vermutlich einen größeren lateralen Luftstrom bei Beibehaltung der apikalen Verengung. Die Zungenbewegung von /e/ wies einen sogenannten 'trough-effect' auf. Sowohl bei vokalischen als auch konsonantischen Nuklei konnte im akzentuierten Fall ein höherer F1 beobachtet werden, der darauf hindeutet, dass der Kiefer gesenkt wurde und dadurch der Mundraum geweitet wurde. Unsere Daten deuten darauf hin, dass es nicht die 'vokalische' dorsale Geste ist, die im Falle des silbischen Konsonanten die Prosodie trägt.

Zur Prosodie gehört auch der Aufbau von Vokalen und Konsonanten zu Silben. So beobachteten unter anderem Sproat and Fujimura (1993) und Krakow (1999) für das amerikanische Englisch den systematischen Unterschied von /l/ in verschiedenen Silbenpositionen. Im Anlaut wurde /l/ mit einer deutlichen apikalen Verengung aber weniger ausgeprägten dorsalen Geste artikuliert, während im Auslaut die apikale Geste geschwächt war und dafür eine deutliche Rückwärtsbewegung des Zungendorsums sichtbar war. Auch die zeitliche Koordination der beiden Gesten zueinander unterscheidet sich, und so sind die apikale und dorsale Geste im Anlaut synchron zueinander, während im Auslaut die apikale Geste der dorsalen folgt. Generell zeichnet sich die Silbe durch bestimmte zeitliche Koordination von Anlaut und Auslaut mit dem Nukleus aus (Browman & Goldstein, 1988; Nam, Goldstein, & Saltzman, 2009). Für das Slowakische wurde in einer Studie von Pouplier and Beňuš (2011) gezeigt, dass der zeitlich Abstand zwischen einem Konsonanten und der apikalen Geste des /l/ oder /r/ größer ist, wenn /l/ oder /r/ im Nukleus stehen. In ihrer Studie wurde jedoch die dorsale Geste nicht genauer untersucht. Im Kapitel 4 wurde der Einfluss des Satzakzentes und der Silbenposition auf die beiden Gesten von /l/ untersucht. Unsere Erwartung war, dass vor allem für /l/ im Nukleus im akzentuierten Fall der Abstand vom vorangehenden Konsonanten und der apikalen Geste erweitert wird. So kann die 'vokalische' Geste zum Vorschein treten, was auch zur höheren Sonorität führt. Die Resultate deuteten darauf hin, dass /l/ im Anlaut und Auslaut sich zwar vom /l/ im Nukleus unterscheidet, dieser Unterschied sich aber auf den stärkeren Einfluss des Vokal im Nukleus zurückführen lässt. Im Gegensatz zum Englischen konnte auch keine systematische Veränderung der Koordination der beiden Gesten in unterschiedlichen Silbenpositionen festgestellt werden. Auch der zeitliche Abstand von der apikalen Geste zum benachbarten Konsonanten wurde unter Akzentuierung nicht signifikant vergrößert. Wir schlossen daraus, dass im Slowakischen die ‘vokalische’ dorsale Geste anders als im Englischen keine aktive prosodische Funktion trägt. Im Gegenzug konnte gezeigt werden, dass /l/ mit seiner ‘konsonantischen’ Eigenschaft als Prosodieträger fungieren kann.

Chapter 1

Introduction

The larger context within which this research is situated is to understand the overall restricted occurrence of syllabic consonants in the languages of the world. Universally, syllables with consonantal nuclei are not rare, yet they are much more restricted than vocalic nuclei. As pointed out by Hyman (2008), the most basic language universal is that all languages have vowels and consonants. Although utterances may exclusively consist of vowels or consonants, respectively, all languages alternate vowels and consonants in a more or less regular basis. This is so because vowels make consonants pronounceable and perceptible, while consonants, especially obstruents contribute to a “salient modulation” of the speech signal. Delattre, Liberman, and Cooper (1955) showed that the place of articulation of consonants is coded in the formant transition of the neighbouring vowel which points towards a fixed position, or locus, which indicates the place of articulation of the consonant. Vowels are optimal carriers of transitional cues of surrounding segments and at the same time are fully perceptible and distinguishable in isolation. However, alternating loud resonant vowels with consonants, particularly stops or fricatives, which differ substantially from vowels in terms of acoustics, optimizes the modulation of speech signal and enhances the “auditory and attentional impact” (Henke, Kaisse, & Wright, 2012, p. 76). It can be assumed that the basic consonant-vowel alternation pattern is biologically grounded in speech production and perception. This fundamental distinction of vowels and consonants is also recognized in the IPA-table where vowels and consonants are represented in two separate charts. Consonants are organized in a table based on their manner and place of articulation. Vowels are organized in a continuous chart in the form of a trapezium based on the openness of the vocal tract and the position of the tongue, so that the position of the vowel within the trapezium indicates the highest point of the arch of the tongue. While it is clear that consonants and vowels are the universal building blocks of spoken language, pinpointing the difference between these two sound categories is astonishingly difficult. As implied by the IPA-table, the difference between vowels and consonants can partially be based on their acoustic and articulatory properties. A very simplistic definition of a vowel is that it is a sound "produced without any kind of obstruction of the outgoing breath" (Ladefoged, 2001, p. 24). Correspondingly, consonants refer to all other sounds which are produced with a greater stricture than vowels, ranging from complete closure as for stops to only a narrow constriction as for approximants. However, not all vowels and consonants can be distinguished solely based on articulation or acoustics. Vowel-approximant pairs, in particular /u/ and /w/ or /i/ and /j/, have been described as being produced with identical or near identical vocal tract shapes. Approximants are also referred to as semi-vowels and in "The sounds of the world's langugages" Ladefoged and Maddieson (1996) classify them as "vowel-like consonants" and discuss these sounds within the chapter dedicated to vowels. It has been debated whether these sounds should be treated as allophones differing only in syllable position (Lehiste, 1964; Lisker, 1995; Swadesh, 1947). Thus, despite certain acoustic and articulatory features considered typical for either vowels or consonants, the two categories cannot be clearly separated based solely on these features. In most cases vowels and consonants also differ in their distribution within a syllable. Vowels usually occupy the syllable nucleus while consonants usually occupy the onset or coda position.

Many studies have investigated the functional differences between vowels and consonants and supported the idea that vowels carry the prosodic information while consonants contribute to lexical processing (Bonatti, Peña, Nespor, & Mehler, 2005; Nazzi, 2005; Nespor, Peña, & Mehler, 2003). In experiments in which participants had to recognize words from a continuous speech stream of an artificial language, they relied more on consonants than on vowels (Bonatti et al., 2005; Toro, Nespor, Mehler, & Bonatti, 2008).

The existence of phonological rules like vowel harmony (van der Hulst & van de Weijer, 1995) and consonant harmony (Rose & Walker, 2004) where either only vowels or only consonants are affected, skipping the segments in between, also suggests that the two categories are processed separately. Further support for the idea that vowels and consonants are fundamentally distinct categories has been found in studies looking at cortical activity. Bouchard, Mesgarani, Johnson, and Chang (2013) found that segments sharing similar phonetic features showed similar spatial patterns of cortical activity. In addition, they observed that vowels and consonants occupy distinct regions. Differences in vowel-consonant comparisons were always greater than in vowel-vowel or consonant-consonant comparisons. Another study using fMRI observed that vowels activate regions that are also responsible for prosodic processing, while consonants are processed in areas responsible for lexico-semantic processing (Carreiras & Price, 2008).

In sum, there are many reasons to treat consonants and vowels as categorically distinct, ranging from neurological, physiological to grammatical arguments. Thereby “physiological form” and “grammatical function” seem to be tightly interwoven. Against this background, the existence of syllabic consonants in Slovak words like *chlp* or *vrt* seems surprising. From all we know, about both, speech motor control and grammar, these kinds of words should not be part of spoken language. Correspondingly there are theoretical frameworks which by definition do not allow for vowel-less syllables, e.g. Government Phonology (Kaye, 1995; Lowenstamm, 1996). Nonetheless there is a general consensus that they exist, but little is known about the circumstances in which these types of words may appear and what their phonetic characteristics are. In this dissertation, we will investigate the syllabic consonants of Slovak since phonological rules of Slovak unequivocally proof these consonants to be in the nucleus. Slovak ideally lends itself to go beyond the usual association of the syllable edge occupied by consonants vs. syllable nucleus occupied by vowels. We will use Slovak as a window onto the question of whether syllabic consonants can take on one of the major functions of (vocalic) syllable nuclei, i.e. act as carriers of prosody.

The next section will provide an overview of the occurrence of syllabic consonants in the world’s languages and also summarize the small existing body of work regarding the phonetic properties of syllabic consonants. We will then turn to the articulatory and acoustic correlates of prosodic emphasis for consonants and vowels before presenting the phonological patterning of Slovak syllabic consonants.

1.1 Syllabic consonants

The main challenge with syllabic consonants probably is that it is not trivial to identify them. Their occurrence is highly dependent on the phonotactics of the language and the notion of syllables. In the majority of cases the syllable can be described as being organized according to a sonority scale. The sonority scale from the least sonorant to the most sonorant sound postulated by Clements (1990), also referred to as the Sonority Sequencing Principle and generally agreed upon is **obstruent < nasal < liquid < glide < vowel**. Often, obstruents are further divided into **plosives < fricatives** (e.g. M. K. Gordon, 2016, p. 99). The nucleus is occupied by the most sonorous sound, usually the vowel. The consonants are organized around the nucleus in a way that the sonority raises towards the nucleus and falls towards the margins. This suggests that syllable nuclei can be defined as sonority peaks, but many languages allow exceptions like ***sc**andal* in English and ***sp**aghetti* in Italian (fricative-plosive cluster in onset), or *Mo**ps*** in German (plosive-fricative coda). In these examples the fricative occupies a more peripheral position than the plosive within a syllable, resulting in a second sonority peak at the edge, in addition to the vocalic nucleus. Georgian allows long consonant clusters, such as *marts'q'v-s* (strawberry (DAT)), which dramatically violates the Sonority Sequencing Principle. Native speakers have been reported to syllabify such words inconsistently, and syllabification is not systematically accounted for in most Georgian grammars (Chitoran, 1998), with some linguists viewing the syllable as irrelevant for the description of Georgian phonotactics (Butskhrikidze, 2002). Interestingly, within the Slavic languages, Polish or Russian allow so called two-peak syllable patterns while Czech or Slovak do not (Sawicka, 2001): In Polish, words like *trwać* (to last), *bóbr* (beaver) or *myśl* (thought) are recognized as having only one syllable, while in Czech the liquids in the words *trvat* (to last), *bobr* (beaver) and *mysl* (sense) are syllabic, thus the words consist of two syllables (examples from Scheer (2008)).

Secondly, in many languages syllabic consonants arise as a consequence of vowel loss in unstressed syllables, as a possible realization in fluent speech. For example, in German, words like *haben* or *reden* can be produced as [ha:bən] or [redən] with a schwa + consonant sequence, but also as [ha:bm̩] or [redn̩] with a syllabic consonant, depending on speaking style (rules of reduction processes given for example in Kohler (1990)). In English words like *bottle* and *button* can be realized with a syllabic consonant or a schwa+consonant sequence (or in case of *bottle* also as [bɒʔɰ], with a vocalized /l/), depending on speaking style, dialect (Roach, Sergeant, & Miller, 1992; Wells, 1965), and also context (Toft (2002) observed that /n/ is usually produced without a preceding schwa after /t/ as in *beaten*, but with schwa after /p/ or /k/ as in *deepen* or *beacon*). In such languages where syllabic consonants alternate with a schwa + consonant realization and the syllable peak shifts to the vocalic segment if it is present, not all linguists recognize the consonant to be syllabic.

In contrast to languages with syllabic consonants as fluent speech phenomena in prosodically weak positions, there are also languages which have lexical syllabic consonants. Tashlhiyt Berber and Slovak belong to these languages. The status of vocalic transitions has been much discussed for Tashlhiyt Berber as well, a language known for allowing all kinds of segments to occupy the syllable nucleus, even voiceless stops. Syllabic consonants of Tashlhiyt Berber have been of particular interest to several researchers, because they are an exception compared to the majority of languages which allow only sonorants to occupy the syllable nucleus (Bell, 1978). Coleman (2001) claimed that syllabic consonants in Tashlhiyt Berber are phonologically schwa vowel + consonant sequences, where the vocalic elements should be analysed as the phonetic realization of schwa, while in other cases it is hidden behind the overlapping consonant gestures. As Bell (1978, p. 157) put it in his typological overview of syllabic consonants:

> The greatest difficulty was in connection with the concept of ‘segment of phonetic representation’ and the representation of transitional elements. Some investigators apparently feel that a consonant cannot be syllabic if even the slightest vocalic transition or release accompanies it.

However, versification and assibilation / spirantization provided strong arguments against a phonological representation of syllabic consonants in Tashlhiyt Berber as being schwa vowel + consonant sequences as proposed by Coleman (2001). These arguments are further supported by acoustic and articulatory analyses (Ridouane, 2008). Ridouane and Fougeron (2011) systematically investigated the occurrence of schwa in word-initial consonant sequences and found that at least one of the consonants had to be voiced and the transition from one consonant to another had to allow for a sufficiently open vocal tract to allow for an emergence of a vocalic transition. But above all, the occurence of the vocalic element was not dependent on the syllable affiliation of the consonants. While phonotactics of Tashlhiyt Berber has been a topic of several studies, stress and accent have been less analyzed with a recent exception (M. Gordon & Nafi, 2012; Grice, Ridouane, & Roettger, 2015; Roettger, 2017). Tashlhiyt Berber has been described to lack word stress (Roettger, 2017). Grice et al. (2015) came to the conclusion that tonal placement can be described as a probabilistic distribution determined by several factors. F0 peak was in general attracted to the final syllable, heavy syllables were preferred over light syllables and more sonorant segments over less sonorant segments (vowels over sonorant consonants, liquids over nasals). Thus, if the final syllable of the target word had a vocalic nucleus or was heavy, it was always produced with a F0 peak. However, if the nucleus of the final syllable was occupied by a sonorant consonant but the penultimate syllable had a vocalic nucleus, some speakers shifted the F0 peak to the penultimate syllable. In words consisting entirely of obstruents, the F0 peak was either placed on the preceding word, was not realized at all, or it was placed on a vocoid between the obstruents in the target word. This patterning contrasts with languages such as German and English in which tonal events align with stressed syllables. In the majority of languages, syllabic consonants occur in unstressed syllables or grammatical particles and affixes only (Bell, 1978), which suggests that in general, consonants are less suitable as carriers of prosody. Slovak, however, has fixed word stress, compared to no word stress in Tashlhiyt Berber, and tonal placement takes place on the stressed syllable. This latter point probably makes it more comparable to English and German in terms of prosodic patterning, with the important difference that in Slovak syllabic consonants can occupy the stressed syllable, and allows us to investigate the effect of phrasal accent on syllabic consonants as opposed to vowels. In Section 1.2 we will present some aspects of prosody relevant for this study, namely organization of segments into syllables and the effect of prosodic strengthening, and in Section 1.3 what is already known about word stress and phrasal accent in Slovak in more detail.

1.2 Prosody

Prosody generally refers to suprasegmental phenomena such as word stress, intonation or speech tempo, among others, and more generally the organization of segments into larger units.

In this study we will focus on the effect of phrasal accent. Accented syllables are one of the prosodic units where effects of prosodic strengthening or emphasis can be found, along with domain-initial and domain-final positions. At these positions, it has been observed that specific features or gestures of segments are enhanced. Studies focusing on domain edges have observed lengthening of vowels phrase finally (Beckman et al., 1992; Edwards, Beckman, & Fletcher, 1991), or longer constrictions for consonants phrase initially (Bombien, Mooshammer, Hoole, Rathcke, & Kühnert, 2007; Cho & Keating, 2009; Fougeron & Keating, 1997). Fougeron and Keating (1997) suggested that the phrase-final vowel and phrase-initial consonant strengthening work together to enhance the vowel-consonant contrast which work in opposite direction – greater constriction for consonants and less constriction for vowels – and consequently the contrast across boundaries. Similarly, it has been shown that in stressed and accented syllables, also referred to as prominent positions, vowels and consonants are articulated with greater spatio-temporal expansion compared to unstressed and unaccented syllables (Beckman et al., 1992; De Jong, 1995). In general, syllables in prominent positions are longer, with the strongest effect being on vowels (Fletcher, 2010; Greenberg, Carvey, Hitchcock, & Chang, 2003). However, the degree of lengthening of individual segments within a syllable is also language dependent. There are cases where the vowel is not lengthened. Heldner and Strangert (2001) compared the effect of accent in Swedish syllables consisting of a short onset consonant and either a long vowel and a short coda consonant (CV:C) or a short vowel and a long coda consonant (CVC:). They found that in syllables with long vowels all segments were lengthened, while in syllables with short vowels almost all lengthening occurred on the onset and coda consonants. This suggests that the nucleus is not necessarily lengthened but maybe the length contrast is enhanced in the accented condition. Vowels in accented position are generally also more peripheral in the acoustic F1-F2 vowel space and are articulated with more extreme gestures (Fletcher, 2010). Often they are also accompanied by a lower jaw position, which widens the oral cavity and enhances the sonority feature (Erickson, 2002; Harrington, Fletcher, & Beckman, 2000). In general consonant-vowel sequences in accented position are associated with greater and faster articulator movements (Harrington, Fletcher, & Roberts, 1995; Mücke & Grice, 2014). However, when it comes to the effect of phrasal accent on consonant articulation regarding constriction strength or feature enhancement the interpretation of results is more complex. Cho and Keating (2009) investigated the effect of boundary and accent on consonants and found that the degree and duration of linguo-palatal contact measured by means of EPG of /n/ and /t/ was affected by boundary strength but not by accent. Yet, a greater RMS burst energy for /t/ and nasal energy for /n/ was observed as an effect of accent.

By signalling boundaries and prominent segments prosodic emphasis aids to organizing syllables and words into greater constructs like sentences. Another aspect of prosody is the organization of smaller segments into syllables, or in most cases the organization of consonants around a vocalic nucleus. Based on acoustic data, Öhman (1966) observed that within a VCV sequence "traces of the final vowel are observable already in the transition from the initial vowel to the consonant" (p.165) and suggested that consonants are superimposed on a continuous vowel articulation. More recently, it has been shown that segments are organized into syllables following certain timing patterns (Browman & Goldstein, 1988; Krakow, 1999). For syllables in American English it has been observed that the timing between the onset and the vocalic nucleus differs significantly from the timing between the vocalic nucleus and the coda. Onset consonants, but not coda consonants are timed synchronously with the nucleus, meaning that the movement of the tongue towards the target for the nucleus vowel already begins with the movement towards the target for the onset consonant. However, this synchronous timing in particular is possible due to the kinematic differences between vowels and consonants (Goldstein, Byrd, & Saltzman, 2006, p. 232):

> (1) Vowel gestures are less constricted than (stop or fricative) consonant gestures, so the vowel can be produced during consonant production without interfering with the acoustic properties of the consonant that signal its presence (because the more narrow a constriction, the more it dominates the source and resonance properties of an acoustic tube). (2) Vowel gestures are formed more slowly and are active longer than consonant gestures, so they dominate the acoustics of the tube during a time with no overlapping or only weakly competing consonant

gestures.

It has further been shown that consonants differ systematically, depending on whether they occupy the onset or coda position (Krakow, 1999). Consonants in onset position are articulated with a stronger excursion, while coda consonants are more often reduced (Byrd, 1996). As a result, allophonic variation like light and dark /l/ (Giles & Moll, 1975; Sproat & Fujimura, 1993) in English or neutralization of phonemic contrasts as in final devoicing in German (Wiese, 2000) can be observed. In the context of syllabic consonants, Fougeron and Ridouane (2008) and Pouplier and Beňuš (2011) investigated for Tashlhiyt Berber and Slovak, respectively, whether consonants show additional allophonic variations making them more "vowel like" when they occupy the nucleus position. They also asked whether the temporal coordination of consecutive consonants depends on their syllable affiliation, showing patterns similar to sequences with vocalic nuclei. Contrary to expectation both, Fougeron and Ridouane (2008) and Pouplier and Beňuš (2011) did not find that the consonant became more "vowel like" kinematically when occupying the syllable nucleus. Fougeron and Ridouane (2008) even observed a stronger constriction and faster movement for the velar consonant when it occupied the nucleus position. Furthermore, consonants in nucleus position were not acoustically longer than when they occupied the onset or coda position. Pouplier and Beňuš (2011) found that the main kinematic differences were between the consonants in onset and coda positions, and consonants in nucleus position patterned with either the onset or the coda. Thus in Tashlhiyt Berber as well as Slovak, consonants in nucleus position were not found to have a less narrow constriction, become slower or be longer than when they occupied the onset or coda conditions. However, differing temporal patterning could be observed as an effect of syllable affiliation. Fougeron and Ridouane (2008) found that the consonant in nucleus position is more overlapped with the onset consonant and less overlapped with the following consonant. Pouplier and Beňuš (2011) looked at the temporal lag between two consonants, where the consonants were either an onset or coda cluster or where one of the consonants occupied the nucleus. The lag was greater when one of the consonants was syllabic compared to when both consonants were part of the onset or coda cluster. The lag was also greater between consonants in the onset cluster compared to the lag between consonants in the coda cluster. Likewise the lag was greater between the consonants in onset and nucleus positions than between the consonants in nucleus and coda positions. The temporal coordination for consonants in Tashlhiyt Berber and Slovak seem to oppose each other, however it should be noted that different measures were used in the two studies. Fougeron and Ridouane (2008) using EPG data measured overlap as the percentage of the duration of the syllabic consonant during which contact associated to either the preceding or following consonant could be measured. Pouplier and Beňuš (2011) used EMA and measured overlap in terms of plateau lag, thus the time between the release of the first consonant and the achievement of the target of the second consonant. The two measures would yield different results for example depending on how slowly the articulator moves away from the target position of the first consonant. Most importantly, in both studies systematic differences in temporal coordination as an effect of syllable position was observed. Pouplier and Beňuš (2011) looked at only the apical gesture and did not investigate the properties of the dorsal gesture and consequently the temporal coordination of the apical and dorsal gesture of /l/ and /r/, the two consonants that can occupy the syllable nucleus in Slovak. Slovak /l/ has been described to be 'dark', for which a retracted tongue dorsum is characteristic (Sproat & Fujimura, 1993). Apical trills have also been described to require a dorsal retraction of the tongue (Kavitskaya, Iskarous, Noiray, & Proctor, 2009). Sproat and Fujimura (1993) investigated the articulatory strength in terms of constriction degree and temporal coordination of the two gestures of /l/ for American English word initially and in different pre-boundary positions. They found different patterns for the two gestures not only between the word-initial onset position and coda positions but also for different pre-boundary rime durations, suggesting that not only syllable affiliation but other prosodically conditioned factors affect gestural coordination. We will thus try to investigate both, the apical and the dorsal gesture to add to the findings by Pouplier and Beňuš (2011) and also look at the effect of phrasal accent.

1.3 Slovak

Slovak is particularly interesting when it comes to syllabic consonants as it has morpho-phonological rules targeting the syllable nucleus, regardless of whether it is occupied by a vowel or a consonant. These rules, which stand in contrast to rules such as vowel or consonant harmony which target either vowels or consonants, unequivocally proof that consonants can occupy the nucleus. In this section we will give a short introduction of Slovak including its orthography and phoneme inventory and the context in which syllabic consonants can occur, followed by some relevant grammatical aspects. As will be pointed out, these grammatical phenomena allow the identification of syllabic consonants. Finally, we will present relevant phonetic studies on Slovak prosody.

Slovak is a West Slavic language which belongs to the group of Indo-European languages and is mainly spoken in Slovakia. According to Simons and Fennig (2017) it is spoken by 5,149,000 people as L1 world wide. Slovak can be divided in three dialect groups: West Slovak, Central Slovak and East Slovak, of which Central Slovak is regarded as the standard language which has been described by linguists. An overview of the phonemic inventory of Slovak consonants and monophthong vowels as provided by Hanulíková and Hamann (2010) is shown in Tables 1.1 and 1.2. Additionally, Slovak has four diphthongs: <*ia, ie, iu*> and <*ô*> [ɪ̯a ɪ̯ɛ ɪ̯u̞ ʊ̯ɔ]. Examples in this study will be given using Slovak orthography, because it reflects pronunciation quite well. The pronunciation of letters which are less intuitive are given in Table 1.3.

Table 1.1: Consonant inventory of Slovak.

	Bilabial	Labiodental	Alveolar	Postalveolar	Palatal	Velar	Glottal
Plosive	p b		t d		c ɟ	k g	
Nasal	m		n		ɲ		
Fricative		f v	s z	ʃ ʒ		x	ɦ
Affricate			t͡s d͡z	t͡ʃ d͡ʒ			
Trill			r r:				
Approximant					j		
Lateral approximant			l l:		ʎ		

Table 1.2: Vowel inventory of Slovak.

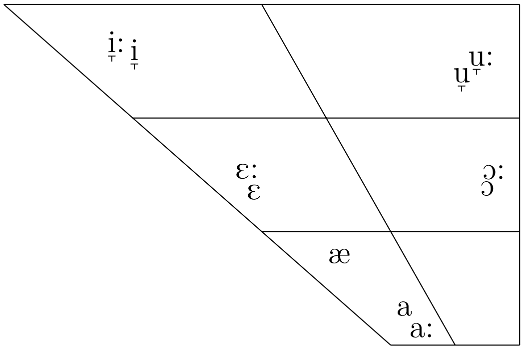


Table 1.3: Correspondences between Slovak orthography and pronunciation.

Orthography	IPA symbol	Description
´	ː	phonemic length
y, i	i	high front vowel
ä	æ	low front vowel
ô	ʊ̯ɔ	rising diphthong
c	ts	alveolar affricate
č, š, ž	tʃ, ʃ, ʒ	palatal affricate and fricatives
t’, d’, l’, ň	c, ɟ, ʎ, ɲ	palatal plosives, lateral and nasal
ch	x	voiceless velar fricative
h	ɦ	voiced glottal fricative

Slovak allows the consonants /l/ and /r/ to occupy the nucleus of a lexically stressed or unstressed syllable. According to the description of Slovak phonotactics given by Pouplier and Beňuš (2011), syllabic consonants require at least one onset consonant and can take an onset cluster consisting of up to three consonants. Syllabic consonants must also be followed by at least one consonant, which can either be a coda or an onset consonant of the following syllable in polysyllabic words. The set of coda clusters is more limited than the set of onset clusters, but in Slovak complex codas are generally infrequent, also with vocalic nuclei. Syllabic consonants never occur word initially or word finally. In nucleus position, /l/ and /r/ can be phonemically short or long, while consonants in onset or coda position do not show phonemic length alternation. Orthographically, the long counterparts are marked with an acute accent, analogously to vowels. There are no minimal pairs for /l/ and /l:/ or /r/ and /r:/, but the length alternation is triggered by the same laws that apply for vowels, usually when words are inflected (Hanulíková & Hamann, 2010; Kenstowicz & Rubach, 1987). A list of minimal pairs and near minimal pairs of short and long nuclei is given in (1), and some examples of words with /l/ and /r/ in onset, nucleus and coda are given in (2). The sample of rules that trigger length alternation include the formation of genitive plural (3), formation of a diminutive (4), certain verb suffixes (5) and the Rhythmic Law (6). The Rhythmic Law states that if through affixation two consecutive syllables happen to have a long nucleus, the second of the two is shortened. The phonemic length contrast is robust and long nuclei are about twice as long as short nuclei, which has been empirically shown for stressed syllables in accented position (Beňuš & Mády, 2010; Beňuš, 2011).

(1) Minimal pairs or near minimal pairs of words with short and long nuclei (taken from Rubach (1993, p. 39)).

krik (shout)	krík (bush)
kur+a (chicken)	kúr+a (cure)
leno (but)	lén+o (feud)
mol (mole)	mól+o (pier)
rad (row)	grád (degree)
vlk (wolf)	tĺk (pestle)
krčm+a (inn)	kŕč (cramp)

(2) Minimal pairs or near minimal pairs of words with l or r in onset, nucleus and coda.

Klak (geographic name)	klk (villus-Nom.Sg.)	kalk (calque-Nom.Sg.)
mrak (cloud-Nom.Sg.)	mrk (wink-Nom.Sg.)	park (park-Nom.Sg.)

(3) Lengthening in genitive plural (taken from Rubach (1993, p. 167)).

a. Genitive plural of feminine nouns.

hal+a (hall)	hál
slin+a (saliva)	slín
ryb+a (fish)	rýb
slz+a (tear)	sĺz
črt+a (feature)	čŕt

b. Genitive plural of neuter nouns.

blat+o (mud)	blát
piv+o (beer)	pív
koryt+o (trough)	korýt
jablk+o (apple)	jabĺk
zrn+o (grain)	zŕn

(4) Lengthening through suffixation of diminutive ending (taken from Rubach (1993, p. 168)).

a. Masculine:

hlas (voice)	hlás+ok
sud (cup)	súd+ok
chlp (hair)	chĺp+ok
srp (sickle)	sŕp+ok

b. Feminine:

hlav+a (head)	hláv+k+a
žrd (stick)	žŕd+k+a

(5) Shortening through suffixation of verbs as in forming imperfective verbs (taken from Pouplier and Beňuš (2011, p. 248).

zváž-i-t' (think)	zvaž-ova-t'
zníž-i-t' (lower)	zniž-ova-t'
predĺž-i-t' (lengthen)	predlž-ova-t'
vykŕm-i-t' (feed)	vykrm-ova-t'

(6) Rhythmic Law (taken from Pouplier and Beňuš (2011, p. 248).

a. Suffix following a short nucleus.

ryb-a (fish)	ryb-ách	ryb-ám
ruk-a (hand)	ruk-ách	ruk-ám
srn-a (deer)	srn-ách	srn-ám
vln-a (wave)	vln-ách	vln-ám

b. Suffix following a long nucleus.

tráv-a (grass)	tráv-ach	tráv-am
lúk-a (meadow)	lúk-ach	lúk-am
dĺžk-a (length)	dĺžk-ach	dĺžk-am
vŕb-a(willow)	vŕb-ach	vŕb-am

The realization of phrasal accent on syllables with consonantal nuclei in particular has to our knowledge not been empirically studied yet, but according to standard literature on Slovak, the accented syllable is in general marked primarily by a raised F0 (Kráľ, 2005). A few studies exist which looked at general intonation patterns in terms of F0 contours in Slovak but they were not concerned with the acoustic and articulatory realization of smaller segments (Beňuš, Reichel, & Mády, 2014; Rusko, Sabo, & Dzúr, 2007). There are however studies which examined the effect of other prosodic aspects on acoustics and articulation of syllable nuclei like Lombard speech (Šimko, Beňuš, & Vainio, 2016), speech tempo and word stress (Beňuš & Pouplier, 2011; Beňuš & Mády, 2010; Beňuš, 2011). In Slovak, word stress is fixed on the first syllable making stress postlexical. Although based on only two speakers, Beňuš and Mády (2010) examined the effect of word stress and speech rate on the quality and duration of phonemically long and short vocalic nuclei. They found that the phonemic quantity contrast is robust in all conditions, with long vowels being twice as long as short vowels on average and at least 1.5 times longer. Vowels are shorter in fast rate compared to normal rate and in unstressed syllables compared to stressed syllables. However, greater durational differences were observed on long vowels. Contrary to claims by the standard literature (Kráľ & Sabol, 1989) and intuitions by natives speakers Beňuš and Mády (2010) also observed that vowel quality is affected by prosody. Vowels in unstressed position are more centralized than in stressed position with the effect being greater for short vowels. Yet phonemic vowel categories remained separated in all conditions. The effect of phonemic quantity and speech rate has been tested for consonantal nuclei as well, which does not differ from vocalic nuclei (Beňuš & Pouplier, 2011; Beňuš, 2011). While the acoustic properties (except for duration) of Slovak consonantal nuclei has so far not been examined, several studies focused on the kinematic and coordination patterns of syllables with consonantal nuclei obtained by means of elctromagnetic articulography (EMA) (Beňuš, 2014; Beňuš & Pouplier, 2011; Beňuš, 2011; Pouplier & Beňuš, 2011). The measures such as plateau duration, peak velocity, time to peak velocity and stiffness of the tongue tip gesture of consonants in nucleus position did not differ substantially from when they occupied the onset or coda position. As mentioned already in Section 1.2, the syllabicity of consonants was related to their temporal coordination with preceding and following consonants which was different to when /l/ or /r/ were part of the onset or coda cluster. However, the fact that syllabic consonants are less overlapped with the preceding or following consonant compared to when they are part of the onset or coda cluster is counterintuitive, particularly given that onset consonants are said to be timed synchronously with the vocalic nucleus. Beňuš (2014) thus compared whether the 'vocalic' dorsal gesture of syllabic /l/ and /r/ was timed to the onset consonant in a similar manner as the tongue gesture for a vowel. While the temporal patterning of the dorsal gesture of syllabic consonants differed from the tongue gesture of vowels - the dorsal gesture for /l/ was more overlapped with the onset consonant and the dorsal gesture for /r/ was less overlapped than the tongue gesture of vowels with the onset consonant - it was more similar than the patterning of the tongue tip gesture. Finally, Beňuš and Pouplier (2011) compared the effect of speech rate and phonemic quantity on jaw displacement of vocalic and consonantal nuclei. A lower jaw allows for a higher acoustic energy to be radiated, increasing the sonority. The ordering of segments in syllables preferably correlates with their intrinsic jaw position. Indeed, Beňuš and Pouplier (2011) showed that the jaw displacement of syllabic consonants patterned with the jaw displacement of high vowels.

1.4 Research aims

Slovak is not as unusual regarding the consonants allowed in the syllable nucleus as for example Tashlhiyt Berber, which is known for allowing even obstruents as syllabic consonants. However, Slovak phonotactics clearly define the position of the syllable nucleus and the primarily stressed syllable, which allows us to carry out a highly controlled experiment.

One the one hand, we compare vowels with consonants in syllable nucleus position. On the other, we compare consonants in nucleus position with those in onset and coda position. For both comparisons, we will investigate how the segments are affected by phrasal accent.

In Chapter 2 we will examine how F0 and acoustic duration are affected by phrasal accent to ascertain that the syllable with the consonantal nucleus is indeed accented if necessary. In Chapter 3 the articulatory implementation of prosodic emphasis will be examined. Current theories base their predictions on the assumption that the nucleus is vocalic. Thus our main motivation is to investigate, whether these predictions apply to consonantal nuclei as well. In Chapter 4 the effect of phrasal accent on the consonant in different positions within the syllable will be compared. In Chapter 5 a summary of the findings and some concluding remarks will be given. All chapters use data from the same recordings, so in the following section we will present the entire data set as well as the preprocessing procedure. In each chapter, the subset used will be presented briefly.

1.5 General Methods

1.5.1 Participants

Eight native speakers of Slovak participated in the experiment. They had not been living outside Slovakia for more than two years and spoke Slovak regularly with other native speakers of Slovak.

The first speaker served as a pilot subject and her results are not included in the analysis. Based on the pilot recording, some adjustments were made to the recording procedure. Another participant turned out to have a uvular /r/ or sometimes a combination of uvular and tongue tip movement. We noticed it during the acoustic segmentation and a closer look at the ultrasound recordings confirmed it. We decided to exclude this participant from all analyses. Thus, data for six speakers (five females and one male) have been fully analyzed and will be presented.

1.5.2 Stimuli

The full data set will be presented, although parts of it were not analyzed within the scope of the thesis.

The target words consisted of phonotactically valid bisyllabic nonsense words to control for coarticulatory effects as much as possible. Target sequences were designed as to compare the long - short distinction of consonantal and vocalic nuclei and to compare the liquids in nucleus, onset and coda positions. The target nucleus was either /e/ for the vocalic condition or /l/ or /r/ for the consonantal condition. When /l/ or /r/ occupied the onset or the coda position, the nucleus of that syllable was also /e/. The nucleus, liquid + vowel sequence or vowel + liquid sequence were flanked by /p/, a voiceless bilabial plosive to allow for easy acoustic segmentation and minimize coarticulatory effects on the tongue tip as well as the tongue back. This was important as /l/ and /r/ both consist of a tongue tip gesture and a velar retraction gesture. Pouplier and Beňuš (2011) noted that they could not systematically evaluate the behaviour of the velar retraction gesture and consequently the coordination between the tongue tip and tongue retraction gesture, because of the velar consonants in their stimuli. According to Sproat and Fujimura (1993) and Krakow (1999) this coordination differs systematically for different syllable positions.

As already mentioned, Slovak words are always stressed on the first syllable. To analyze the effect of word stress, the target sequence was either in the first syllable (stressed position), or in the second syllable (unstressed position). If the target syllable was the first syllable, the nucleus of the second syllable was occupied by /a/, such as in ***plp****ap*. If the target syllable was the second syllable, the first syllable was occupied by /i/, for example *pi****plp***. The full set of target words is presented in Table 1.5. The target sequence is presented in bold. Regarding syllabification it should be noted that for example in ***pep****ap* the second *p* can be syllabified as the onset of the second syllable rather than the coda of the first syllable. In ***pelp****ap* it is also not clear whether the *lp*-sequence can be viewed as a coda cluster or a heterosyllabic consonant sequence with *l* belonging to the coda of the first syllable and *p* belonging to the onset of the second syllable. In order to elicit the two accentuation patterns, target words were presented in two carrier phrases. The carrier phrases were chosen so that the vowels preceding and following the target word were also /i/ and /a/, respectively, so the target sequence was always in the same vocalic context. Front vowels were chosen so that the tongue retraction movement could be reliably identified to belong to /l/ or /r/:

(1) Pozri, ved' on mi ***plpap*** dal.
‘Look, in the end he gave me ***plpap***.’
(2) Pozri, aj **Ron** mi *plpap* dal.
‘Look, also **Ron** gave me *plpap*.’

In the first carrier phrase (1), the target word (*plpap*) receives the main phrasal accent. In the second carrier phrase (2), *Ron* is accented and the target word is not primarily accented.

Table 1.5: List of target words: The target sequence is presented in bold text. When the target is in the first syllable, it receives word stress, while in second syllable position it is unstressed. In nucleus position, the target can be either short or long. Long nuclei are orthographically marked with an acute accent mark. In addition, /l/ and /r/ can be in onset or coda position as part of a cluster.

	/e/	/l/	/r/
short	**pep**ap pi**pep**	**plp**ap pi**plp**	**prp**ap pi**prp**
long	**pép**ap pi**pép**	**pĺp**ap pi**pĺp**	**pŕp**ap pi**pŕp**
onset		**plep**ap pi**plep**	**prep**ap pi**prep**
coda		**pelp**ap pi**pelp**	**perp**ap pi**perp**

For each speaker, the order of the sentence pairs was randomized individually. For each target word, the sentence in which it was accented was presented first, then followed by the sentence in which it did not carry the primary accent. The word that was supposed to be primarily accented was written in red script when presented to the speakers to read during the recording.

1.5.3 Recording

Setup

Acoustic and articulatory data were recorded simultaneously in a soundproof booth at the Phonetics Institute in Munich.

Stimuli presentation, data collection and synchronization was coordinated using the AAA software (Articulate Instruments Ltd, 2012a). Ultrasound was used to acquire articulatory data. An ultrasound image is created by ultra high frequency waves emanating from the piezoelectric crystals in the probe. When the probe is attached to the chin, the ultrasound beams travel upwards and are reflected at the upper surface of the tongue. Then these ultrasound beams get back to the probe along the scan lines. Based on the time a beam needs to return to the probe the distance of the upper surface from the probe is calculated (Stone, 2005). For our recordings an Ultrasonix SonixTouch scanner and a C9-5 transducer were used. The transducer was fixed under the speaker's chin using the AAA stabilization headset (Articulate Instruments Ltd, 2012b), so that a mid-sagittal image of the tongue was recorded. Scan depth was adjusted to get the optimal picture for each participant, and the scan sector and consequently the number of scan lines were limited to depict the section between the shadows thrown by the jaw and hyoid bone as recommended by Stone (2005). Shallower depth settings and smaller sectors provide faster scan rates so that the frame rate varied across the participants between 60.0 Hz and 85.34 Hz. Individual ultrasound settings are presented in Table 1.6. It is unclear why in the case of Speaker 5 the frame rate was only 60.0 Hz although all other settings are identical to those of Speaker 1. In addition to the ultrasound data, lip and jaw movement was recorded using a video camera which was attached to the stabilization headset. Small beads were used as fixed landmarks by glueing them onto the upper and lower lips and the jaw of the participants. The frame rate of the the video recording was 29.97 Hz. The headset was attached to a pulley string hanging from the ceiling to avoid the entire weight of the headset weighing on the head of the participant. Acoustic data was recorded using a Sennheiser MKH40 microphone at a sampling rate of 22 050 Hz. Stimuli were presented on a computer screen which was visible through a glass window from the booth one sentence at a time for speakers to read them aloud.

Table 1.6: Ultrasound settings for each speaker.

Speaker	Frame rate [Hz]	Number of scan lines	Pixel resolution [mm]	Field of view [rad]	Depth [mm]
1	80.6	108	0.154	124.7	70
2	76.78	102	0.154	117.7	80
3	72.52	108	0.154	124.7	80
4	85.34	102	0.154	117.7	70
5	60.0	108	0.154	124.7	70
6	85.34	102	0.154	117.7	70

Recording procedure

Before the actual recording, we tested whether the participant's tongue was well visible using ultrasound. Then, all participants heard two to four sentence pairs read by a native Slovak speaker who did not participate in the experiment as part of the instruction on accentuation. After recording the first participant, we noticed that she made several mistakes during the first repetition, so all further participants read out loud the entire list of sentences to get acquainted with the stimuli before the recording.

Five to six repetitions were recorded (speakers 1, 2 and 4 five repetitions, speakers 3, 5 and 6 six repetitions), depending on whether the participants seemed tired and whether they were willing to complete one more round after having completed five repetitions. They could take breaks in between to drink water or just take a break but could not take off the headset the ultrasound probe was attached to in order to keep the probe position stable. We did not control for speech rate.

1.5.4 Preprocessing

Acoustic segmentation

Acoustic data were first automatically segmented using the WebMAUS forced alignment system (Kisler, Schiel, & Sloetjes, 2012) and then manually corrected in Praat (Boersma & Weenink, 2014). During manual correction it was verified once more whether there were no misread tokens. It was then checked whether word boundaries for the entire utterance and segment boundaries within the target word were segmented correctly. /p/ was always flanked by voiced segments, so it was defined as the portion with no voice bar and visible formants. Vowels and liquids were defined as segments with a visible voice bar and formants. Voice onset time between /p/ and the following sonorant segment where the voice bar or formants were missing was segmented separately, but for later analysis it was interpreted as belonging to /p/. Setting the segment boundary between vowels and liquids was less straightforward. The vowel usually had more stable and clearer formant values. The segmentation was also verified acoustically. TextGrids that identified the relevant segments (bold parts of the target words in Table 1.5) for articulatory analysis were imported to the AAA software (Articulate Instruments Ltd, 2012a). The TextGrids were also transferred to an EmuR database (Winkelmann, Jaensch, Cassidy, & Harrington, 2016) for acoustic analysis, using the TextGridTools toolkit for python (Buschmeier & Wlodarczak, 2013).

Tongue contour tracking

Tongue contours were tracked using the AAA software (Articulate Instruments Ltd, 2012a). Before tracking the tongue contours, a fan shaped grid with 42 evenly spaced axes was superimposed on the ultrasound images, where the origin of the axes matches the theoretical origin of the ultrasound beams. The fan was adjusted for each participant to match their ultrasound image automatically. Representing the tongue contour on a fan grid using polar coordinates matches the way an ultrasound image is retrieved. For each ultrasound frame within the target sequence the tongue contour was tracked semi-automatically using the tracking methods available within the AAA software and manual correction. The tongue contour was represented as a spline with 42 control points, each marking the point where the tongue surface crossed one of the axes of the superimposed fan grid. Within AAA it is

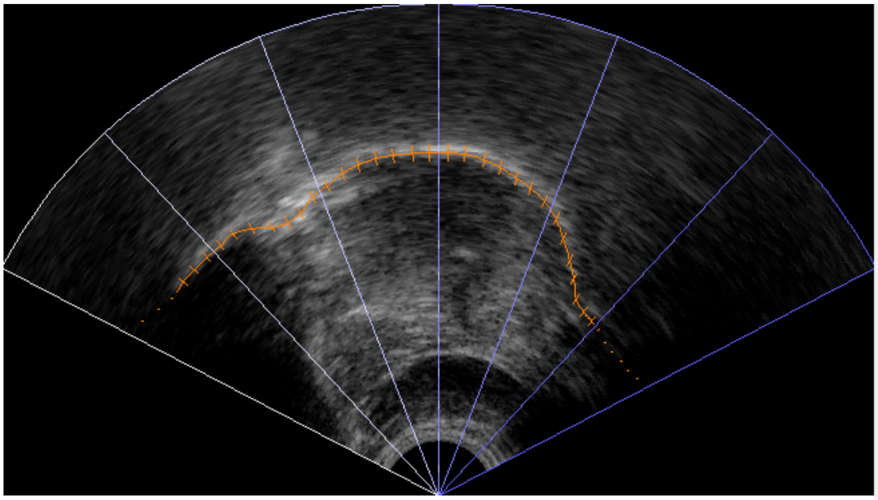


Figure 1.1: Tongue contour tracking within AAA: Tongue tip is to the left. The superimposed fan is indicated in gradient white to blue color. The tongue contour is represented in orange. The tick marks indicate where the fan spline crosses one of the 42 axes of the fan grid. In regions where the tongue contour is not visible (very left and very right regions) the tick marks are only small dots and are not connected with each other. At these points the confidence value is set to zero.

possible to indicate regions where the tongue was not visible, so at those axes an additional value indicating confidence was set to zero (see also Figure 1.1). The tongue contour data was exported and statistical analyses were carried out in R.

Additional recordings were made where subjects swallowed water to be able to trace the palate and the occlusal plane using a bite plate were retrieved. However, these data were not used for our analysis. More detailed measures used for articulatory analysis will be presented in each chapter.

Chapter 2

F0 and acoustic duration as correlates of phrasal accent on vocalic and consonantal nuclei

2.1 Introduction

In this chapter we compare whether nuclear accent is implemented on syllables with consonantal nuclei in the same manner as for syllables with vocalic nuclei by looking at the F0 and acoustic duration of the nucleus.

As laid out in the introduction, there is next to no published work on the prosody of syllabic consonants. In general, tones marking phrasal accent occur on the primary stressed syllable. However, even in the case of vocalic nuclei, it is not unusual that the F0 peak is attracted by a sonorant consonant. For example, it has been shown for German that the F0 peak occurs earlier for syllables with sonorant onsets than voiceless obstruent onsets and later when the coda is a sonorant than when it is a voiceless obstruent (Jilka & Möbius, 2006). In Dutch, the peak occurs on the sonorant following the accented vowel when the vowel is phonologically short, while it occurs on the vowel when it is long (Ladd, Mennen, & Schepman, 2000). In the Australian languages Warlpiri and Mawng, not only does the F0 peak occur on the consonant following the vowel of the accented syllable, but so do other differences associated with prosodic strengthening usually observed on vowels. For example, longer duration and hyperarticulation are found on the consonant rather than on the vowel (Butcher & Harrington, 2003; Fletcher, Stoakes, Loakes, & Singer, 2015). However, in these studies the nucleus was always a vowel, and studies which compare vocalic and consonantal nuclei are less common. In Tashlhiyt Berber, a language known for allowing all types of consonants in nuclear position, the placement of F0 peak is more variable (Grice et al., 2015; Roettger, 2017). As already pointed out in Section 1.1 of the introduction, Grice et al. (2015) showed that in Tashlhiyt Berber, the F0 peak could shift from a syllable occupied by a consonantal nucleus to a syllable with a vocalic nucleus even if the consonantal nucleus was a sonorant. It thus seems that although sonorant consonants can attract F0 peaks, syllables with vocalic nuclei are preferred for F0 peak placement. Barnes, Brugos, Veilleux, and Shattuck-Hufnagel (2014) noted that F0 contours can be affected by segmental context. We mentioned above that F0 peaks are attracted by sonorant consonants. However, an earlier peak in syllables with obstruent coda compared to sonorant coda can be interpreted as a strategy “to ensure optimal expression of F0 contours in a given segmental context” (Barnes et al., 2014, p. 1125). This enables the speaker to realize the entire pitch contour during the sonorous segment of the syllable. Languages can also restrict certain tone patterns to particular segments. Cross-linguistically, syllables with long vowels are preferred over syllables with short vowels and syllables with sonorant consonant codas to syllables with obstruent codas as carriers of contour tones (Zhang, 2004). Generally, F0 contours are acoustically not perceivable during obstruents’ voiceless or closure portions. Instead of filling in or interpolating the missing F0 over these segments, it has been shown that listeners ignore and skip over these segments (Barnes, Brugos, Veilleux, & Shattuck-Hufnagel, 2011). In sonorant consonant sequences, the F0 contour is present. Yet the vocalic nucleus often seems to be preferred for the realization of critical pitch information. To examine how F0 contours are perceived on less sonorous sequences like nasals or voiced fricatives compared to vowels, Barnes et al. (2014) compared the perception of F0 height (pitch) in target words *day*, *Dane* and *Dave* with systematically differing sonority word-finally. In their stimuli the F0 peak was during the final vowel or consonant. Indeed, they showed that equal F0 on less sonorous sequences was perceived as having a lower F0 than on more sonorous sequences. This finding is analogous to observations that at equal F0, more sonorous low vowels like /a/ are perceived to have a higher F0 than less sonorous high vowels like /i/. This phenomenon is also referred to as intrinsic pitch (Fowler & Brown, 1997; Stoll, 1984). For vowels, it has also been observed that high vowels are usually produced with somewhat higher F0 than low vowels, particularly in the speakers' higher F0 range, usually referred to as intrinsic F0 (Fowler & Brown, 1997; Silverman, 1987; Whalen & Levitt, 1995). Because intrinsic F0 can be observed in different languages, regardless of the size of vowel inventory and because it occurs in tone languages as well, it has been proposed that (at least partly) it is related to the supraglottal articulation of vowels (Whalen & Levitt, 1995). The muscle activity during a high vowel moves the hyoid bone forward, by which a force on the thyroid cartilage is applied and the vocal cords are lengthened, resulting in a higher F0 (Honda, 1983). This is also referred to as the tongue pull theory. In addition to the physiological circumstances which result in higher F0 for high vowels, some speakers may also actively enhance this effect (Hoole & Honda, 2011). It is thus possible that speakers phonologize the effect of intrinsic F0 and further emphasize this difference. The intrinsic F0 of sonorant consonants has to my knowledge not been investigated yet.

Assuming that F0 on consonantal nuclei is perceived as being lower than the equal F0 on vocalic nuclei, considering the lower intrinsic pitch observed for nasals and voiced fricatives in coda position (Barnes et al., 2014), the question is whether a higher F0 is produced on consonantal nuclei. Articulatorily, the two syllabic consonants in Slovak (a dark /l/ and /r/) are both sounds that should have a retracted tongue dorsum (e.g. Kavitskaya et al. (2009); Recasens (2013); Sproat and Fujimura (1993)). Thus, at least in the dorsal part the tongue configuration is more similar to low vowels like /a/ which does not predict a higher intrinsic F0. The raised tongue tip for /l/ or /r/ is formed using the intrinsic tongue muscles (mainly the superior longitudinal muscles). A contraction of the genioglossus posterior, which according to the tongue pull theory is responsible for the higher intrinsic F0 in high vowels, is not expected for the tongue configuration for /l/ and /r/. While for vowels intrinsic F0 and intrinsic pitch go in the opposite direction so that it can be said that their effects cancel each other out more or less, this can not be predicted for Slovak /l/ or /r/. One reason why consonantal nuclei are less preferred cross-linguistically could be that F0 is not as salient on consonants as on vowels.

In many languages phrasal accent is accompanied by lengthening as a secondary cue to F0. Beňuš (2011) has previously investigated vocalic and consonantal nuclei of lexically stressed syllables without considering the effect of phrasal accent and shown that the phonemic length contrast is robust, with long nuclei being about twice as long as short nuclei. This suggests that syllabic consonants pattern with vocalic nuclei. We will add to this knowledge by examining the effect of phrasal accent on the acoustic duration of phonemically short and long vocalic and consonantal nuclei[1].

This chapter will be the ground work for the following chapters which investigate articulatory correlates of prosodic emphasis. Because not many languages offer the opportunity to directly compare the effect of phrasal accent on vocalic and consonantal nuclei, we will investigate whether the main correlates, namely F0 and duration, are implemented on consonantal nuclei as they are on vocalic nuclei.

2.2 Methods

Details on recordings and segmental principles of the entire corpus were given in Section 1.5. The target words that will be analysed in this chapter are in Table 2.1 and the target nuclei are printed in bold font. According to Slovak orthography long nuclei are marked with an acute accent mark. In Slovak, stress is fixed on the first syllable, thus in our target words the target nucleus is stressed.

Table 2.1: List of target words used for F0 and duration analyses: The target nucleus is presented in bold font, word stress is on the first syllable in all cases.

	/e/	/l/	/r/
short	p**e**pap	p**l**pap	p**r**pap
long	p**é**pap	p**ĺ**pap	p**ŕ**pap

In order to elicit the two accentuation patterns, the target words were presented in two carrier phrases, here with the accented words in bold:

---

[1] These results on acoustic duration have previously been presented at ICPhS in Glasgow (Bučar Shigemori, Pouplier, & Beňuš, 2015).

(1) Pozri, ved' on mi ***plpap*** dal.
‘Look, in the end he gave me ***plpap***.’
(2) Pozri, aj **Ron** mi *plpap* dal.
‘Look, also **Ron** gave me *plpap*.’

In the first carrier phrase, the target word carries the main phrasal accent. In the second carrier phrase, *Ron* is accented.

Data from six native speakers of Slovak will be analysed. Three speakers repeated each token 5 times while the other three speakers made 6 repetitions. Due to mispronunciation or recording errors, three tokens were eliminated; *pŕpap* in accented condition and *pepap* and *pépap* in unaccented condition. In sum, 393 tokens will be analysed ((3 target nuclei x 2 phonemic lengths x 2 accent conditions x (3 speakers x 5 repetitions + 3 speakers x 6 repetitions)) - 3 tokens). The acoustic data were automatically segmented using the WebMAUS (Kisler et al., 2012) forced alignment system and then manually corrected in Praat (Boersma & Weenink, 2014). For further analysis the data was converted into an EmuR database (Winkelmann et al., 2016). F0 was extracted using wrassp (Bombien, Winkelmann, & Scheffers, 2016) within EmuR. Statistical analysis was carried out in R (R Core Team, 2015). Details are described separately for each measurement.

2.2.1 Measurements for F0

F0 was measured from 30ms before the beginning of the target word (the last 30ms of /i/ in *mi* of the carrier phrase) to the end of /a/ in the second syllable of the target word. The time points were mapped onto a time interval from 0 to 1 resulting in time-normalized curves consisting of 58 to 206 equidistant data points. The data are visualized on the left of Figure 2.1. The F0 curves of the vocalic nucleus /e/ are depicted in the top row and of the consonantal nuclei /l/ and /r/ in the middle and bottom row, respectively. The left column shows the unaccented condition for each nucleus type and the right column the accented condition. Long and short nuclei are plotted together. Subjects are colour coded. During the voiceless segments of /p/ in the target word F0 is zero. The only male subject stands out, having an overall lower F0. Normalization for speaker was not carried out, because the statistics were expected to account for speaker-specific differences in the form of random effects. Correction for octave jumps and smoothing has not been carried out either. Overall, the pattern seems consistent across subjects. In accented condition, the section between the voiceless segments (from around time point 0.25 to 0.75,) which is the target nucleus, has the highest F0 for all nucleus types.

For statistical analysis, linear interpolation was applied to the voiceless segments where F0 was zero. The interpolated data are visualized on the right of Figure 2.1. The interpolated data can be better approximated by splines than treating the voiceless segments as missing data which yield more meaningful statistical results using functional linear mixed modelling.

Functional linear mixed modelling (FLMM)

Functional linear mixed modelling (FLMM) is a method analogous to regular mixed models, capable of evaluating curve dynamics (Cederbaum, 2017; Cederbaum, Pouplier, Hoole, & Greven, 2016; Pouplier, Cederbaum, Hoole, Marin, & Greven, 2017). Because FLMM is not wiedely known, an account of the most relevant parts from Cederbaum et al. (2016) and Pouplier et al. (2017) will be given. Instead of a single measure for each effect as is the case in mixed models, in FLMM effects are represented as curves. This enables an analysis of two-dimensional data and allows us to capture the dynamics of, for example in this chapter, F0 as a function of time. Covariates are analogous to fixed effects in mixed effects modelling. However, they can consist of only two levels, which are dummy coded as 0 or 1[2]. FLMM estimates a reference mean curve, covariate effect curves and interaction effect curves, each with their confidence band. The reference mean curve represents the condition where all dummy codings are set to 0. The covariate and interaction effect curves represent the effect of each of them on the reference mean. An effect is significant when confidence bands differ from zero. This can be the case for only a section and not necessarily for all time points. In addition, the model calculates functional random effects for Speaker, Item and a smooth error term which captures speaker-by-item interaction and a repetition-specific deviation. The general form of the model is given in (2.1):

$$Y_{ijh}(t) = \mu(t, x_{ijh}) + B_i(t) + C_j(t) + E_{ijh}(t) + \varepsilon_{ijh}(t) \tag{2.1}$$

---

[2] Covariates in FLMM can handle continuous variables as well, but in case of categorical data they can only have two levels.

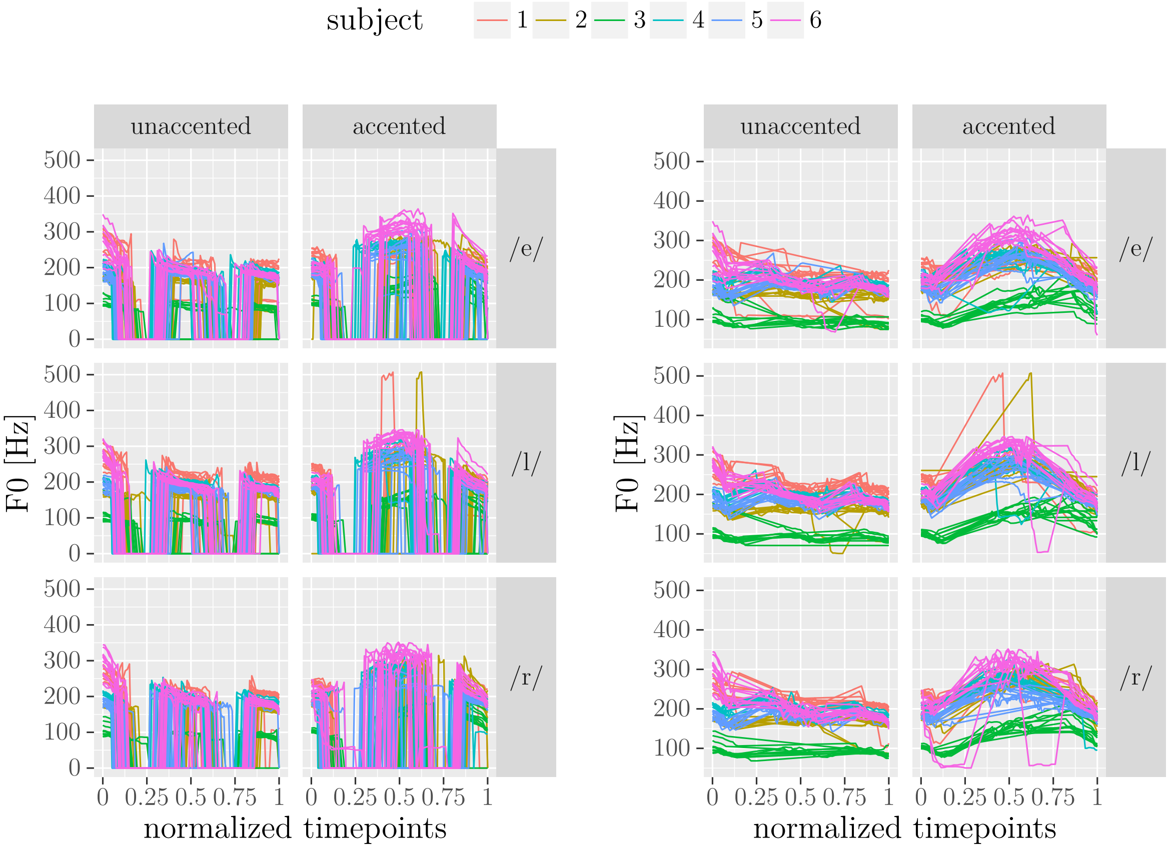


Figure 2.1: On the left F0 as extracted from the data over normalized time, before interpolation, on the right after interpolation: from top to bottom by nucleus type (vocalic /e/, consonantal /l/, consonantal /r/), in unaccented (left) and accented (right) condition, colour coded by speaker.

$Y_{ijh}(t)$ is the index curve for speaker i, item j, and repetition h observed at time $t \in T = [0, 1]$. $\mu(t, x_{ijh})$ is a curve specific smooth mean function, where x_{ijh} are known covariates and interactions of covariates. $B_i(t)$ and $C_j(t)$ are random functional intercepts for Speaker and Item, respectively, and $E_{ijh}(t)$ is a Speaker-, Item-, and Repetition-specific smooth random deviation which also includes the interaction between Speaker and Item. $\varepsilon_{ijh}(t)$ is white noise measurement error. An example of the mean function $\mu(t, x_{ijh})$ with two covariates and their interaction is given in (2.2).

$$\mu(t, x_{ijh}) = f_0(t) + f_1(t) \cdot covariate1_j + f_2(t) \cdot covariate2_j + f_3(t) \cdot covariate1_j \cdot covariate2_j \quad (2.2)$$

$f_0(t)$ is the reference mean curve, $f_1(t)$ and $f_2(t)$ the covariate effect curves and $f_3(t)$ the interaction effect curve. $covariate1_j$ and $covariate2_j$ are either 0 or 1, depending on how the item j was coded. So the mean curve of a specific condition is calculated by adding the covariate effect curves of the effect in question to the reference mean curve.

What needs to be kept in mind is that time is normalized. For each curve, time points are mapped to a sequence from 0 to 1. FLMM allows individual curves to differ in number of time points so resampling is not necessary.

2.2.2 Measurements for acoustic duration

For the stimuli at hand, in all target sequences the nucleus was flanked by voiceless plosives, which made its acoustic segmentation straightforward: we defined the nucleus as the part in which the voice bar and formant structure were clearly visible. For details see Section 1.5.4. For the statistical analysis mixed effect models were used to test the effect of nucleus type, phonemic quantity and phrasal accent on the acoustic duration of the nucleus. While the fixed effects PHRASAL ACCENT and PHONEMIC QUANTITY had two levels, analogously to the dummy codings of the covariates in FLMM for F0 analysis, NUCLEUS TYPE had three levels: /e/, /l/ and /r/.

2.3 Results

2.3.1 F0

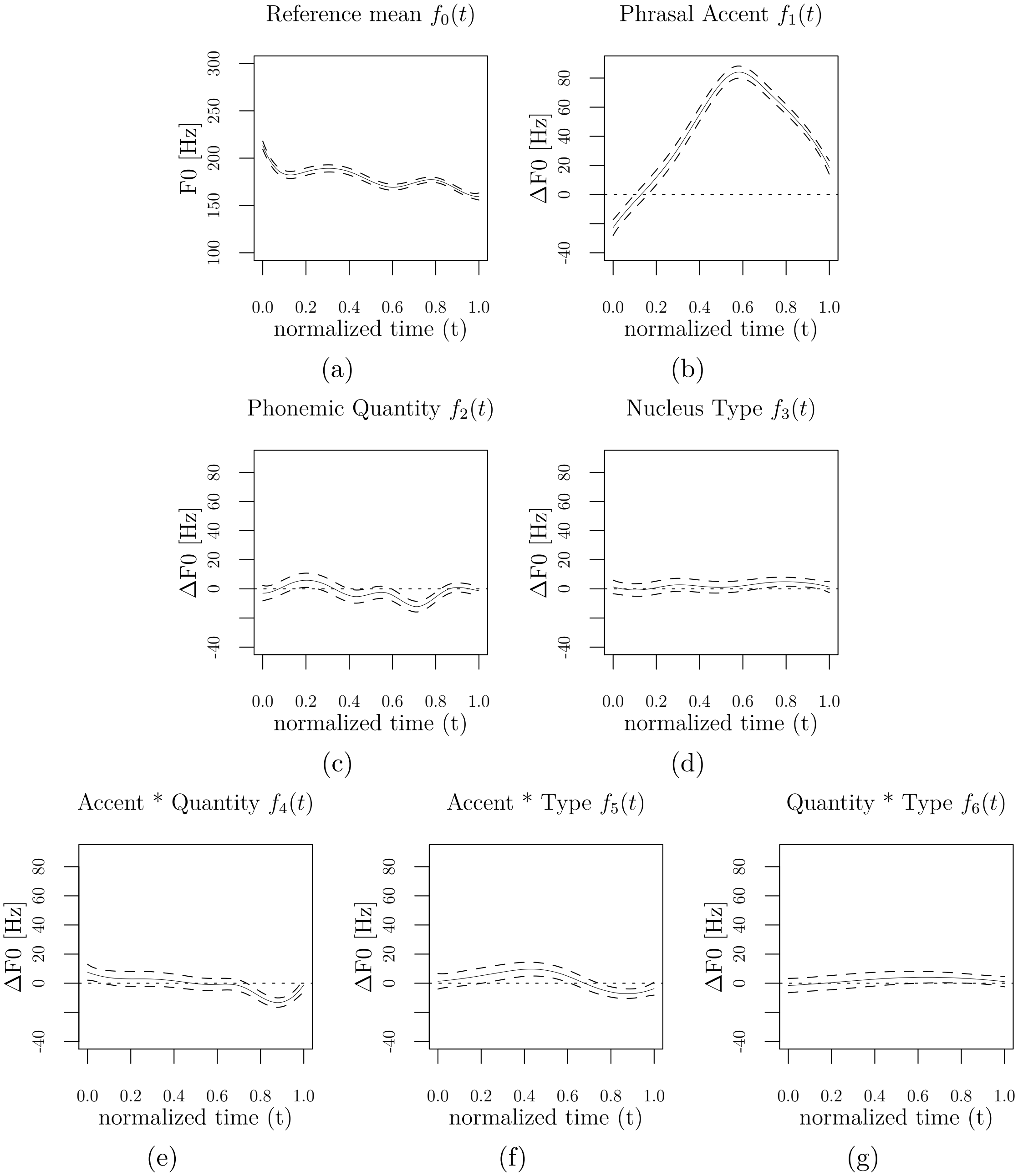
To test the effect of phrasal accent, phonemic quantity and nucleus type on the F0 curve functional linear mixed modelling (sparseFLMM package in R (Cederbaum, 2017)) was used. The model consisted of three covariates: PHRASAL ACCENT (0 = unaccented; 1 = accented), PHONEMIC QUANTITY (0 = short; 1 = long) and NUCLEUS TYPE (0 = vowel; 1 = consonant). The reference mean corresponds to the mean F0 curve of the unaccented condition in which the first syllable of the target word was occupied by a short vocalic nucleus. For each effect a pair wise interaction effect curve was estimated as well, resulting in six effects. The overview of all covariates for the current analysis is given in Table 2.2.

Table 2.2: Overview of covariates for F0 analysis.

$f_0(t)$	Reference mean	unaccented short vocalic nucleus
$f_1(t)$	Phrasal accent	0 = unaccented, 1 = accented
$f_2(t)$	Phonemic quantity	0 = short, 1 = long
$f_3(t)$	Nucleus type	0 = vocalic, 1 = consonantal
$f_4(t)$	Phrasal accent * Phonemic quantity	
$f_5(t)$	Phrasal accent * Nucleus type	
$f_6(t)$	Phonemic quantity * Nucleus type	

The reference mean, corresponding to the F0 curve for the short vocalic nucleus (target word *pepap*) in unaccented condition, is shown in Figure 2.2 (a). It is rather flat and corresponds to the mean of the left column of the interpolated curves in Figure 2.1. Plots (b) to (d) in Figure 2.2 show the covariate effects while plots (e) to (g) show the interaction effects. In each of these plots the solid lines represent the mean effect and the dashed lines represent the confidence bands. An effect is significant when the covariate mean with its confidence bands differs from zero.

The largest effect can be observed for PHRASAL ACCENT (Figure 2.2 (b)). The covariate effect curve is significantly above zero from time point 0.2 until the end of the target se-



(a) (b) (c) (d) (e) (f) (g)

Figure 2.2: Reference mean and covariate effect curves of F0 measured from 30ms before the beginning of the target word to the end of /a/ in the second syllable of the target word. The reference mean corresponds to the F0 contour of the unaccented target word occupied by a short vocalic nucleus.

quence with the peak around time point 0.6, while it is significantly below zero at the very beginning. Figure 2.3 shows the summed estimated effect curve for PHRASAL ACCENT. The red dashed line is the reference mean, corresponding to the unaccented short vocalic condition. The red solid line corresponds to the accented short vocalic condition. It is obtained by adding the covariate effect curve for PHRASAL ACCENT $f_1(t)$ (Figure 2.2 (b)) to the reference mean $f_0(t)$ (Figure 2.2 (a)). While the red solid line corresponding to the accented condition starts lower than the red dashed line for the unaccented condition, it keeps rising and reaches the peak around time point 0.5 where it starts to decline slowly. However, it does not get as low or lower than for the unaccented condition again. The F0 contour was measured from the last 30ms of the vowel preceding the target sequence until the end of the vowel occupying the second syllable of the target word. This means that the nucleus of the second (unstressed) syllable also has a significantly higher F0 in accented condition than in unaccented condition. However the main difference for the two accent conditions is observed during the first (stressed) syllable nucleus. The segment at the very beginning during which F0 is lower in the accented condition than in the unaccented condition belongs to the vowel preceding the target word.

The covariate effect curve $f_2(t)$ for PHONEMIC QUANTITY (Figure 2.2 (c)) has a significant lowering effect at around time point 0.7. We considered an effect to be at trend level when the edge of the confidence band overlapped with zero, as is the case around time point 0.2 and between time points 0.4 and 0.5. The result of the effect is visualized as the summed effect curve corresponding to the long unaccented vocalic nucleus condition in Figure 2.4 using the blue dashed line. Compared to the red dashed line which corresponds to the short unaccented vocalic nucleus it is higher between around time points 0.1 and 0.35. At around time point 0.35 the two lines cross each other and the blue dashed line is lower until around time point 0.85. However, the differences are significant only around time point 0.7. For both dashed lines three falling slopes can be observed. Comparing the summed effect curves curves with the raw data in Figure 2.1 we can assume that for both unaccented conditions the peaks roughly correspond to the beginning of the voiced segments. The first decline begins at time point 0 for both conditions and ends at around time point 0.1 for the long unaccented vocalic condition (blue dashed line) and slightly later for the short unaccented vocalic condition (red dashed line). This corresponds to the last 30ms of the word *mi* preceding the target word.

The second decline which presumably corresponds to the target nucleus lasts from around time point 0.3 to time point 0.6 for the short unaccented vocalic condition and from time point 0.2 to time point 0.7 for the long unaccented vocalic condition. The last decline begins slightly before time point 0.8 for the short unaccented vocalic condition and slightly later for the long unaccented vocalic condition and corresponds to the vowel /a/ of the target words. The effect of Phonemic quantity thus results in a longer declining slope for the phonemically long target nucleus which as a result reaches a lower F0 value.

The interaction of phrasal accent and phonemic quantity $f_4(t)$ (Figure 2.2 (e)) has a significant effect at the very end of the measured interval, after time point 0.8. The summed effect curve corresponding to the long accented vocalic nucleus is visualized using a blue solid line in Figure 2.4. It is obtained by adding the covariate effect curves $f_1(t)$ and $f_2(t)$ for phrasal accent and phonemic quantity, respectively, and the interaction effect curve $f_4(t)$ to the reference mean $f_0(t)$. After reaching the high peak in accented condition, F0 for long accented nuclei declines more, probably due to its longer duration, reaching a lower F0 value at the end of the target nucleus. After time point 0.8 the difference between the two blue lines is significantly smaller than between the two red lines according to the interaction effect curve $f_4(t)$. This means that difference in F0 on the unstressed vowel /a/ between the two accent conditions is smaller when it follows a long vowel than when it follows a short vowel.

What we are most interested in is the effect of nucleus type, or whether vocalic and consonantal nuclei pattern together. The covariate $f_3(t)$ for Nucleus type has a lower boundary just slightly above zero at around time point 0.8 (Figure 2.2 (d)). This corresponds roughly to the beginning of the vowel in the second syllable of the target word. However, we observe a significant interaction effect for phrasal accent and nucleus type (Figure 2.2 (f)). $f_5(t)$ has a significant raising effect between time points 0.3 and 0.6. Consequently, accented consonantal nuclei have a higher F0 than accented vocalic nuclei. The summed effect curve for accented consonantal nuclei are obtained by adding $f_1(t)$ (covariate effect curve for phrasal accent), $f_3(t)$ (covariate effect curve for nucleus type) and $f_5(t)$ (interaction effect curve for nucleus type and phrasal accent) to the reference mean curve $f_0(t)$. The summed estimated effect curves for phrasal accent and nucleus type are shown in Figure 2.5. The dashed lines represent the short unaccented vocalic and consonantal conditions in red and blue, respectively. They are mostly overlapping, but after time point 0.6 the blue dashed line is slightly higher than the red dashed line. The solid lines represent the accented condition. The difference between the accented (solid line) and unaccented (dashed line) condition is bigger for consonantal nuclei represented in blue than for vocalic nuclei in red. The raising slope for consonantal nuclei in accented condition (solid blue line) is steeper than the slope for vocalic nuclei (solid red line), so that the peak is reached at roughly the same time, with the peak for consonantal nuclei being higher than for vocalic nuclei. The solid blue line reaches the peak slightly earlier than the solid red line. However, it is not possible to make reliable assumptions about peak alignment since time is normalized and the onset consonant is voiceless. In addition, $f_5(t)$ has a significant lowering effect after around time point 0.7. F0 during the second nucleus of the target word (vowel /a/) is not higher when following the consonantal nuclei compared to when following the vocalic nuclei in accented condition, in contrast to the unaccented condition. At the beginning and end of the measured section, corresponding to the preceding and following vowels, the two solid lines are overlapping.

The interaction between PHONEMIC QUANTITY and NUCLEUS TYPE was not significant (Figure 2.2 (g)).

In sum, a robust implementation of a raised F0 contour in accented syllables was observed for vocalic and consonantal nuclei. In accented condition F0 was higher also during the nucleus of the second syllable, but the main difference and the peak were during the nucleus of the first syllable. In the accented condition F0 was even higher when the nucleus was occupied by a consonant. This difference affected the syllable with the consonantal nucleus only and did not affect the nucleus of the preceding or following syllable.

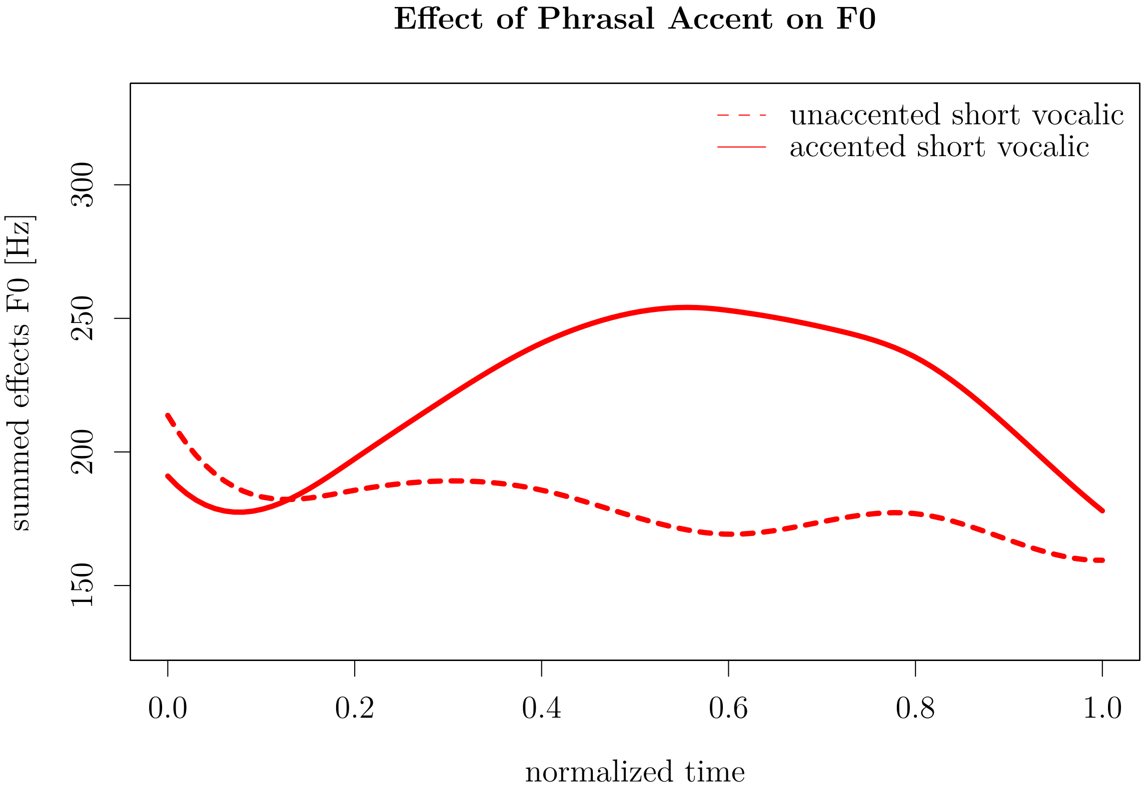


Figure 2.3: Summed effect curves showing the effect of PHRASAL ACCENT on F0. The red dashed line represents the reference mean $f_0(t)$ corresponding to the unaccented short vocalic nucleus condition. The red solid line represents the effect of PHRASAL ACCENT $f_1(t)$ added to the reference mean and corresponds to the accented short vocalic nucleus condition.

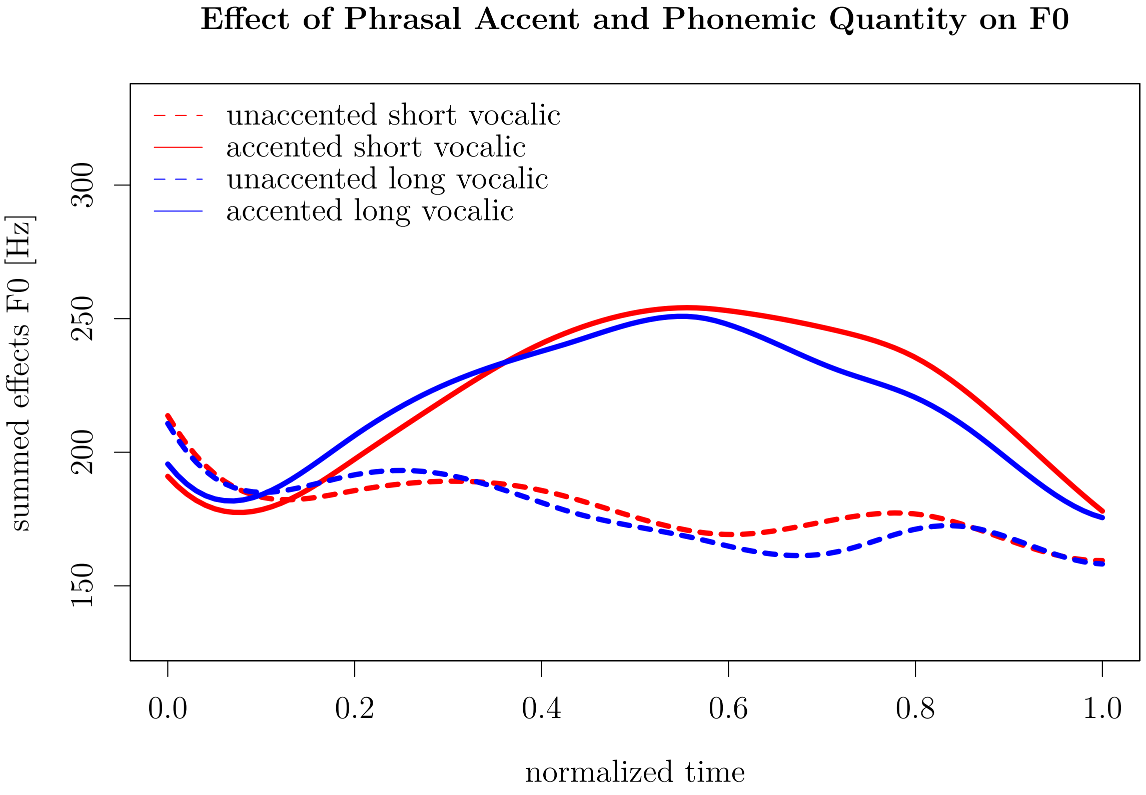


Figure 2.4: Summed effect curves showing the effect of PHRASAL ACCENT and PHONEMIC QUANTITY on F0. The red dashed line represents the reference mean $f_0(t)$ corresponding to the unaccented short vocalic nucleus. The red solid line represents the effect of PHRASAL ACCENT $f_1(t)$ added to the reference mean and corresponds to the accented short vocalic nucleus. The blue dashed line corresponds to the unaccented long vocalic nucleus and is obtained by adding the effect of PHONEMIC QUANTITY $f_2(t)$ to the reference mean $f_0(t)$. The blue solid line corresponds to the accented long vocalic nucleus and is obtained by adding the effects of PHRASAL ACCENT $f_1(t)$, PHONEMIC QUANTITY $f_2(t)$ and their interaction effect $f_4(t)$ to the reference mean $f_0(t)$.

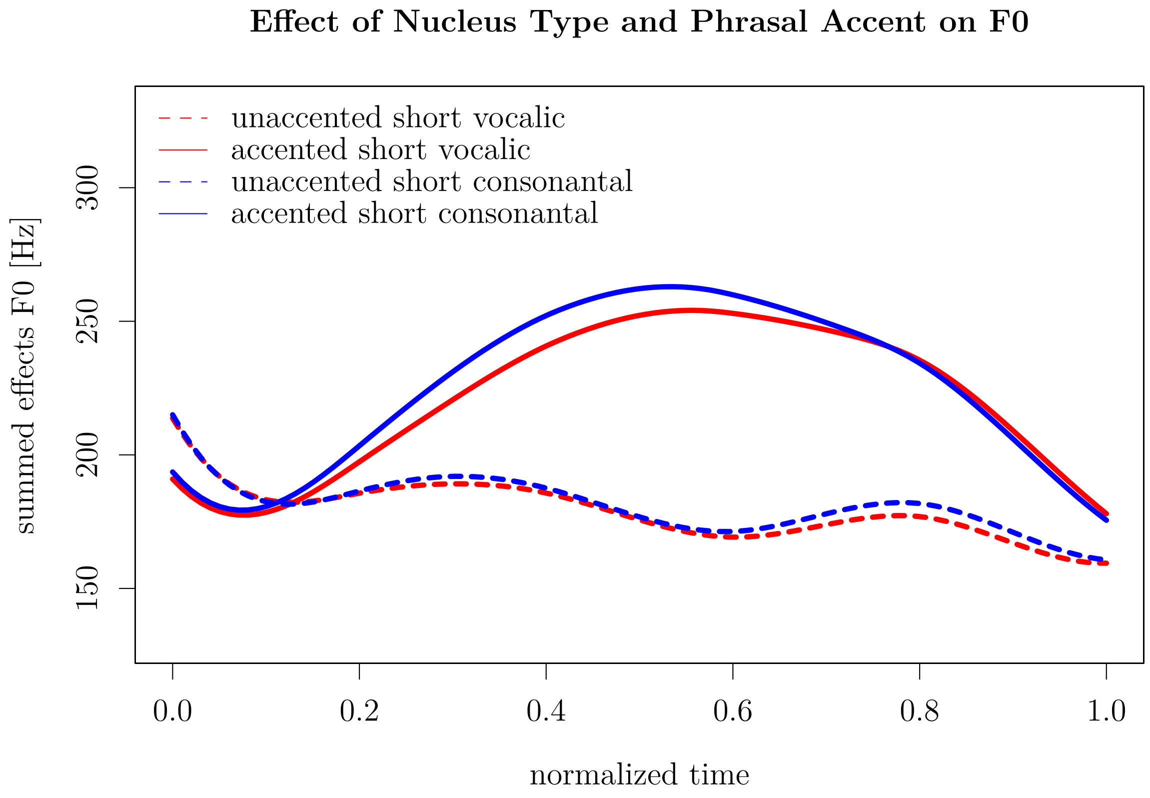


Figure 2.5: Summed effect curves showing the effect of NUCLEUS TYPE and PHRASAL ACCENT on F0. The red dashed line represents the reference mean $f_0(t)$ corresponding to the unaccented short vocalic nucleus. The red solid line corresponds to the accented short vocalic nucleus $(f_0(t) + f_1(t))$. The blue dashed line corresponds to the unaccented short consonantal nucleus $(f_0(t) + f_3(t))$ and the blue solid line to the accented short consonantal nucleus $(f_0(t) + f_1(t) + f_3(t) + f_5(t))$.

2.3.2 Acoustic duration

To test the effect of nucleus type, phonemic quantity and phrasal accent on the acoustic duration of the nucleus, linear mixed effects analysis was performed using the lme4-package (Bates, Maechler, Bolker, & Walker, 2014) in R. PHRASAL ACCENT (two levels: accented, unaccented), PHONEMIC QUANTITY (two levels: short, long) and NUCLEUS TYPE (three levels: /e/, /l/, /r/) were fixed factors, with REPETITION and SPEAKER (with by-SPEAKER intercept and slope for PHONEMIC QUANTITY and NUCLEUS TYPE) as random factors. P-values were then obtained by a likelihood-ratio-test in which the final model was compared to a model with or without the effect in question. The final model consisted of PHRASAL ACCENT (two levels: accented, unaccented), PHONEMIC QUANTITY (two levels: short, long) and their interaction as fixed effects and REPETITION and SPEAKER (with by-SPEAKER intercept and slope for PHONEMIC QUANTITY and NUCLEUS TYPE) as random factors.

Accent is usually accompanied by durational lengthening. Our main interest is whether the effect of phrasal accent on consonantal nuclei is the same as on vocalic nuclei. Figure 2.6 illustrates the nucleus durations for each nucleus type (/e/, /l/ or /r/) and phonemic quantity (long nuclei in light grey, short nuclei in dark grey), for accented and unaccented words. The phonemic quantity contrast of long nuclei being about twice as long as short nuclei as reported by Beňuš (2011) in accented condition could be replicated. In the unaccented condition, the phonemic quantity contrast was also robust, although slightly smaller.

There was no significant effect of NUCLEUS TYPE ($\chi^2[2] = 0.34, p > 0.1$). The effect of PHONEMIC QUANTITY ($\chi^2[2] = 38.80, p < 0.01$) was significant with long nuclei being about $65.3ms \pm 10.8$ (standard errors) longer than short nuclei. PHRASAL ACCENT was significant as well ($\chi^2[2] = 39.2, p < 0.001$). On average, if the target word was accented, the nucleus was $20.5ms \pm 2.8$ (standard errors) longer. The interaction between PHRASAL ACCENT and PHONEMIC QUANTITY was also significant ($\chi^2[1] = 30.27, p < 0.001$) and the post-hoc Tukey test revealed that PHRASAL ACCENT had no significant effect on short nuclei. Accented phonemically short nuclei did not differ significantly from unaccented short nuclei, while phonemically long nuclei were significantly longer in duration when accented compared to when unaccented.

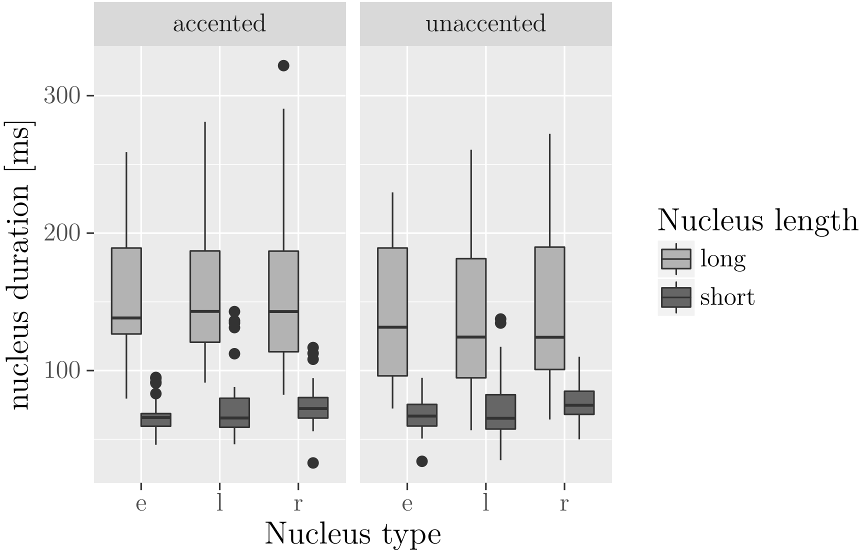


Figure 2.6: Effect of phrasal accent on the acoustic duration of different nucleus types. Acoustic duration is shown separately for each nucleus type (/e/ - vocalic and /l/, /r/ - consonantal), grouped by phonemic length. Left panel shows durations for the accented context and the right panel for the unaccented context.

2.4 Discussion

The aim of this chapter was to investigate whether phrasal accent is implemented on consonantal nuclei in the same manner as on vocalic nuclei in terms of F0 and acoustic duration.

Based on the standard literature on Slovak (Kráľ, 2005), F0 was expected to be a major correlate of phrasal accent. This was confirmed in our data. For our analysis F0 was measured from the last 30 ms of the vowel preceding the target word to the end of the vowel occupying the second syllable of the target word. Word stress in Slovak being fixed on the first syllable, the first nucleus of the target word was expected to be affected most by phrasal accent. As expected, in accented conditions a higher F0 could be observed during almost the entire target sequence with the greatest difference for the two accent conditions during the segment corresponding to the first nucleus of the target word, regardless of whether it was occupied by a consonant or a vowel. Only during the very beginning of the target sequence, during the segment corresponding to the vowel /i/ preceding the target word, F0 was significantly higher for the unaccented condition. The reason might be that the vowel preceding the target word is also the syllable following the primary accented word *Ron* in this condition. Rusko et al. (2007) eliminated the label *L+H** from their set of the phonological intonation annotation scheme for Slovak. We thus assume that F0 is not intentionally lowered before the high F0 in the accented condition, but rather it is higher in the unaccented condition due to the accent being on the preceding word.

In addition to the effect of phrasal accent, several smaller, yet significant effects were observed. Our main interest was the interaction effect between phrasal accent and nucleus type. The F0 difference between the accented and unaccented conditions was greater on consonantal nuclei than on vocalic nuclei. The difference in F0 could be observed for the accented condition. Thus, while no difference in F0 was observed for the unaccented condition, in accented conditions F0 was higher if the target nucleus was consonantal. As mentioned in the introduction, F0 has been shown to be perceived as less prominent on segments lower in the sonority scale (Barnes et al., 2014). It is possible that speakers actively compensate for a perceptually lower F0 on less sonorant sequences by producing a higher F0. The assumption that the higher F0 for consonantal nuclei is controlled for actively is especially likely if we assume that the tongue during the consonantal nucleus is retracted, which is expected for

Slovak /r/ or /l/. Without a perception study, however, it is not possible to tell whether the perceptual differences are fully compensated for in production, so that consonantal nuclei are indeed perceived equally prominent compared to vocalic nuclei. Still, we could show that the accented-unaccented contrast of F0 was robustly produced regardless of whether the nucleus was a vowel or a consonant. Interestingly, nucleus type also affected F0 of the vowel occupying the second syllable within the target word, which was the same in all conditions, namely /a/. In the unaccented condition, when the target nucleus was occupied by a consonant, F0 on /a/ was slightly higher than when the target nucleus was a vowel. The interaction between phrasal accent and nucleus type had an effect in the opposite direction. Thus, while F0 on /a/ was generally higher as an effect of phrasal accent, the difference was smaller for the two accent conditions when following a consonantal nucleus than when following a vocalic nucleus.

Phonemic quantity as well as the interaction of phonemic quantity and phrasal accent also had a significant effect. For phonemically long nuclei, a lower F0 was observed towards the end of the target nucleus and in the accented condition also during the following vowel. The object of our investigation was not F0 movement, which is why we chose voiceless instead of voiced plosives to allow easier acoustic segmentation. Therefore we could not conduct analysis on F0 alignment. A possible explanation for the lower F0 towards the end of the target nucleus and the following vowel is the execution of the phrase final fall.

The acoustic durations of vocalic and consonantal nuclei did not differ from each other. The phonemic length contrast was robust in both accented and unaccented conditions. An effect of phrasal accent was observed in phonemically long nuclei only. When accented, phonemically long nuclei were lengthened additionally, creating an even greater contrast to short nuclei. It can be assumed that short nuclei are not further compressible which leaves them with less room for variation. Similar effects have been also observed for tense and lax vowels in German (Siddins, Harrington, Reubold, & Kleber, 2014), where tense vowels, associated with longer duration, showed greater differences over changes in different stress positions than lax vowels. This asymmetry has been observed in other prosodic aspects in Slovak as well. While the phonemic quantity contrast was always robust, in fast speech or unstressed position compared to normal speech rate and stressed position, long nuclei were shortened more than short nuclei (Beňuš & Mády, 2010; Beňuš, 2011).

Although phrasal accent affected only the duration of phonemically long nuclei and not the short nuclei, the high F0 signalling phrasal accent is reached on the accented syllable, regardless of its acoustic duration.

To conclude, in terms of F0 and acoustic duration, phrasal accent can be realised on syllabic consonants just as for vowels. Although a significant difference of phrasal accent on F0 was observed for the two nucleus types, the effect was not in the direction that would suggest that consonantal nuclei are less capable of carrying prosodic information.

In the next chapter we will investigate the articulatory effect of phrasal accent on vocalic and consonantal nuclei. In particular, the effect of phrasal accent is associated with specific modifications of the articulators involved (Beckman et al., 1992; De Jong, 1995). As will be presented in the next chapter, these hypotheses are made based on the assumption that the nucleus is occupied by a vowel. However, given that these hypotheses also describe the functional requirements of the nucleus, we want to examine if and how they also apply to consonantal nuclei.

Chapter 3

Effect of phrasal accent on acoustics and articulation of vocalic and consonantal nuclei

3.1 Introduction

In the previous chapter we showed that phrasal accent was realized consistently with regard to F0 and acoustic duration on syllables with vocalic and consonantal nuclei. In this chapter, articulatory and formant data will be analysed in order to study the effect of prosodic emphasis on /e/ and /l/ in nucleus position. Data for target words with /r/ will not be analysed, because in many cases it was difficult to identify the apical tap or trill in the ultrasound recordings. In this chapter and the next, syllabic consonant will refer to syllabic /l/.

There are two main theories explaining prosodic emphasis: the Sonority Enhancement Hypothesis (Beckman et al., 1992) and the Hyperarticulation Hypothesis (De Jong, 1995). More precisely, the Sonority Enhancement Hypothesis predicts that the nucleus becomes more sonorous by lowering the jaw and opening the vocal tract to a greater degree, while the onset becomes less sonorous by tightening the constriction. The Hyperarticulation Hypothesis predicts an enhancement of distinctive features of the phonemes to maximize lexical distinction. In other words, assuming that the nucleus is occupied by a vowel and the onset and coda are occupied preferably by stops, the Sonority Enhancement Hypothesis makes predictions based on positions of the segment within the syllable while the Hyperarticulation Hypothesis makes predictions based on phonological properties of its segments. Yet in this constellation, it is not possible to disentangle the positional and segmental properties. The only difference between the two hypotheses is that the Hyperarticulation Hypothesis makes additional predictions about features such as backness or roundness, which are also enhanced under accentuation. The two hypotheses go hand in hand when the nucleus is occupied by a low vowel, for example /a/. A lower tongue body in accented context can be interpreted as both oral cavity expansion, as predicted by the Sonority Enhancement Hypothesis, as well as enhancement of the place feature ([+low]), as predicted by the Hyperarticulation Hypothesis. However, a nucleus occupied by non-low front vowels like /i/ presents a case with conflicting predictions. The Hyperarticulation Hypothesis predicts a higher position of the tongue in accented position based on its place features, [+front] and [+high], which would result in a narrower constriction in the palatal region, usually also with a higher jaw than for low vowels. The Sonority Enhancement Hypothesis predicts lowering of the jaw, which requires a movement in the opposite direction to the tongue. Harrington et al. (2000), Erickson (2002) and Cho (2005) have each pointed out this conflict and investigated acoustically and articulatorily the effect of prosodic emphasis on high and low vowels in connection with the two hypotheses. Acoustically, higher sonority has been reported to correlate with higher F1 (Keating, 1983; Parker, 2002), while hyperarticulated vowels occupy a more peripheral position in the F1-F2 vowel space. For /i/, a more peripheral position corresponds to a lower F1 and a higher F2. While all studies reported higher F1 values and lower jaw position for accented low vowels, results for /i/ were inconsistent. The production of accented /i/ was usually accompanied by lower jaw position, higher F2 and fronted tongue position (measured in the dorsal region in studies by Harrington et al. (2000) and Erickson (2002) and in dorsal and mid region by Cho (2005), all by means of EMA), while inconsistencies were observed in F1 and the vertical dimension of the tongue position. In the study by Harrington et al. (2000) one speaker had a higher F1 and lower tongue dorsum in accented condition. The other speaker showed no effect of accent on F1 but had a higher tongue position for /i/ in bilabial and dental context but not in velar context. Erickson (2002) recorded high vowel /i/ for three speakers. F1 was lower for all speakers but no significant effect on the vertical tongue position was observed. Cho (2005) investigated the effect of accent on /i/ separately for the preboundary and postboundary positions. None of the four speakers showed significant effects for the vertical position of the dorsal region. In preboundary condition, a lower midsection of the tongue in accented condition was observed for two speakers, and for one of the speakers a higher F1 was observed. For the other two speakers a lower F1 was observed in the accented condition, but no significant articulatory effects in the vertical dimension. In postboundary accented condition one speaker produced /i/ with a higher midsection tongue position and lower F1. The other three speakers had a higher F1 but no significant articulatory effects in vertical dimension. The opposing requirements for hyperarticulation and sonority enhancement for high vowels might result in speakers using different strategies to implement prosodic emphasis. Cho (2005) interprets the results for high vowels as evidence that in accented contexts the vowel's sonority is enhanced by opening the jaw and lips more and place features are enhanced only in the direction that does not contradict sonority enhancement. So overall, the consistently lower jaw when accented for all vowels in particular, but also inconsistencies in the vertical articulatory dimension and F1 for high vowels, suggest that the main effect of prosodic emphasis is sonority enhancement. This might be one explanation why universally, as noted by Kenstowicz (2004, p. 191) , "lower vowels are more optimal stress-bearing units than higher vowels". Low vowels, for which hyperarticulation also fulfils the requirements for sonority enhancement, are consequently the ideal candidates for prosodic strengthening compared to high vowels. Consonants, even sonorant nasals or liquids, would be less optimal because they require a constriction in the oral cavity.

Indeed, the cross-linguistic distribution of syllabic consonants supports this idea. While syllabic consonants are quite common among languages, in most cases, their distribution is restricted. Often they occupy only lexically unstressed positions, which are not the primary carriers of prosodic emphasis (Bell, 1978). For example, in English or German, they are post-lexical and a result of vowel reduction in unaccented contexts. Consonants are characterized by an obstruction in the oral cavity. Although sonorants can signal F0 variation, the obstruction in the oral cavity would become more narrow when hyperarticulated, which opposes the requirements for sonority enhancement. The difficulty in meeting the requirements of the Sonority Enhancement Hypothesis and the Hyperarticulation Hypothesis simultaneously could be one of the reasons why syllabic consonants are more common in unstressed positions.

In this regard, Slovak is a special language which provides the opportunity to analyse the implementation of prosodic strengthening of syllables with consonantal nuclei. As already mentioned, in Slovak the consonants /l/ and /r/ can occupy the nucleus of a lexically stressed syllable. In this study, we will investigate the effect of phrasal accent on /l/ when it occupies the syllable nucleus and compare it with the effect on a vocalic nucleus occupied by /e/. /l/ in Slovak is dark in all syllable positions, a segment which is often described as consisting of a "consonantal" apical (or laminal) gesture and a "vocalic" dorsal retraction gesture (Sproat & Fujimura, 1993). It has been suggested that the presence of the dorsal gesture (in addition to the apical gesture) with its vowel-like tongue body control allows /l/ to function as a syllable nucleus (Pouplier & Beňuš, 2011). However, similar to the case of /i/, it is difficult to predict how phrasal accent should be implemented in nuclei occupied by /l/ when trying to satisfy the Sonority Enhancement Hypothesis and the Hyperarticulation Hypothesis.

Originally, sonority has been used to rank segments in order to explain the cross-linguistic preference for certain syllable structures. Clements (1990) suggests that the appropriate way to see the sonority hierarchy is of it to not be "a single, multivalued feature or parameter, but rather that the hierarchical scale is built up out of independently-needed linguistic categories linked by implicational relations" (p.18). Instead, according to Clements (1990), relative sonority is defined based on the binary features [±syllabic], [±vocoid], [±approximant] and [±sonorant]. The more + specifications apply to a segment, the more sonorous it is. While this approach suggests that /l/ would become more sonorous when occupying the syllable nucleus than onset or coda, it is not clear whether its realization in terms of acoustics or articulation is affected. Although there have been attempts to find a phonetic basis for sonority (see Parker (2002, p. 44-48) for an overview of proposed correlates), one single parameter usually fails to explain all phonotactic preferences across languages. A very popular parameter is openness of the vocal tract. Strictly speaking, low vowels, which are considered to be most sonorous, in particular low back vowels, can also have a strong narrowing in the dorsal region, similar to close front vowels in the palatal region, which are usually ranked as less sonorous. In contrast, the Sonority Enhancement Hypothesis predicts primarily a lower jaw position in more sonorous contexts which is probably more likely to co-occur with a low back vowel than with a high front vowel. It thus seems that tongue retraction in the dorsal region, as long as it does not result in a fricative or stop closure, does not stand in opposition to sonority enhancement, even though it narrows the vocal tract. Even though Beckman et al. (1992) proposed that sonority consists of two dimensions, "the impedance of the vocal tract looking forward from the glottis" and "time" (p. 84), we will consider the primary correlate to be the openness of the jaw (which they measured in their paper) and its correlates, such as F1 and lower tongue position in the anterior region.

As already mentioned, /l/ in American English has been described as consisting of two articulatory gestures which vary as an effect of prosodic variation. In coda position, where it is considered to be more vowel like or more sonorous, the "vocalic" dorsal gesture is more prominent than in onset position, while the "consonantal" apical gesture is usually weakened. Proctor and Walker (2012) who looked at phonotactic constraints of American English liquids and rhotics and their articulation in different vowel and syllable contexts, argue that the more stable realization of the dorsal gesture of rhotics across conditions might explain why they are ranked higher in the sonority hierarchy than laterals in Germanic languages. We will take this as an additional argument as to why the dorsal retraction movement of the tongue does not conflict with sonority enhancement despite resulting in a more narrow back cavity of the vocal tract.

We hypothesize that in accented contexts, a stronger retraction of the dorsal gesture will be observed as it can be interpreted as both, sonority enhancement and hyperarticulation. For the apical gesture, the Hyperarticulation Hypothesis and the Sonority Enhancement Hypothesis make contradicting predictions: the former predicts a stronger apical constriction, while the latter predicts weakening of the apical constriction to allow for a more open oral cavity in the anterior region together with a lower jaw. Acoustically, dark /l/ has been described as having a smaller F2-F1 value (distance between F1 and F2) than light /l/ (Sproat & Fujimura, 1993), which Lin, Beddor, and Coetzee (2014) have shown to be correlated with a weaker apical constriction. More generally, a higher tongue body, which narrows the anterior part of the oral cavity, leads to a lower F1 and vice versa, while F2 is more sensitive to the front-back movement of the tongue body. A movement of the tongue body to the front leads to a higher F2 (Stevens, 1997). Studies by Sproat and Fujimura (1993) or Lin et al. (2014) seem to suggest that the apical constriction is weakened the more the tongue dorsum is retracted. This would make the prediction made by the Hyperarticulation Hypothesis, that both, the apical and dorsal gesture are enhanced, rather unlikely. However, Narayanan, Alwan, and Haker (1997) showed that the degree of weakening of the apical constriction for dark /l/ compared to light /l/ depended on how speakers articulated /l/ in general. The study was based on only four speakers and the articulation was examined for sustained speech sounds recorded with magnetic resonants imaging (MRI). They also carried out a second recording with electropalatography (EPG). Speakers were able to be grouped into those who produced a laminal /l/, where the entire tongue blade was raised, and those who produced an apical /l/, where only the tongue tip was raised. Those speakers with a laminal production showed significantly greater contact in the front region for light /l/ compared to dark /l/, and one speaker produced a dark /l/ without any contact in the alveolar region. For speakers with an apical /l/, the contact in the front region was comparable for the two varieties of /l/. All speakers showed a narrowed uvular or upper-pharyngeal region for dark /l/. In addition to the consonantal nucleus, we will also analyse words with /e/ in the target nucleus position in order to compare how prosodic emphasis is implemented on vowels in Slovak. /e/ is a front mid vowel, and the predictions made by the Hyperarticulation Hypothesis and Sonority Enhancement Hypothesis do not necessarily go hand in hand. For Slovak, the standard literature claims that vowel quality is not affected by prosody (Král' & Sabol, 1989), however Beňuš and Mády (2010) found that vowel quality was affected by phonemic quantity and word stress. As mentioned already in Section 1.3, unstressed vowels as well as short vowels were more centralized than stressed and long vowels, respectively. In particular, /e/ was found to centralize the most. Statistics were given for Euclidian distances from the centroid in the vowel space. However, visual inspection of Figure 1 in the paper by Beňuš and Mády (2010) suggests that for unstressed and short vowels, /e/ had a lower F2, while effects on F1 were less consistent. One speaker seemed to produce /e/ with higher F1 for the phonemically short condition and unaccented condition, while for the other speaker it was the opposite. Hanulíková and Hamann (2010) placed the vowel /e/ between the second and third cardinal vowel in the vowel diagram and mid way between the Slovak vowels /i/ and /a/. Possibly, /e/ tends to be less affected in the F1 dimension to maintain distance to the vowels /i/ and /a/. The effect of prosodic emphasis on /e/ has previously been studied by Erickson (2002), who observed a higher and more fronted tongue position, a lower jaw, and lower F1 and higher F2 for the accented condition. Although jaw lowering is in general expected to raise F1, Erickson (2002) proposes that a lower F1 and higher F2 can be generated by making the tongue higher an more fronted at the same time. The Sonority Enhancement Hypothesis predicts a lower jaw position in the accented condition, which acoustically results in higher F1. However, to avoid a greater acoustic overlap with /a/ in the F1 dimension the tongue might be positioned somewhat higher, which could be interpreted as an extra effort to maintain its acoustic features and hyperarticulation. Similarly to the effect of prosodic emphasis on /i/ in English, it is difficult to predict what will happen with /e/ regarding the vertical position of the anterior part of the tongue and F1. In addition, the Hyperarticulation Hypothesis predicts a lower F2 and a more fronted tongue position which should be captured at the posterior radius of the ultrasound data.

Our main question is whether /l/ in nucleus position lowers the tongue in anterior region in the accented condition to enhance sonority but at the cost of weakening the apical constriction, contrary to the prediction made by the Hyperarticulation Hypothesis. Pouplier and Beňuš (2011) showed that the phonological quantity contrast is visible for the consonantal apical gesture and that the apical gesture is indeed part of the nucleus. They suggested that the fact that /l/ and /r/ have two gestures might be related to the fact that they can function as syllabic consonants. So, while the apical gesture is present for syllabic /l/, it might be the vocalic gesture for which the main effect of accent will be observed, as predicted by both the Sonority Enhancement Hypothesis and the Hyperarticulation Hypothesis. We will also look at the vocalic nucleus, because to our knowledge next to nothing is known about the effect of phrasal accent on the quality of vowels in Slovak.

We will look at both articulatory and acoustic data in order to see the relationship between articulation and acoustics. Additionally, some articulatory data which we did not investigate such as jaw position, might be captured in the acoustics.

3.2 Methods

3.2.1 Stimuli

For four bisyllabic target words, shown in Table 3.1, the effect of prosodic strengthening will be analysed. The nucleus is either vocalic (/e/) or consonantal (/l/) and phonemically short or long. As already mentioned, /r/ was not analysed because it was difficult to identify the apical gesture in the ultrasound recordings. In Slovak, long nuclei are orthographically marked with an acute accent mark. Lexical stress in Slovak is fixed on the first syllable; thus in our target words, the target nucleus is stressed. In this position, the effect of phrasal accent is expected to be strongest.

Table 3.1: List of target words: The sequence used for articulatory analysis is presented in bold text.

	/e/	/l/
short	**pep**ap	**plp**ap
long	**pép**ap	**pĺp**ap

In order to elicit the two accentuation patterns, target words were presented in two carrier phrases:

(1) Pozri, ved' on mi ***plpap*** dal.
‘Look, in the end he gave me ***plpap***.’
(2) Pozri, aj **Ron** mi *plpap* dal.
‘Look, also **Ron** gave me *plpap*.’

In the first carrier phrase, the target word receives the main phrasal accent. In the second carrier phrase, *Ron* is accented.

For details on acoustic and articulatory recordings and segmental principles see Section 1.5.

3.2.2 Measurements

Ultrasound data

Based on acoustic segmentation, tongue contours were tracked from the beginning of the acoustic closure of the first *p* of the target word to the end of the closure of the second *p*, (the bold part in Table 3.1). We will refer to this part as the target sequence. The contour data represented in polar coordinates were exported and further analysis was carried out in R. Analysis of the articulatory data has been carried out throughout the entire acoustic target sequence, because the main movement can be seen during the flanking /p/s.

For each speaker, two angles were selected, at the front and back part of the tongue, separately for target words with vocalic and consonantal nuclei. This was necessary, because the position of the tongue relative to the probe depended on the rotation of the probe during the recording and because of individual anatomical differences. The angle was chosen based on whether the contour was visible at that angle throughout the entire target sequence for all tokens, and by visual verification that it captures a comparable region for all speakers. Conceptually, the angles correspond to scan lines of the ultrasound beam that radiates out of the probe. For an impression, tongue contours for each frame within the first repetition of ***pĺp****ap* in accented condition with the angles for each speaker are given in Figure 3.1 and for ***pép****ap* in Figure 3.2. The angles for each speaker are also listed in the caption. In AAA tongue contours are represented in radians with 0 radians corresponding to 90° in the plots. The contours were rotated only for plotting reasons and the angles listed in the caption correspond to the angles indicated in the plot. The posterior angle should capture the tongue retraction movement, while the anterior angle captures the height of the tongue body and especially for speakers who have a very curled tongue during the production of /l/ the dip is captured (see Figure 3.1). At these angles the tongue was visible throughout the entire target sequence for all tokens. The measurement at the anterior part of the tongue differs from measures used traditionally where tongue tip is used to capture the constriction degree. However, Smith and Lammert (2013) investigated this curling of the tongue blade during the production of /l/ in American English. While they only report the degree of curving, they noted that “[d]uring vocalized /l/ in syllable nucleus and coda position, tongue curling was produced, not by raising of the tongue tip, but by lowering of the tongue blade” (p.3233). This can be interpreted that in their data the tongue blade was lower in nucleus and coda position. In our recordings, the tongue tip was not always visible. For some speakers this was sometimes the case during the crucial apical constriction of /l/, where the tongue tip was probably parallel to the ultrasound beams. In such a case, the ultrasound beams radiating out of the probe cannot be reflected. In other cases, the tongue would not reach the region of the angle that would capture tongue tip raising for /l/ during other segments. So the region where the tongue curving takes place seems to be a good alternative gestural landmark. For each ultrasound frame within the target sequence the distance of the tongue from the origin of the superimposed grid was retrieved at the two angles. This resulted in two curves per token, representing the radius over time. At the anterior angle, a greater radius indicates that the tongue is closer to the palate, so we will refer to it as tongue height. At the posterior angle, a higher radius indicates a more narrow constriction in the uvular region and we will refer to it as tongue retraction. The radius was z-normalized for each speaker.

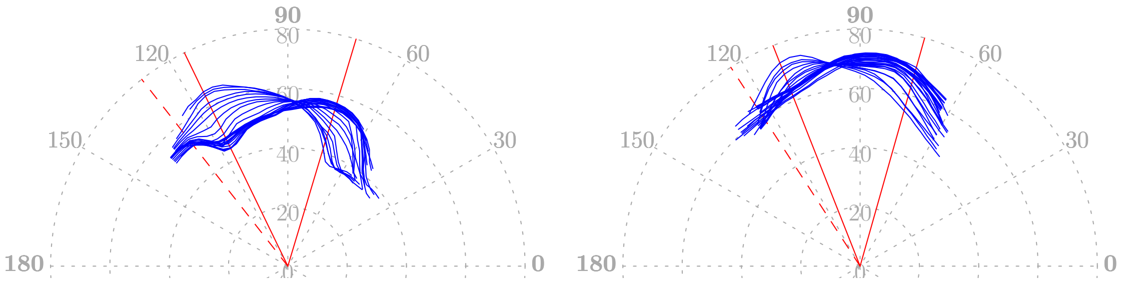
To analyse tongue movement over time (normalized from 0 to 1) as captured at selected angles, functional mixed modelling was used, separately for vocalic and consonantal nuclei and for tongue height and tongue retraction, resulting in four model analyses. Each model consisted of two covariates: PHRASAL ACCENT (0 = unaccented, 1 = accented) and PHONEMIC QUANTITY, and their interaction. The overview of covariates and their interaction is given in Table 3.2. The reference mean represents the condition where all dummy codings are set to 0, corresponding to the tongue movement measured at the selected angle for the unaccented target sequence with short nucleus.

Table 3.2: Overview of covariates for articulatory and acoustic analyses.

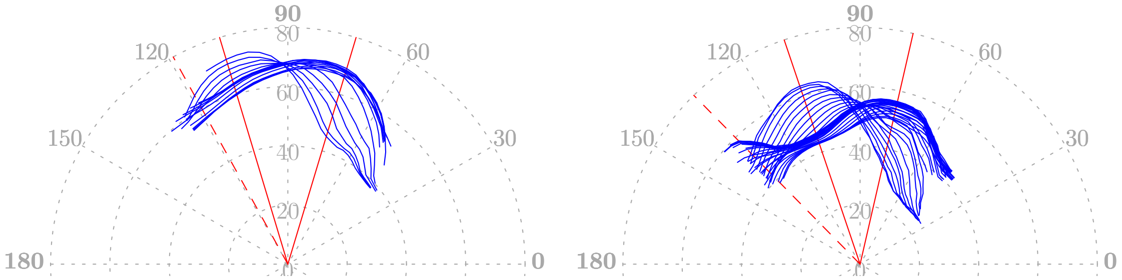
$f_0(t)$	Reference mean	unaccented short nucleus
$f_1(t)$	Phrasal accent (covariate 1)	0 = unaccented, 1 = accented
$f_2(t)$	Phonemic quantity (covariate 2)	0 = short, 1 = long
$f_3(t)$	Phrasal accent * Phonemic quantity	

3.2.3 Formant analysis

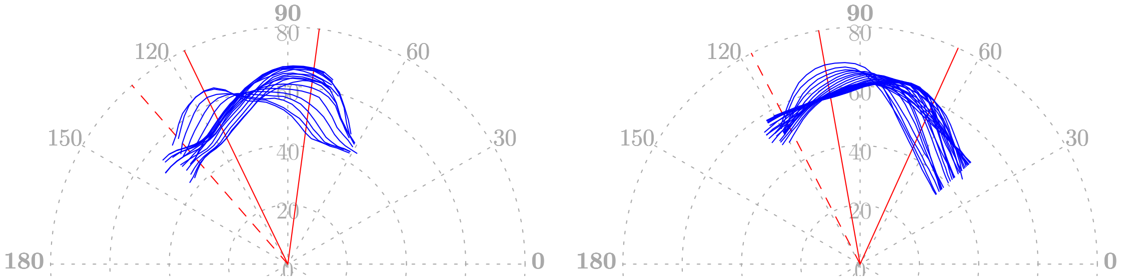
For formant analysis, we only looked at the nucleus portion of the target sequences listed in Table 3.1, as it does not make sense to search for formants in the voiceless flanking consonants. Formant data was extracted using the forest function within the wrassp package within EmuR (Bombien et al., 2016; Winkelmann et al., 2016). The forest function uses a Blackman window with the window size of 20 ms which is shifted for every 5 ms. We examined



(a) Speaker 1 (b) Speaker 2

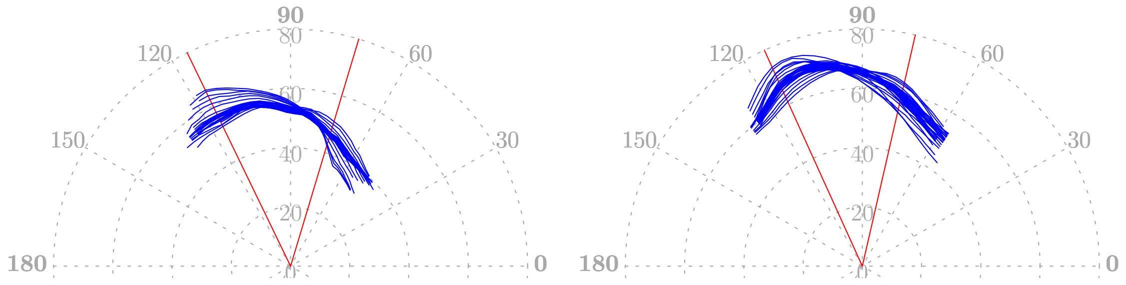


(c) Speaker 3 (d) Speaker 4

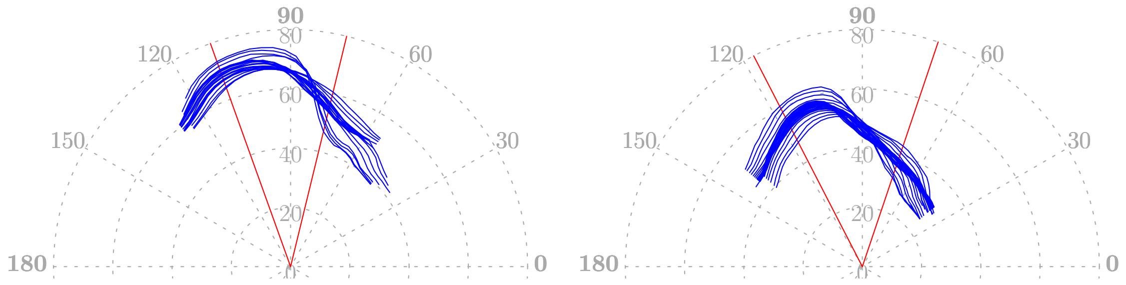


(e) Speaker 5 (f) Speaker 6

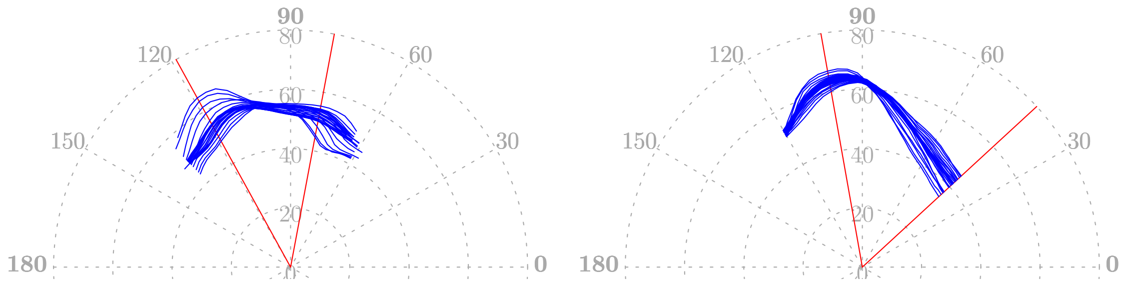
Figure 3.1: Contour data of first repetition of ***pĺp****ap* for each speaker. Tongue tip is on the left. The angles at which the data for the articulatory analysis of the anterior and posterior region were retrieved are indicated using red solid lines. The red dashed line indicates the angle at which the tongue tip movement for the timing analysis in Chapter 4 was captured. The angles for the tongue tip, anterior and posterior region are as follows: Speaker 1: 128.0°, 112.8°, 64.15°; Speaker 2: 123.0°, 111.5°, 74.2°; Speaker 3: 118.9°, 106.7°, 73.3°; Speaker 4: 132.5°, 108.7°, 77.1°; Speaker 5: 131.1°, 115.9°, 82.4°; Speaker 6: 117.3°, 100.1°, 65.6°



(a) Speaker 1 (b) Speaker 2



(c) Speaker 3 (d) Speaker 4



(e) Speaker 5 (f) Speaker 6

Figure 3.2: Contour data of first repetition of ***pép****ap* for each speaker. Tongue tip is on the left. The angles at which the measurements for the anterior and posterior region were captured are indicated in red. The angles for the anterior and posterior region are as follows: Speaker 1: 115.85°, 73.28°; Speaker 2: 114.4°, 77.1°; Speaker 3: 109.8°, 76.3°; Speaker 4: 117.3°, 71.34°; Speaker 5: 118.9°, 79.4°; Speaker 6: 105.8°, 42.6°

formant contours during the entire nucleus instead of extracting formant values at only one specific time point. The formant data extracted for the acoustically segmented nucleus were time normalized by mapping each data point onto a time interval from 0 to 1. FLMM was used to evaluate the effect of PHRASAL ACCENT and PHONEMIC QUANTITY, separately, on F1 and F2 of vocalic and consonantal nuclei (resulting in four separate analyses). The covariates are the same as for the articulatory analysis, listed in Table 3.2.

3.3 Results

The results will be presented separately for vocalic and consonantal nuclei. Before we begin with the results of the ultrasound data, we will briefly report the acoustic duration of each segment. In the previous chapter, we showed that the main difference was between phonemically short and long nuclei, with long nuclei being about twice as long as short nuclei. Phrasal accent did not have an effect on phonemically short nuclei. However, phonemically long nuclei were longer in the accented condition than in the unaccented condition. Because we will compare articulatory data of sequences with long and short nuclei with normalized time, it is helpful to know how the temporal proportions differ for each context. Separately for the target words with the vocalic and consonantal nucleus, we will first present how phrasal accent and phonemic length of the nucleus affected the acoustic duration of consonants preceding and following the nucleus, and visualize how, on a time scale from 0 to 1, the target sequence is divided into the three segments - initial /p/ (*p1*), the nucleus (*N*) and the second /p/ (*p2*).

3.3.1 Vocalic nucleus

Figure 3.3 shows the absolute duration of each segment in the top plot, and in the bottom plot how the target sequence is proportionally divided in each segment for the different contexts. We will first report how individual segments were affected by phrasal accent and phonemic length of the nucleus and then interpret the proportional make-up when the segment is time normalized.

Acoustic duration of the onset consonant *p1* was affected significantly by PHRASAL AC-

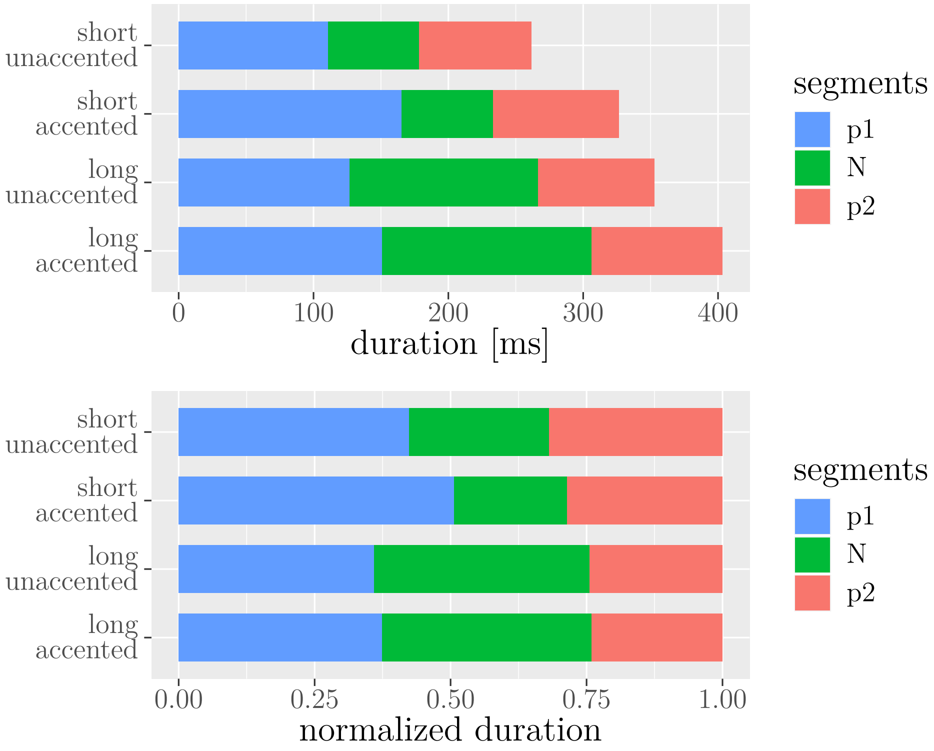


Figure 3.3: Absolute (top) and normalized (bottom) mean durations of segments in target sequences with vocalic nuclei.

CENT ($\chi^2[1] = 39.8, p < 0.001$). PHONEMIC QUANTITY of the nucleus did not have a significant effect ($\chi^2[1] = 2 * 10^-4, p > 0.1$), but the interaction of PHRASAL ACCENT and PHONEMIC QUANTITY of the nucleus did ($\chi^2[1] = 6.7, p < 0.01$). The results are visualized in Figure 3.4 for clarity. We see that PHRASAL ACCENT had the effect of lengthening the duration of *p1*. The effect was greater on *p1* when the nucleus was short. So in the unaccented condition the acoustic duration of *p1* preceding a phonemically short nucleus was shorter than the duration of *p1* preceding a phonemically long nucleus. In the accented condition, the duration of *p1* before a phonemically short nucleus was longer than when it preceded a phonemically long nucleus. The mean duration of *p1* preceding a phonemically long vowel was $151.5ms \pm 11.0$ when accented and $126.6ms \pm 11.0$ when unaccented. The acoustic duration of *p1* before a phonemically short vowel was $166.1ms \pm 11.0$ in the accented condition and $111.6ms \pm 11.0$ in the unaccented condition.

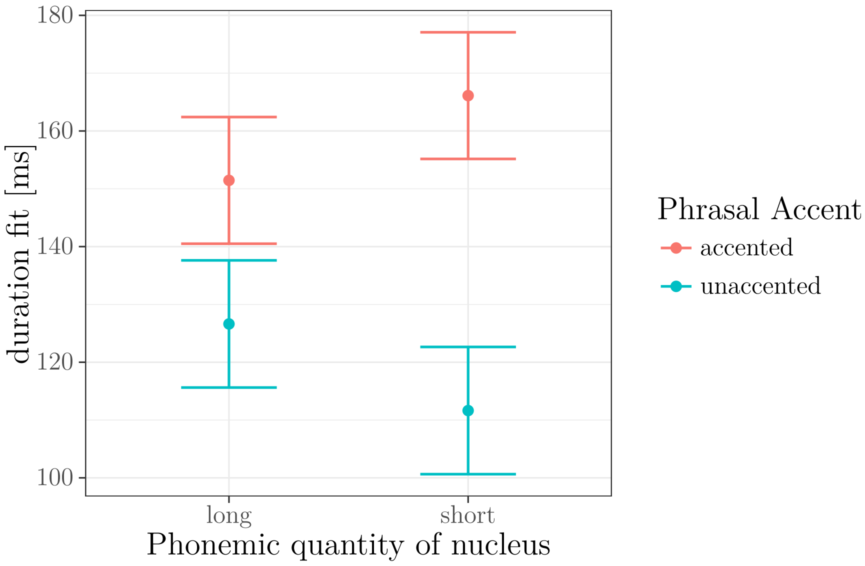


Figure 3.4: Visualized effects of PHRASAL ACCENT and PHONEMIC QUANTITY of the nucleus on the duration of *p1* as modelled using linear mixed effect models. The plot shows the condition mean and standard error.

Acoustic duration of *p2*, the consonant following the target nucleus, was affected significantly only by PHRASAL ACCENT ($\chi^2[1] = 34.7, p < 0.001$). PHONEMIC QUANTITY of the nucleus ($\chi^2[1] = 0.6, p > 0.1$) and the interaction of PHRASAL ACCENT and PHONEMIC QUANTITY ($\chi^2[1] = 0.4, p > 0.1$) did not have a significant effect on *p2*. The mean duration of *p2* in accented condition was $93.2ms \pm 5.9$. In the unaccented context, the duration of *p2* was $82.4ms \pm 5.9$.

Although the duration of short nuclei was not affected by phrasal accent, when time normalized the accented short nucleus takes up a smaller portion of the target sequence than the unaccented short nucleus, because *p1* and *p2* were significantly longer in accented condition. For target words with long nuclei, however, the proportion of the segments within the target sequence is not much affected by phrasal accent, because all three segments are lengthened under accentuation.

Results of the ultrasound recordings of the vocalic nucleus

Posterior angle We begin with the results for the tongue retraction movement in target sequences with vocalic nuclei (*pep* or *pép*). Recall that the tongue contour was represented in polar coordinates and for each speaker one angle in the posterior part of the oral cavity was chosen at which the position of the tongue was measured. This resulted in curves representing the radius over time. A higher radius indicates a more retracted tongue position. /e/ is a front vowel, thus if /e/ is less centralized, it should be more fronted in the accented condition, and we should see a smaller radius for the posterior angle. The radius was z-normalized by speaker to get reasonable results using the FLMM statistics. In Figure 3.5 the data which was used for the statistical analysis is visualized. The z-normalised data is on the right. On the left, the data before z-normalization is shown, where the curves for each individual speaker are more visible due to less overlap. Because the vowel preceding the target sequence is /i/ and has an extremely fronted tongue, we can observe tongue retraction movement into /e/ at the beginning. At the end further retraction can be seen towards the following vowel /a/. During the central part, in some cases a fronting movement can be observed, while in other cases the tongue position is maintained or a slower retraction movement can be observed.

FLMM analysis was carried out with the levels PHRASAL ACCENT and PHONEMIC QUANTITY.

In Figure 3.6 the reference mean, covariate effects and their interaction effect are shown. The reference mean $f_0(t)$ (Figure 3.6 (a)) corresponds to the movement of the posterior part of the tongue for the unaccented condition of the short nucleus (during *pep* of target word *pepap*). A higher value indicates a more retracted tongue position. From time point 0 to about 0.4 and then again from about 0.8 to 1 the radius rises, indicating a retraction gesture. The movements take place mainly during the lip closure of the flanking /p/s. In Figure 3.11 the summed effect curves are presented separately for each condition with their corresponding acoustic boundaries. From time point 0.4 to 0.8, the tongue position is maintained. To better understand the covariate and interaction effects, summed effect curves are shown in Figure 3.7. The red dashed line in Figure 3.7 corresponds to the reference mean curve $f_0(t)$ in Figure 3.6 (a), which is the tongue retraction movement during the unaccented target sequence with short nucleus. The covariate effect curve $f_1(t)$ for PHRASAL ACCENT

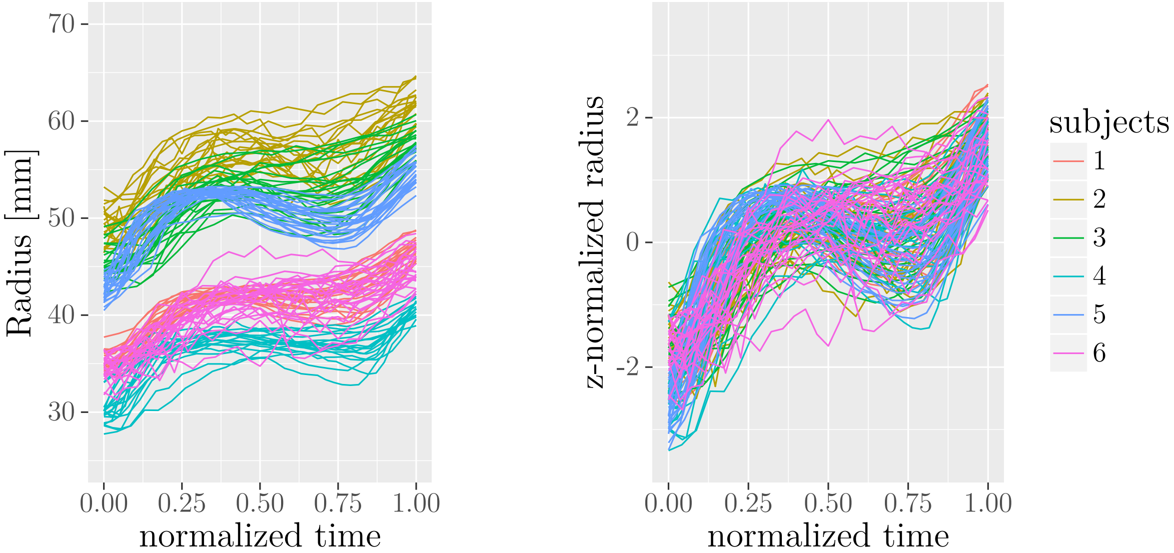
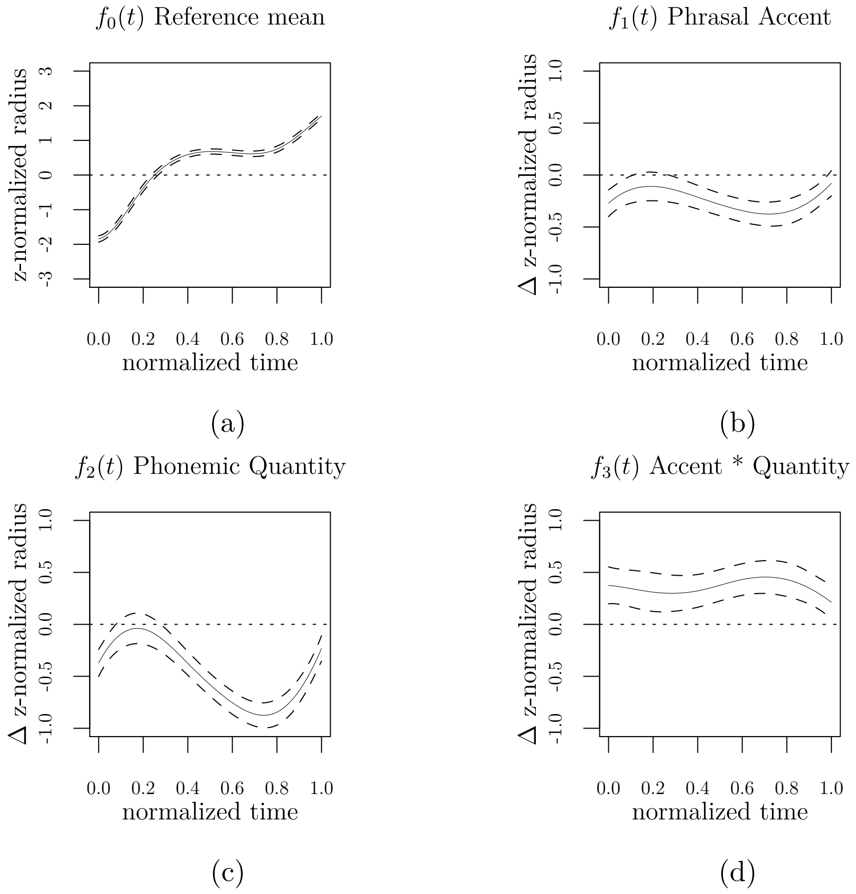


Figure 3.5: Articulatory data captured at the posterior angle of target sequences with vocalic nuclei: *pep* or *pép*. On the left radius over normalized time is plotted, and on the right z-normalized radius over normalized time. A higher radius indicates a more retracted tongue.

(Figure 3.6 (b)) lies below zero which indicates that in the accented condition the tongue is more fronted than in the unaccented condition, with the effect being significant after around time point 0.4. This can also be seen when we compare the red solid line (corresponding to the accented condition) with the red dashed line (unaccented condition) in Figure 3.7. While the red dashed line and red solid line are very close at the beginning until around timepoint 0.35, the solid line slightly declines until around time point 0.75, but then rises again. The effect of PHONEMIC QUANTITY $f_2(t)$ (Figure 3.6 (c)) is similar to that of PHRASAL ACCENT but has a greater magnitude. The summed effect curve corresponding to the unaccented target sequence with long nucleus is visualized in Figure 3.7 with a dashed blue line. For the long nuclei, we can see that after a peak is reached at around time point 0.3, it falls slightly until around time point 0.7, where it starts rising again towards its position for the following /a/. So for long vowels, we observe a slight fronting movement during the nucleus. The interaction between PHRASAL ACCENT and PHONEMIC QUANTITY has a significant effect as well, but in the opposite direction to the two covariates. The interaction effect curve $f_3(t)$ in Figure 3.6 (d) is significantly above zero and the effect is greater during the second half of the target sequence. To get the summed effect curve for the accented target sequence with long nucleus, both covariate effect curves and the interaction effect curve are added to the reference mean. As a result, for the long nucleus in accented condition compared to the long nucleus in unaccented condition, shown in Figure 3.7 using the blue solid and dashed lines respectively, we do not see a fronting effect as for short nuclei. For short nuclei phrasal accent does have a significant effect after time point 0.4 and the difference between the red dashed line and red solid line is significant.

In sum, at the posterior angle we could observe the effect predicted by the Hyperarticulation Hypothesis. For the short vowels, the tongue is fronted during the nucleus in the accented condition. Long vowels have a fronter tongue position in general, but no accent induced fronting effect was observed. Most probably, there is a ceiling effect because the tongue has more time to reach the target, also in the unaccented condition. Recall that the unaccented long vowel is still longer than any of the short nuclei (Chapter 2, Figure 2.6).



(a) (b)

(c) (d)

Figure 3.6: Reference mean and covariate effects of tongue movement captured at the posterior angle in target sequences with vocalic nuclei. A greater radius indicates a more retracted tongue.

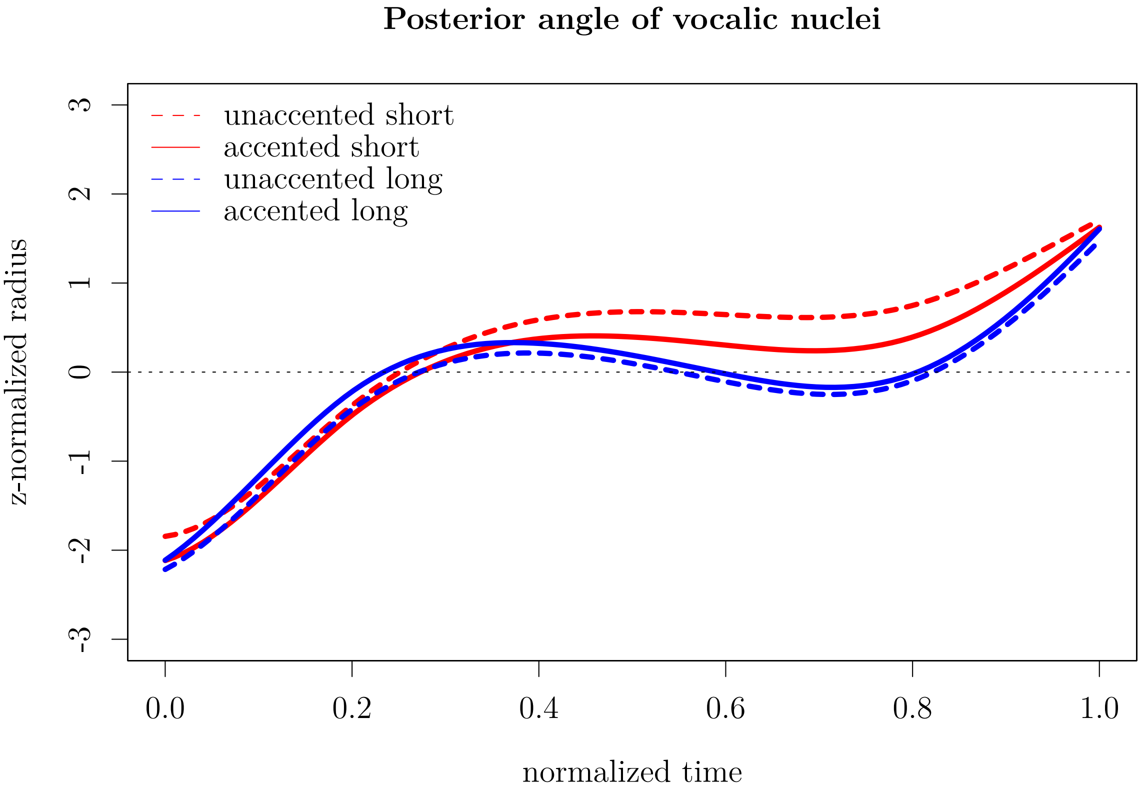


Figure 3.7: Summed effect curves of tongue retraction movement captured at the posterior angle in target sequences with vocalic nuclei. A higher radius indicates a more retracted tongue. The red dashed line represents the reference mean $f_0(t)$ which corresponds to the short unaccented nucleus condition. The red solid line represents the effect of PHRASAL ACCENT $f_1(t)$ added to the reference mean, and corresponds to the short accented nucleus condition. The blue dashed line represents the condition where the covariate effect curve for PHONEMIC QUANTITY, $f_2(t)$, is added to the reference mean, which corresponds to the long unaccented condition. The blue solid line represents the effects of PHRASAL ACCENT $f_1(t)$, PHONEMIC QUANTITY $f_2(t)$ and the interaction effect curve $f_3(t)$ added to the reference mean and represents the long accented nucleus condition.

Anterior angle We now continue with the results of the tongue movement captured at the anterior angle in sequences with vocalic nucleus. At this angle a greater radius indicates a higher tongue position. According to the Sonority Enhancement Hypothesis we expect a lower tongue position to go with a lower jaw for the accented condition. However, according to the Hyperarticulation Hypothesis, it is not clear what to expect, as /e/ is a mid vowel. For hyperarticulation of /e/ we do not expect a change in the acoustic F1 dimension which in general is associated with jaw and tongue height. But if the jaw is lowered as expected for the Sonority Enhancement Hypothesis, the tongue might be raised slightly to avoid F1 movement closer to that of a low vowel. In addition the target position for /e/ might be held for a longer period in the accented condition. The data as it entered the analysis is visualised in Figure 3.8. For tongue height, we see a lowering movement from the preceding high vowel /i/ to /e/, a mid vowel at the beginning during the first *p*, and again at the end towards the following vowel /a/, a low vowel, during the second *p*. In between, in some cases the tongue position is held steady, while in other cases the tongue position is raised.

The reference mean corresponding to tongue height during the unaccented target sequence with short nucleus, covariate effects of PHRASAL ACCENT and PHONEMIC QUANTITY, and their interaction are shown in Figure 3.9. The summed effect curves are shown in Figure 3.10. In Figure 3.11 the summed effect curves are presented separately for each nucleus condition but with the corresponding acoustic boundaries. As observed for the raw data, the reference mean in Figure 3.9 (a) shows that the tongue position is high at the beginning and is being lowered until around time point 0.3, which is during the initial consonant. In Figure 3.11 we can see that for the unaccented short vocalic nucleus the boundary between the initial consonant and the nucleus, retrieved using the mean of the acoustically determined duration, is after time point 0.4. The position is held during the nucleus until around time point 0.7, after which the radius becomes smaller, meaning that the tongue is further lowered during the second consonant. PHRASAL ACCENT has a significant lowering effect during the first half of the target sequence and a significant raising effect in the second half. Comparing the dashed red line for the unaccented condition and the solid red line for the accented condition, we can see that the solid line is lower than the dashed line during the first half of the target sequence and higher during the second half. As visualized in Figure 3.3, the short unaccented nucleus begins acoustically around time point 0.4 while

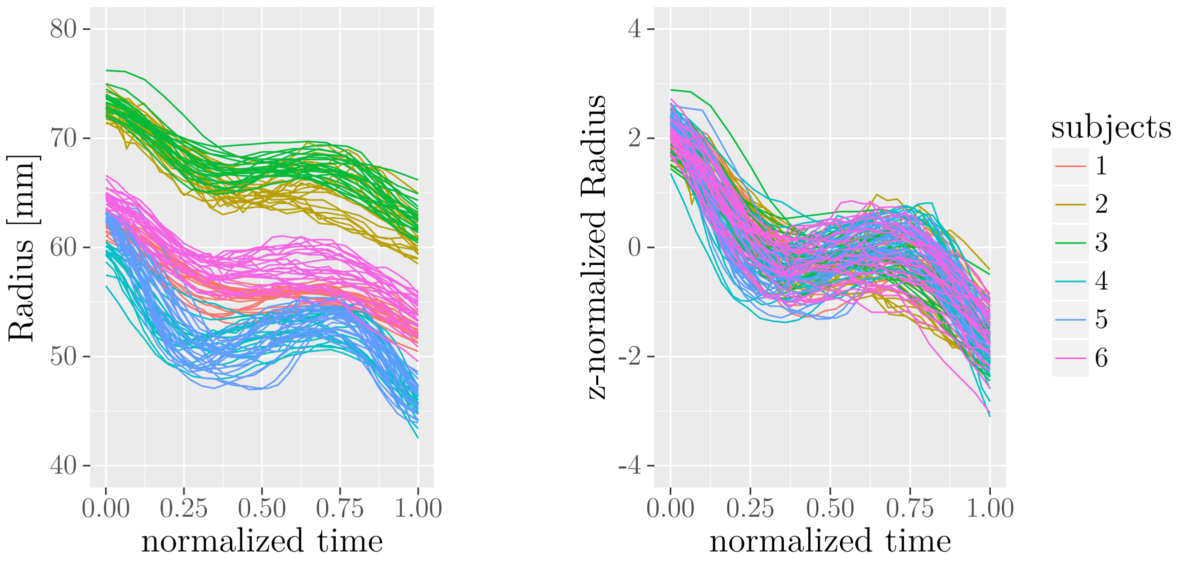
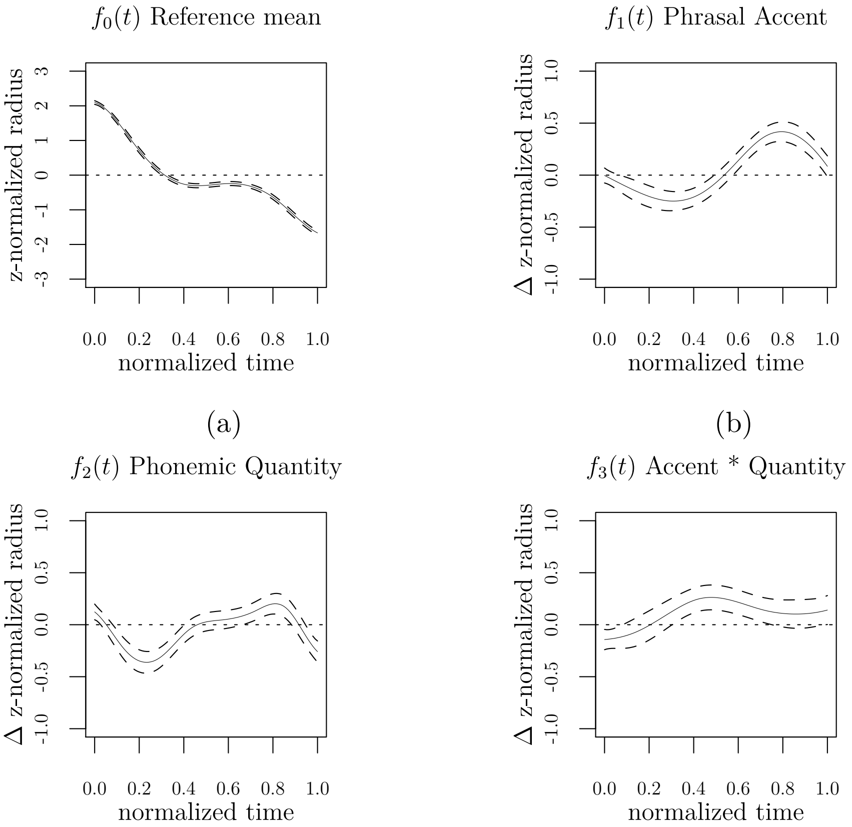


Figure 3.8: Articulatory data measured at the anterior angle of target sequences with vocalic nuclei: *pep* or *pép*. On the left the radius over normalized time, on the right z-normalized radius. A higher radius indicates a higher tongue.

the short accented nucleus begins at around time point 0.5. For the accented condition the tongue reaches a lower position during the initial consonant but rises again, and at the beginning of the nucleus it is about as high as for the unaccented condition. However, in the accented condition the tongue rises during the nucleus, so towards the end of the nucleus it is higher than in the unaccented condition. PHONEMIC QUANTITY has a significant lowering effect between time points 0.1 and 0.4 and a significant raising effect around time point 0.8. When we compare the red and blue dashed lines in Figure 3.10, we see that the falling slope at the beginning is steeper for the condition with the long nucleus. This might be an artefact of time normalization. Due to the longer nucleus, the flanking consonants take up proportionally less time, so the slope is steeper to reach the same position. During the nucleus, which takes up proportionally more time for the long vowel, the position is held. The falling slope at the end of the target sequence begins proportionally later for the sequence with the long nucleus compared to the sequence with the short nucleus. In Figure 3.11 we can see that the tongue lowering begins more or less together with the beginning of the acoustically determined beginning of the following consonant. The interaction of PHRASAL ACCENT and PHONEMIC LENGTH has a significant raising effect from slightly before time point 0.4 to around time point 0.7. After that it has a raising effect at trend level. For long vowels the tongue is higher in the accented condition compared to the unaccented condition with the difference being greater at the end of the nucleus than at the beginning of the nucleus. This is analogous to the effect on short vowels. By comparing the tongue movement during the entire target sequence we see a greater magnitude in tongue movement as a result of accentuation for both the short and long nuclei. In both cases we observe a raising tongue movement during the vowel which can be interpreted as hyperarticulation.



(a) (b)

(c) (d)

Figure 3.9: Reference mean and covariate effects of tongue movement captured at the anterior angle in target sequences with vocalic nuclei. A greater radius indicates a higher tongue position.

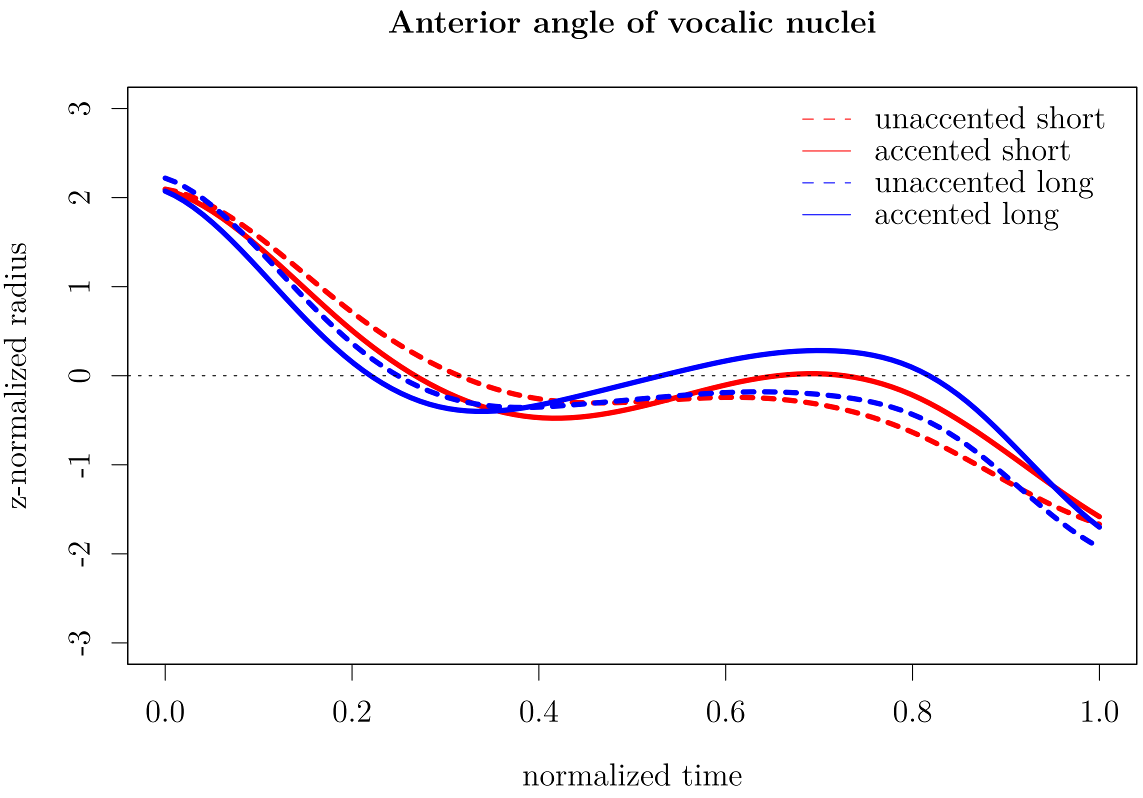


Figure 3.10: Summed effect curves of tongue movement captured at the anterior angle in target sequences with vocalic nuclei. A greater radius indicates a higher tongue position. The red dashed line, which is also the reference mean $f_0(t)$, corresponds to the unaccented short nucleus condition, the red solid line to the accented short nucleus condition $(f_0(t) + f_1(t))$, the blue dashed line to the unaccented long condition $(f_0(t) + f_2(t))$ and the blue solid line to the accented long condition $(f_0(t) + f_1(t) + f_2(t) + f_3(t))$.

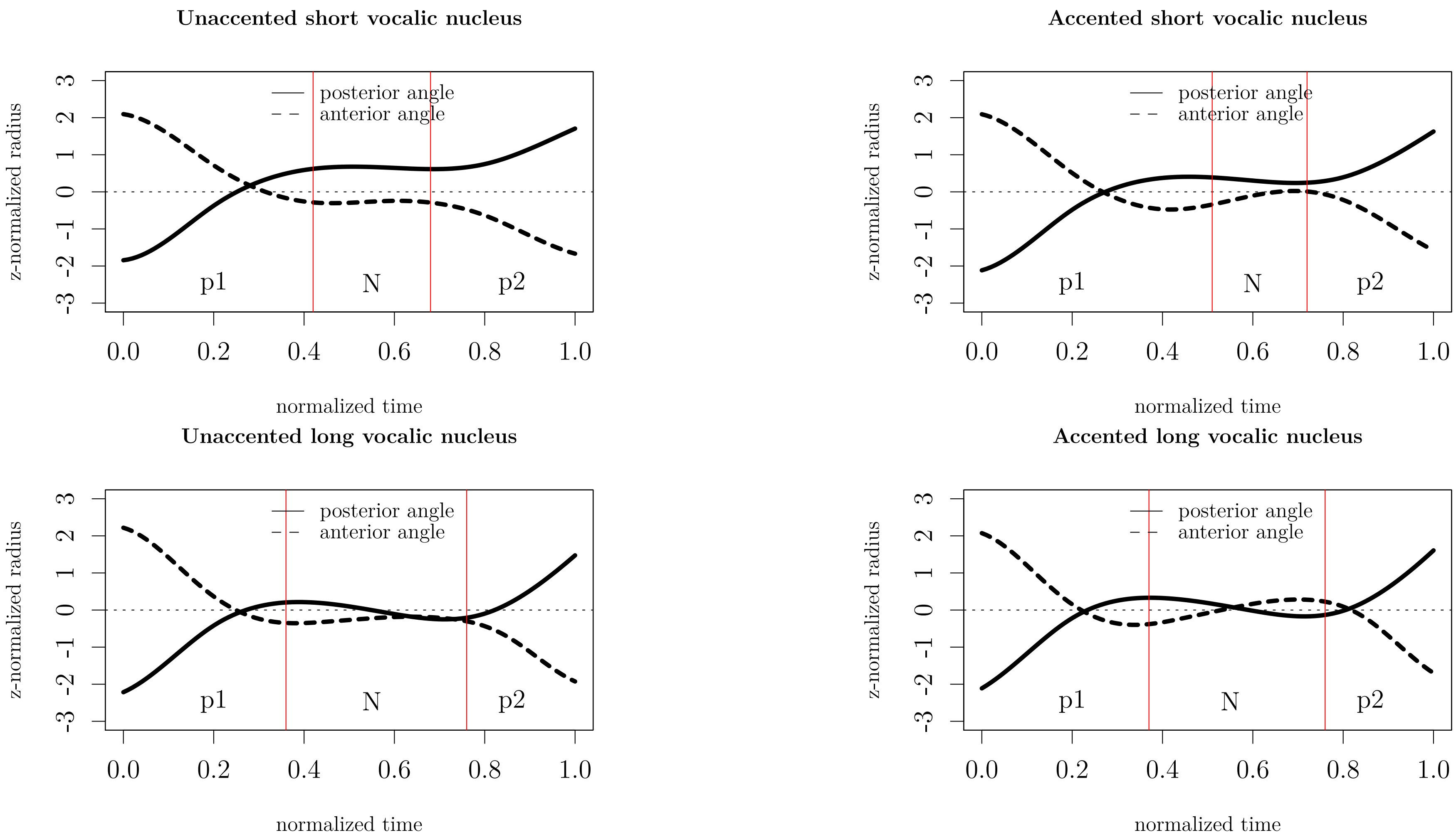


Figure 3.11: Summed effect curves of tongue movement captured at posterior and anterior radii separately for each nucleus condition. The red vertical lines represent the boundaries between the acoustically determined segments *p1* (onset consonant /p/), *N* (vocalic nucleus /e/), *p2* (/p/ following the nucleus). A smaller radius for the anterior angle indicates tongue lowering, a smaller radius for the posterior angle indicates tongue fronting.

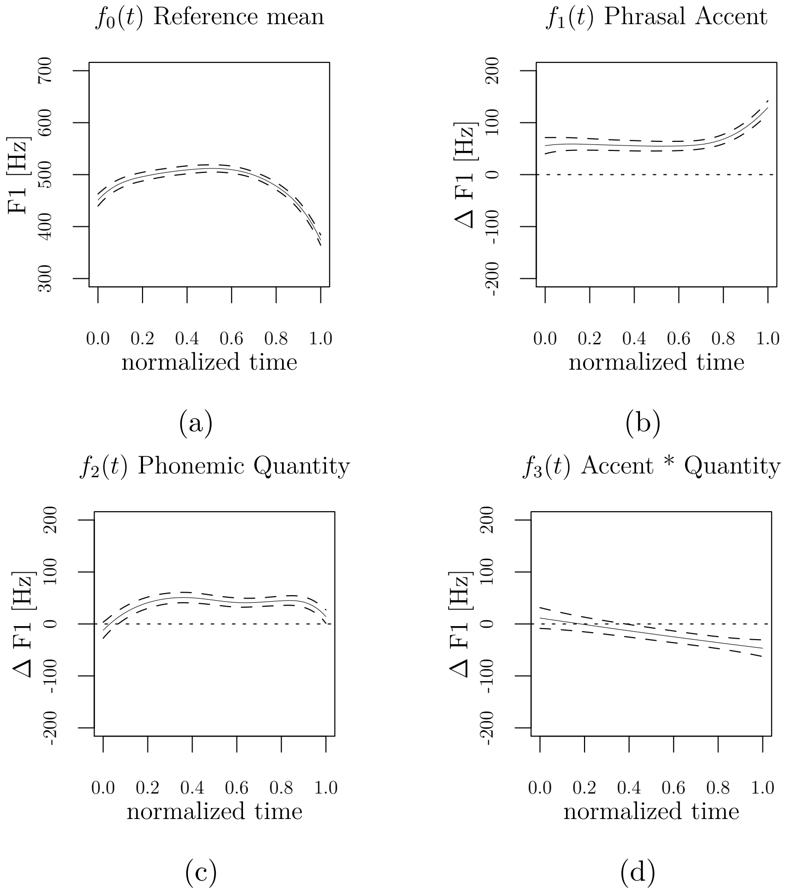
Results of the formant analysis for the vocalic nucleus

For the formant analysis we carried out a FLMM for F1 and F2 during the acoustically segmented /e/. Recall that the articulatory data was analysed not only during the acoustically segmented nucleus but also during the preceding and following consonants. The formant data are measured during the nucleus only and cover a different time span (see also Figure 3.11). The model consisted of the covariate effects PHRASAL ACCENT and PHONEMIC QUANTITY and their interaction. The results for F1 will be presented first, followed by the results for F2. A lower jaw position which is often observed for accented vowels and is predicted by the Sonority Enhancement Hypothesis predicts a higher F1. The Hyperarticulation Hypothesis predicts in general more peripheral formant values in the vowel space, which for /e/ means a higher F2 but not necessarily a different F1 to keep distant from both /i/ and /a/. The articulatory data showed a more fronted tongue in the accented condition for short nuclei. Long nuclei were not more fronted in the accented compared to unaccented condition but were in general more fronted than short nuclei. Based on the articulatory data, a higher F2 is expected for long nuclei compared to short nuclei and an accent induced higher F2 only for short nuclei. At the beginning of the nucleus the anterior part of the tongue was either lower in the accented condition compared to the unaccented condition or at the same height. For the accented condition a rising movement could be observed, and towards the end of the nucleus the tongue was higher in the accented condition compared to the unaccented position. Although a higher tongue position in the anterior region as observed in our articulatory data towards the end of the nucleus in accented condition is associated with a lower F1, F1 is also affected by jaw height which we did not analyse. Based only on the tongue data a declining F1 should be observed for the accented condition, so F1 will be lower towards the end of the nucleus compared to the unaccented condition for which a rather flat F1 is expected.

F1 Figure 3.12 shows the reference mean and covariate effects of PHRASAL ACCENT and PHONEMIC QUANTITY and their interaction effect for F1. The summed effect curves are visualized in Figure 3.13. The reference mean $f_0(t)$, corresponding to the F1 of the unaccented phonemically short /e/, shows a relatively flat formant curve with some rising transition at the beginning of the nucleus and a slightly steeper declining transition at the end of the nucleus. The consonants preceding and following the nucleus are both bilabial consonants

/p/, for which low loci are characteristic (Delattre et al., 1955). PHRASAL ACCENT had an effect of raising F1 throughout the entire nucleus (Figure 3.12(b)). The covariate effect curve $f_1(t)$ is flat at the beginning and rises after time point 0.8 which results in compensation of the falling slope at the end of the reference mean. Thus the F1 curve of the accented condition is flatter towards the end, as can be seen in Figure 3.13. The red dashed line and red solid line are parallel at the beginning but towards the end the red solid line corresponding to F1 of the accented condition does not fall as much as the red dashed line. PHONEMIC QUANTITY also had a significant raising effect on F1 (Figure 3.12 (c)). In Figure 3.13 we can see that the F1 of long nuclei (blue dashed line) begins at the same height as the F1 of short nuclei (red dashed line) but rises more at the beginning. It is significantly higher throughout almost the entire nucleus. Towards the end the declination of the F1 of long nuclei is steeper and consequently F1 of long nuclei reaches the same height as F1 of short nuclei. The interaction effect curve $f_3(t)$ is a linearly falling slope (Figure 3.12 (d)). During the first half of the nucleus the effect is not significant, but during the second half it has a significant lowering effect. In Figure 3.13 we can see that the difference between the red lines (short nuclei) becomes greater after time point 0.8 while the difference between the blue lines (long nuclei) does not become greater.

The Sonority Enhancement Hypothesis predicted a higher F1 in the accented condition which reflects a wider oral cavity, induced by a lower jaw and lower tongue position. The results of the tongue movement data captured at the anterior angle did not show a lower tongue position during the nucleus in accented condition, so we did not expect a higher F1, but the movement was analysed at only one radius. Furthermore, it might have been the jaw that mainly contributed to a higher F1. Because of the interaction effect curve $f_3(t)$ the summed effect curve for long accented nuclei has a slightly steeper declining movement compared to the summed effect curve for long unaccented nuclei. This matches with the articulatory data so far as for the accented condition we observed a rising movement of the tongue in the anterior region for the accented condition. We will return to this apparently contradictory result in the discussion.



(a) (b)

(c) (d)

Figure 3.12: Reference mean and covariate effect curves for F1 in vocalic nuclei.

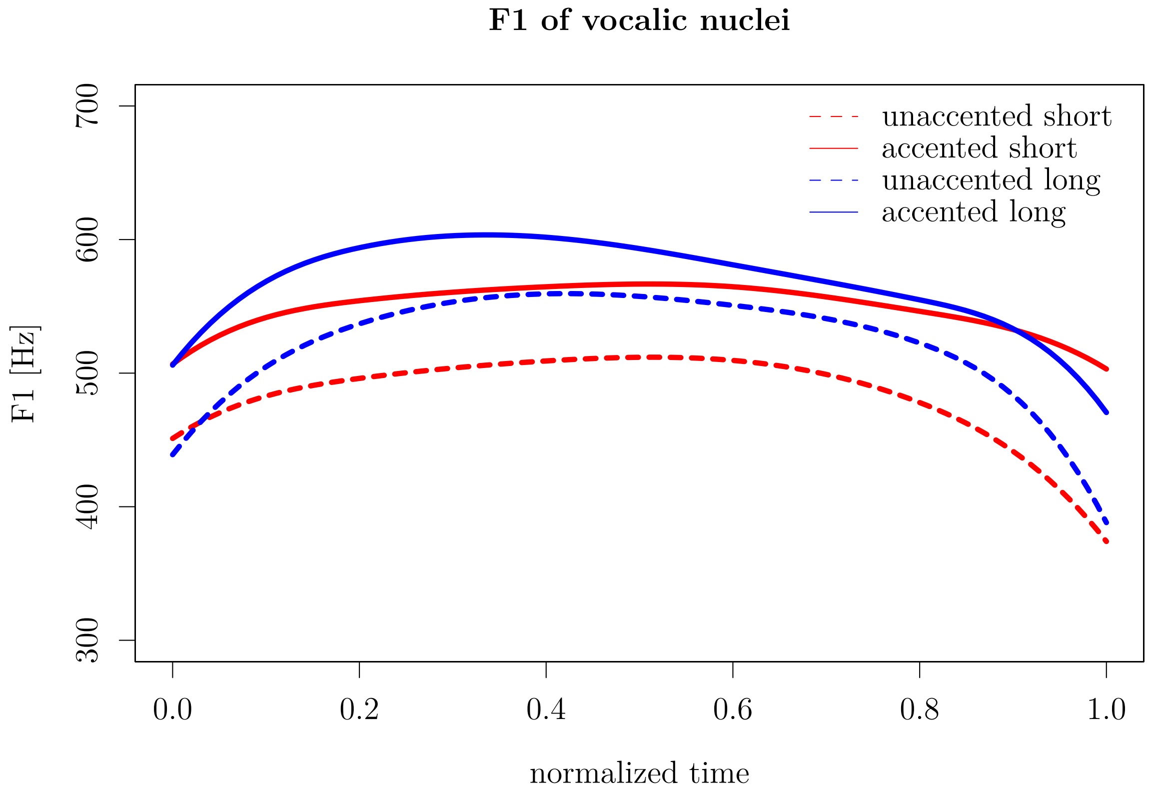
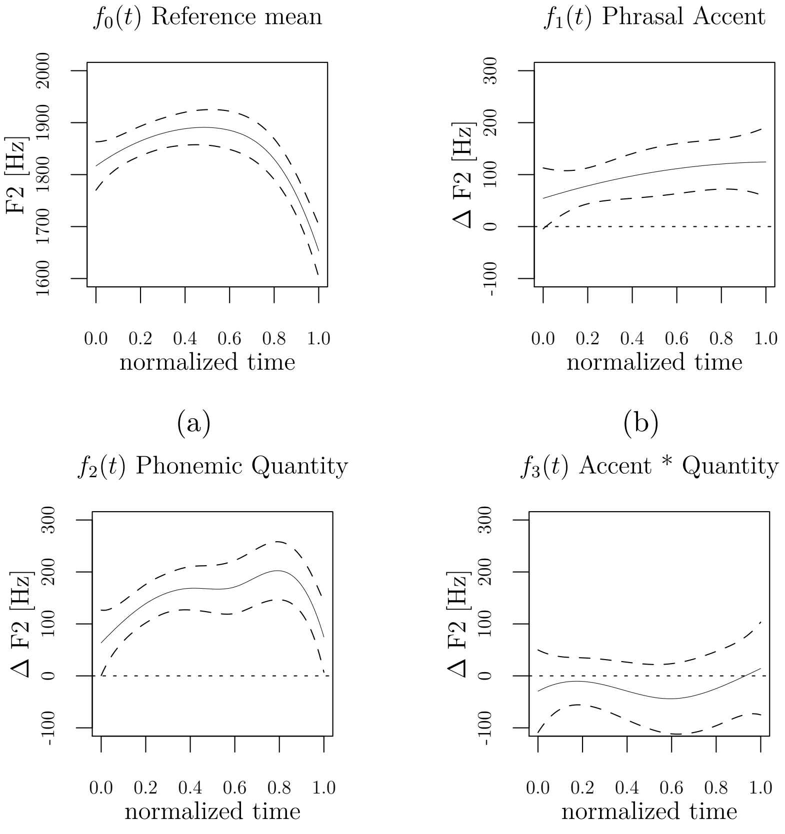


Figure 3.13: Summed effects curves for F1 in vocalic nuclei. The red dashed line represents the reference mean $f_0(t)$ which corresponds to the short unaccented nucleus. The red solid line represents the effect of PHRASAL ACCENT $f_1(t)$ added to the reference mean, and corresponds to the short accented nucleus. The blue dashed line represents the condition where the covariate effect curve for PHONEMIC QUANTITY, $f_2(t)$, is added to the reference mean, which corresponds to the long unaccented nucleus. The blue solid line represents the effects of PHRASAL ACCENT, PHONEMIC QUANTITY and the interaction effect curve added to the reference mean and represents the long accented nucleus.

F2 Figure 3.14 shows the reference mean curve $f_0(t)$, covariate effect curves for PHRASAL ACCENT $f_1(t)$ and PHONEMIC QUANTITY $f_2(t)$, and the interaction effect curve $f_3(t)$ for F2 of the vocalic nucleus. The summed effect curves are shown in Figure 3.15. The reference mean $f_0(t)$ of F2 (Figure 3.14 (a)) rises slightly but overall remains almost flat until around time point 0.6. Then it declines rapidly towards the end. The covariate effect curve for PHRASAL ACCENT $f_1(t)$ (Figure 3.14 (b)) is significantly above zero throughout the entire nucleus. A higher F2 points to a more fronted tongue position, which for the accented condition is expected by the Hyperarticulation hypothesis. $f_1(t)$ has a raising slope, so the difference between the accented and unaccented condition for short nuclei is greater towards the end compared to at the beginning of the nuclei. When we look at the red dashed line (short, unaccented condition) and the red solid line (short, accented condition) in Figure 3.15, they are not entirely parallel. At the beginning they are closer together than at the end. The covariate effect curve for PHONEMIC QUANTITY $f_2(t)$ is also significantly above zero throughout the entire nucleus except for the start and end points (Figure 3.14 (c)). Thus we can see in Figure 3.15 that the F2 of long unaccented nuclei (blue dashed line) is higher than the F2 of unaccented short nuclei (red dashed line), but the effect is smaller at the beginning and end and greatest at around time point 0.8. The confidence band of $f_3(t)$ (Figure 3.14 (d)) overlaps with zero, so the interaction effect of PHRASAL ACCENT and PHONEMIC QUANTITY is not significant. This means that accent has the same effect on short and long nuclei. Consequently, the difference between the red lines and the difference between the blue lines are not significantly different.

A higher F2 indicates a more fronted tongue position. The articulatory data captured at the posterior angle (Figure 3.7) match the acoustic data partially. A significant fronting effect in the accented condition could be observed for short nuclei only. Long nuclei were more fronted than short nuclei regardless of the accent condition which also matches the acoustic data. However, for long nuclei the tongue was not more fronted in the accented condition compared to the unaccented condition, although F2 was higher in the accented condition for long nuclei. The articulatory data was obtained at only one angle, so not all the articulatory information reflected in the acoustics is captured.

The confidence bands for F2 are very wide compared to the confidence bands for F1. One reason might be that some speakers show a greater variation in the F2 dimension.



(a) (b)

(c) (d)

Figure 3.14: Reference mean and covariate effects for F2 in vocalic nuclei.

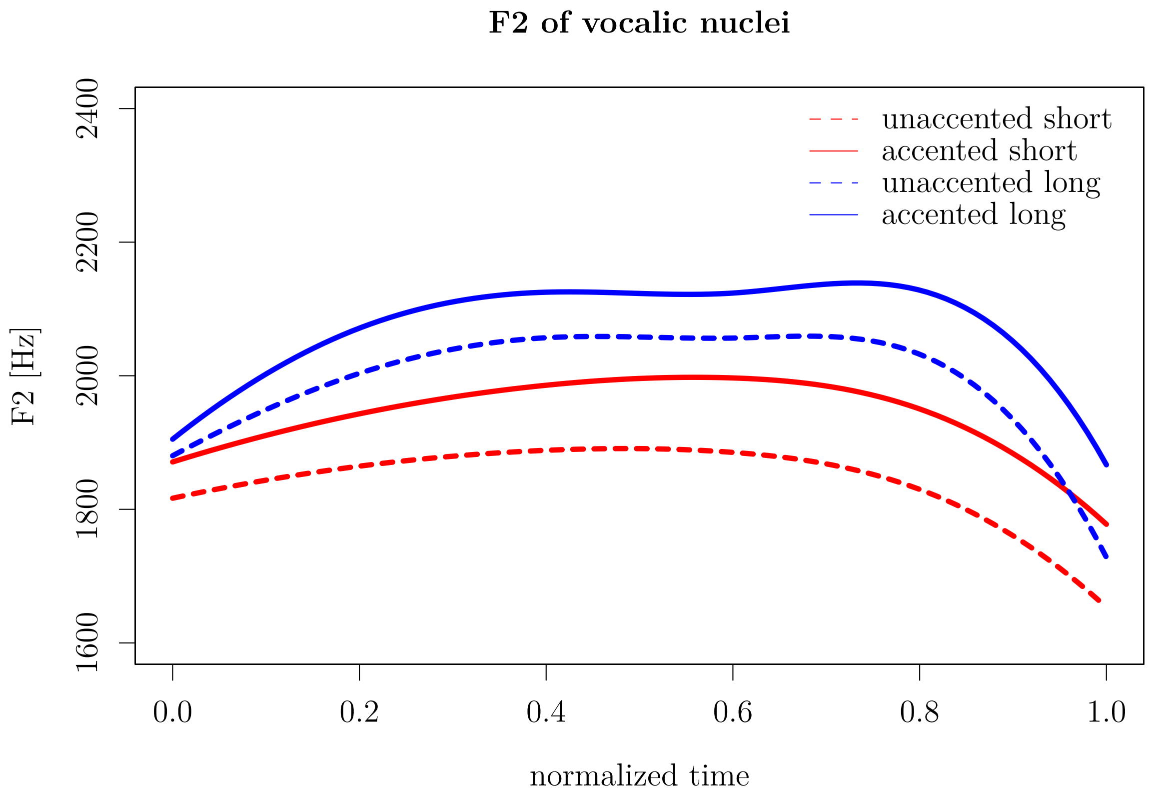


Figure 3.15: Summed effects curves for F2 in vocalic nuclei. The red dashed line corresponds to the reference mean $f_0(t)$ and corresponds to the F2 of the unaccented short nucleus. The red solid line corresponds to the accented short condition $(f_0(t) + f_1(t))$, the blue dashed line to the unaccented long condition $(f_0(t) + f_2(t))$ and the blue solid line to the accented long condition $(f_0(t) + f_1(t) + f_2(t) + f_3(t))$.

3.3.2 Consonantal nucleus

Results will be presented in the same order as for the vocalic nucleus, beginning with the results of the articulatory data, followed by the acoustic analysis. Recall that for the articulatory analysis we looked at tongue movement during the time normalized target sequence *plp* and *pĺp* of *plpap* and *pĺpap*. So we will briefly present how phrasal accent and phonemic quantity of the nucleus affected the duration of consonants, which we will refer to as *p1* and *p2*, and how the time normalized target sequence is divided into each segment for each condition.

Figure 3.16 shows the mean durations of the segments in the target sequence (top plot) and how the time normalized target sequence is proportionally divided into each segment for the different contexts (bottom plot). The difference compared to the sequences with vocalic nuclei is minimal and will be pointed out as we present the details.

Acoustic duration of the onset consonant *p1* was affected significantly by phrasal accent ($\chi^2[1] = 38.9, p < 0.001$). Phonemic quantity of the nucleus did not have a significant effect ($\chi^2[1] = 0.5, p > 0.1$). In contrast to *p1* preceding vocalic nuclei, the interaction of phrasal accent and phonemic quantity was not significant. The mean duration of *p1* in the accented condition was $160.3ms \pm 10.6$. In the unaccented condition, *p1* was shorter and the mean duration was $120.9ms \pm 10.6$. Except for the non-existent interaction effect, *p1* duration preceding consonantal nuclei did not differ substantially from *p1* duration preceding vocalic nuclei. The duration of *p1* in the accented condition preceding the consonantal nucleus was between the duration of *p1* preceding the accented long vocalic nucleus and the accented short vocalic nucleus, and likewise for the unaccented condition.

Acoustic duration of *p2*, the consonant following the target nucleus, was affected significantly by phrasal accent ($\chi^2[1] = 11.7, p < 0.001$). Phonemic quantity of the nucleus ($\chi^2[1] = 1.1, p > 0.1$) did not have a significant effect. Differing from *p2* following vocalic nuclei, the interaction of phrasal accent and phonemic quantity was significant ($\chi^2[1] = 6.0, p < 0.05$). Post-hoc Tukey tests revealed that the effect of phrasal accent on *p2* was greater following long nuclei. In the unaccented condition the acoustic duration of *p2* following a short nucleus and following a long nucleus did not differ, but in the accented condition *p2* following a long nucleus was longer than *p2* following a short nucleus. The mean duration of *p2* in the accented context after a long nucleus was $108.4ms \pm 7.52$ and

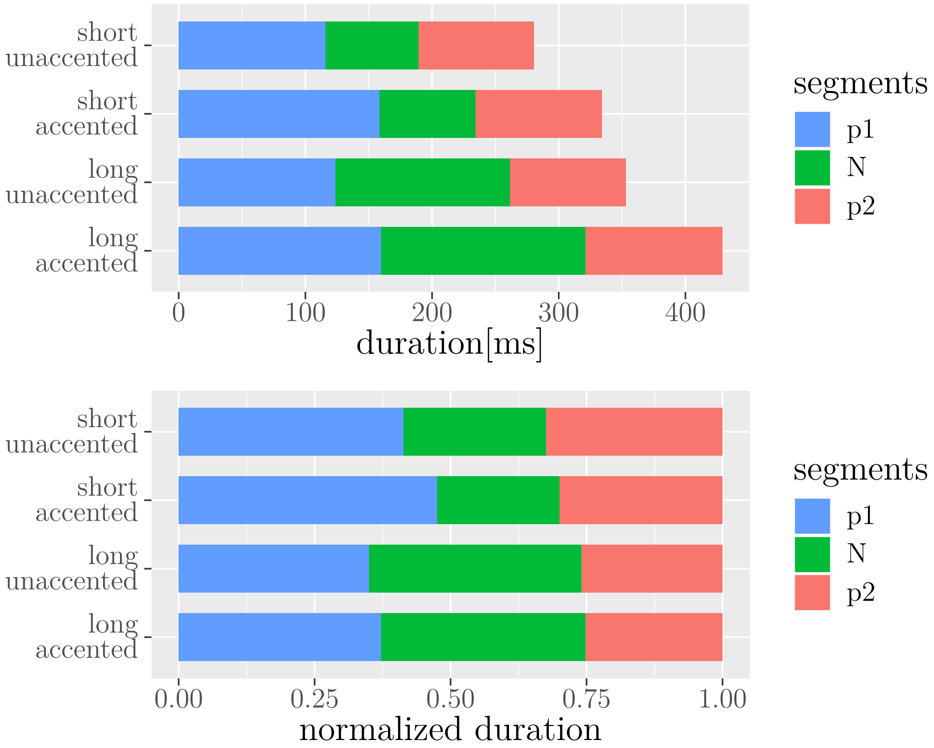


Figure 3.16: Absolute (top) and normalized (bottom) mean durations of segments in sequences with consonantal nuclei.

after a short nucleus it was $100.3ms \pm 7.2$. In unaccented context the acoustic duration of *p2* following a long nucleus was $92.0ms \pm 7.87$ and following a short nucleus it was $91.5ms \pm 8.0$. *p2* durations following consonantal nuclei are longer than following vocalic nuclei, which were $93.2ms \pm 5.9$ for the accented condition and $82.4ms \pm 5.9$.

When we compare the normalized duration of *p1* before short nuclei for the accented and unaccented condition in Figure 3.16 for consonantal nuclei with Figure 3.3 for vocalic nuclei, we see that the difference is slightly smaller when the nucleus is consonantal. Considering the segments within the time normalized sequence, as was the case for vocalic nuclei, the proportion of target sequences with long nuclei is not affected by phrasal accent. When the nucleus was short, the nucleus was proportionally longer in the unaccented condition, because the preceding and following consonants lengthened but not the nucleus.

Results of the ultrasound recordings of the consonantal nucleus

Posterior angle The data used for the tongue retraction analysis of the consonantal nucleus is visualized in Figure 3.17. A higher radius indicates a more retracted tongue. The contours start at a low radius, because the vowel preceding the target sequence is /i/ and the tongue is still fronted at the onset of the initial consonant. It is retracted for the dark /l/ in the nucleus, reflected in high radius values. At the end of the target sequence the tongue is fronted slightly towards the target for /a/, which is the vowel following the target sequence. For the posterior angle both the Sonority Enhancement Hypothesis as well as the Hyperarticulation Hypothesis predict a more retracted tongue in the accented condition. Thus, we expect a higher radius for the accented condition compared to the unaccented.

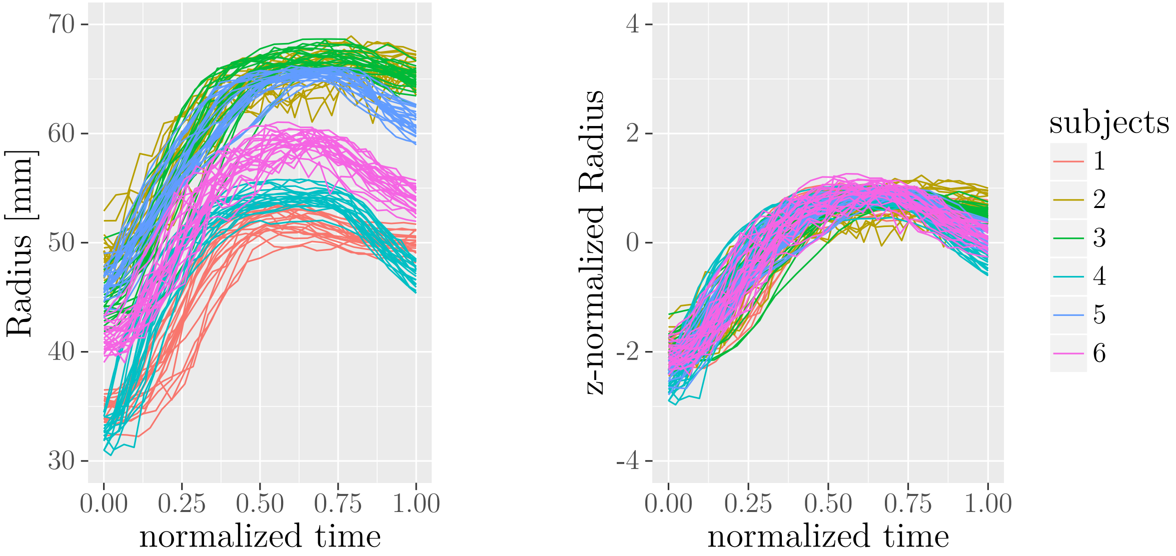
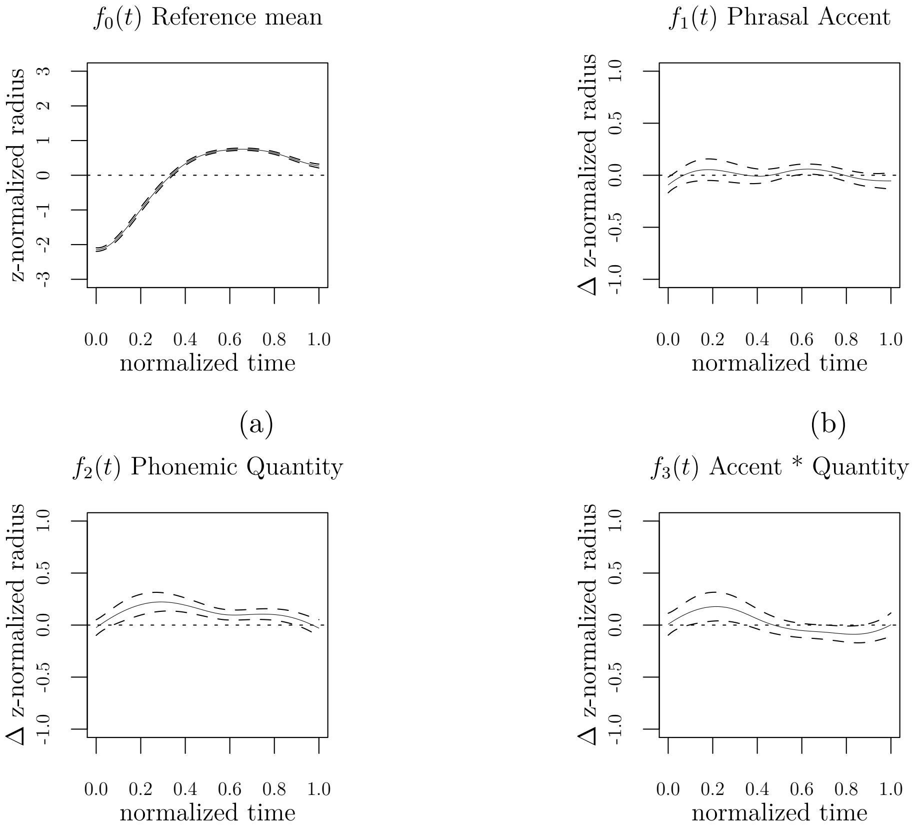


Figure 3.17: Radius measured at the posterior angle of target sequences with consonantal nuclei: *plp* or *pĺp*. On the left the radius over normalized time, on the right z-normalized radius over normalized time. A higher radius indicates a more retracted tongue.

The reference mean curve, the covariate effect curves for PHRASAL ACCENT and PHONEMIC LENGTH as well as their interaction effect curve are shown in Figure 3.18. Recall that the effects are significant when the confidence bands do not overlap with zero. The summed effect curves are shown in Figure 3.19 and in Figure 3.25 with the acoustically determined segment boundaries for each condition. The reference mean corresponds to the tongue retraction movement of the unaccented target sequence with phonemically short nucleus. The effect of PHRASAL ACCENT was not significant except for a trend level effect between around time point 0.6 and 0.7. During this time interval the lower edge of the confidence band of $f_1(t)$ (Figure 3.18 (b)) overlaps with zero. To see how the radius captured at the posterior angle changes over time for the accented short nucleus condition, the covariate effect curve for PHRASAL ACCENT $f_1(t)$ is added to the reference mean $f_0(t)$ In Figure 3.19 the red dashed line corresponds to the unaccented short condition and the red solid line to the accented short condition. The two lines are almost overlapping, but between around time point 0.6 and 0.7 the red solid line is slightly higher than the red dashed line. For the sequences with short nuclei, between these time points the tongue is slightly more retracted (trend level significance) in the accented condition compared to the unaccented condition. However, when we take in account the time normalization, our data suggests that the target position is held longer for the accented condition. Recall that although accent did not have an effect on the acoustic duration of short nuclei, the preceding and following /p/s were longer in accented condition. Thus the movement trajectory for the accented condition has to be viewed as being more compressed compared to the unaccented condition when time normalized.

The effect of PHONEMIC QUANTITY is significant throughout almost the entire target sequence. Only at the very beginning and end the effect is not significant. Thus, when the nucleus is long, the tongue is more retracted. When we compare the red and blue dashed lines in Figure 3.19, the blue dashed line, corresponding to the unaccented long nuclei raises faster. Recall that within the time normalized target sequence, *p1* is shorter when preceding long nuclei although its absolute duration was not affected by the phonemic quantity of the nucleus. So the retraction movement during *p1* for long nuclei is not faster than for short nuclei. The time interval for the accented condition is just more compressed when time normalized. Although the red and blue dashed lines at their maximum values between around 0.5 and 0.8 are very close, according to $f_2(t)$ the difference is significant. The interaction of PHRASAL ACCENT and PHONEMIC QUANTITY $f_3(t)$, shown in Figure 3.18 (d), is significant between around time point 0.1 and 0.3, and at trend level significant between time point 0.6 and 0.8. Beginning with the significant effect between time points 0.1 and 0.3, it results in the summed effect curve for the accented long nucleus condition to have a steeper rising slope at the beginning, compared to the unaccented long nucleus condition as can be seen in Figure 3.19 (blue solid line and blue dashed line). This is at least partly an effect of time normalization. Proportionally, *p1* duration in the two conditions is almost the same (Figure 3.16). However, the mean duration of *p1* in the unaccented condition is shorter than in the accented condition. The target for the retraction movement is thus reached earlier relative to the acoustically determined beginning of the nucleus in the accented condition, because there is more time during *p1*. The trend level significance between time points 0.6 and 0.8 goes in the opposite direction of the trend level significance of PHRASAL ACCENT $f_1(t)$ between time points 0.6 and 0.7. While the tongue was slightly more retracted in the accented short nucleus condition compared to the unaccented short condition, no difference can be observed for long nuclei.

In sum, the tongue was not more retracted as an effect of accent, but as an effect of phonemic quantity. The tongue was more retracted for long nuclei compared to short nuclei with the tendency for the difference being greater in the unaccented condition. However, in accented condition the target position was maintained for a longer period of time.



(a) (b)

(c) (d)

Figure 3.18: Reference mean and covariate effects of radius measured at the anterior angle in target sequences with consonantal nuclei. A greater radius indicates a more retracted tongue.

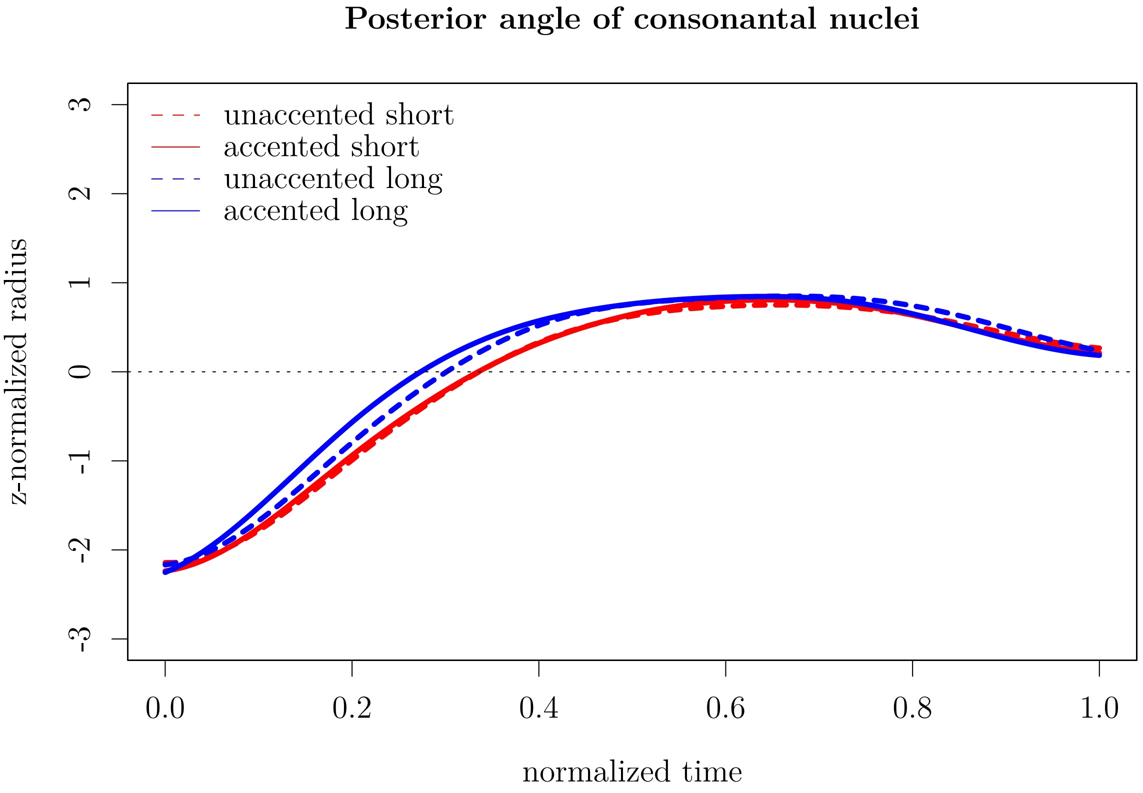


Figure 3.19: Summed effect curves of tongue movement captured at the posterior angle in target sequences with consonantal nuclei. A greater radius indicates a more retracted tongue. Red dashed line: unaccented short nucleus ($f_0(t)$); red solid line: accented short nucleus ($f_0(t) + f_1(t)$); blue dashed line: unaccented long nucleus ($f_0(t) + f_2(t)$); blue solid line: accented long nucleus ($f_0(t)f_1(t) + f_2(t) + f_3(t)$).

Anterior angle Figure 3.20 shows the data that entered the FLMM analysis of the tongue movement at the anterior angle of the target sequence with consonantal nucleus. It shows the radius of the tongue position measured at a predefined angle as a function over normalized time. For the analysis the z-normalized data plotted on the right was used. A higher radius indicates a higher tongue position. Recall that the anterior angle does not capture the apical constriction but a lowered tongue position behind the raised tongue tip (Figure 3.1). The sequence begins with a high tongue position because of the preceding high vowel /i/. The major movement takes place at the beginning, mainly during the initial /p/. The tongue moves from a high position of the preceding vowel /i/ to its lower position for /l/, at which the position is maintained, or in some cases the tongue is raised again, and then further slightly lowered towards the end for the following vowel /a/. For the anterior region, the Sonority Enhancement Hypothesis and the Hyperarticulation Hypothesis make contradicting predictions for the effect of phrasal accent. According to the Sonority Enhancement Hypothesis, the oral cavity should be widened in the accented condition, which is achieved by lowering the jaw and tongue. The Hyperarticulation Hypothesis predicts a stronger apical constriction when accented, for which a higher tongue position is required.

The reference mean, covariate effects of PHRASAL ACCENT and PHONEMIC LENGTH and their interaction effect are shown in Figure 3.21. The reference mean corresponds to the tongue position measured at the anterior radius for the unaccented condition in sequences with phonemically short nuclei. The covariate effect curves indicate how the reference mean is changed. The effect is significant when the confidence bands do not overlap with zero. The summed effect curves are retrieved by adding the covariate effect curves to the reference mean curve. PHRASAL ACCENT has a lowering effect from the beginning of the target sequence until around time point 0.6. This effect can also be seen when we compare the dashed red line corresponding to the reference mean and the solid red line corresponding to the accented condition of short nuclei in Figure 3.22. The red solid line has a steeper falling slope and reaches a lower position around time point 0.4 than the red dashed line. It is however not possible to tell whether the movement in the accented condition is actually faster, because of the normalized time. Between time points 0.4 and 0.6, the red dashed line keeps a relatively constant value but the red solid line rises, so that the difference between the two lines becomes smaller. Looking at Figure 3.16 or Figure 3.25 we can see that this

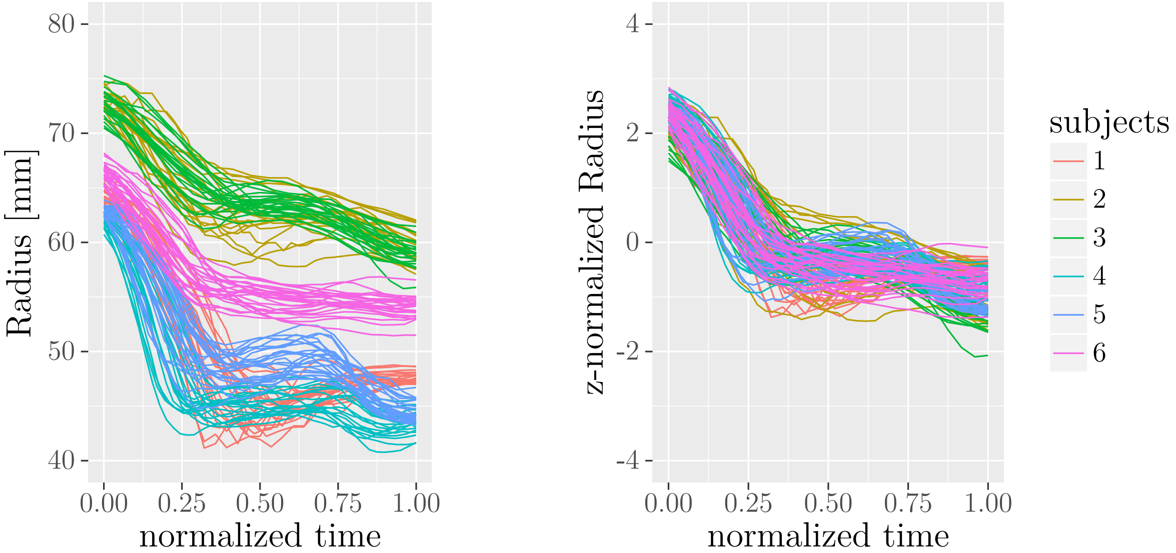
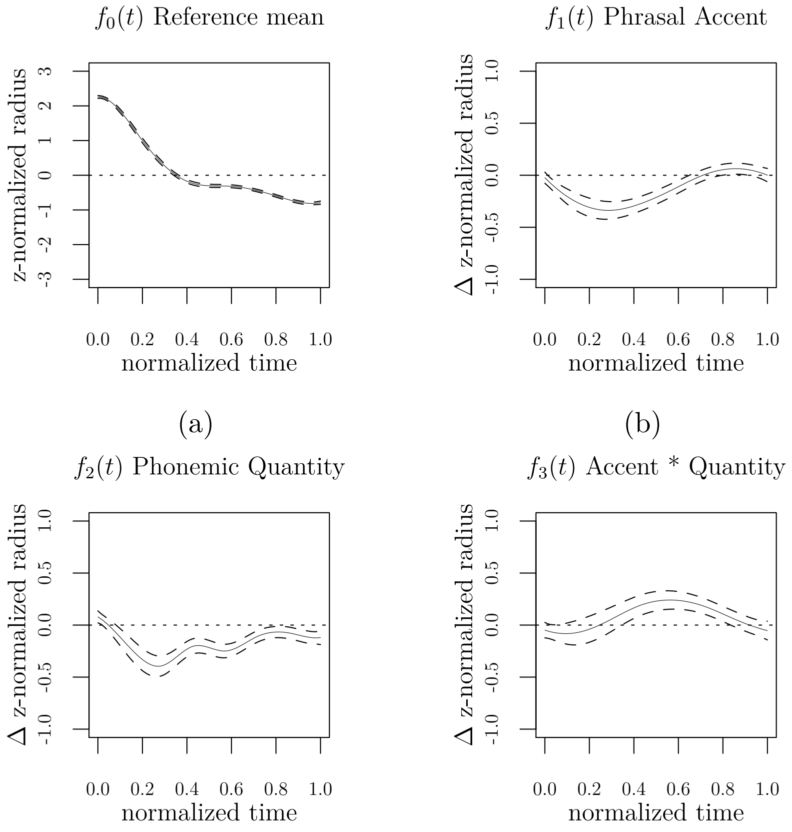


Figure 3.20: Radius measured at the anterior angle of target sequences with consonantal nuclei: *plp* or *pĺp*. On the left the radius over normalized time, on the right z-normalized radius. A greater radius indicates a higher tongue position.

is roughly speaking during the nucleus. Thus, at the beginning of the nucleus, the tongue is lower for the accented condition, but rises during the nucleus, so at the end it reaches the same position as for the unaccented condition. PHONEMIC QUANTITY has a lowering effect throughout the entire segment, except at the very beginning. Around time point 0.8 the lowering effect is significant only at trend level. The effect is easier to understand when we compare the blue and red dashed lines in Figure 3.22. The blue dashed line is lower than the red dashed line throughout the entire segment except for at the very beginning. The two lines begin at about the same position, but the blue dashed line has a steeper falling slope. The blue dashed line eventually reaches a lower position. Because of the time normalization it is not possible to tell whether the movement is faster in either condition. Between around time points 0.4 and 0.6 the two lines run almost parallel at a constant radius. After time point 0.6 the difference between the red and blue dashed lines becomes smaller but the red dashed line is still higher than the blue dashed line. The interaction between PHRASAL

ACCENT and PHONEMIC QUANTITY had a raising effect between time points 0.4 and 0.8. So while for the short nucleus condition the tongue was lower during the nucleus when accented compared to the unaccented condition, for the long nucleus condition the tongue was only slightly lower at the beginning of the acoustically determined nucleus but higher at the end of the nucleus in accented condition. Roughly speaking, based on mean durations of each segment, for long nuclei the nucleus is between time points 0.36 and 0.74 for the unaccented condition and between time points 0.37 and 0.75 for the accented condition, and for short nuclei between 0.41 and 0.67 in unaccented condition and between 0.48 and 0.71 for the accented condition (see also Figure 3.25).

To summarize, it is not possible to say whether the tongue was overall higher or lower in the accented condition compared to the unaccented condition. As an effect of accent in both, short and long nuclei, a raising movement during the acoustically determined /l/ could be observed. The tongue contours for the four conditions are plotted in Figure 3.23 for the beginning of the acoustically determined nucleus and in Figure 3.24 for the end of the acoustically determined nucleus. Each tongue contour was retrieved at the first and last ultrasound frame of the acoustically determined nucleus, respectively, for the same speaker and for each condition during the same utterance. The effects are minimal, but for long nuclei we can see that at the beginning of the acoustically determined nucleus (Figure 3.23) the tongue contours at the anterior angle for the accented and unaccented condition are closer together than at the end of the nucleus (Figure 3.24). In both cases the tongue position is higher for the accented condition compared to the unaccented condition. For this speaker, we can also see the upwards curled tongue tip for all conditions in both plots.



(a) (b)

(c) (d)

Figure 3.21: Reference mean and covariate effects of tongue movement captured at the anterior angle in target sequences with consonantal nuclei. A greater radius indicates a higher tongue position.

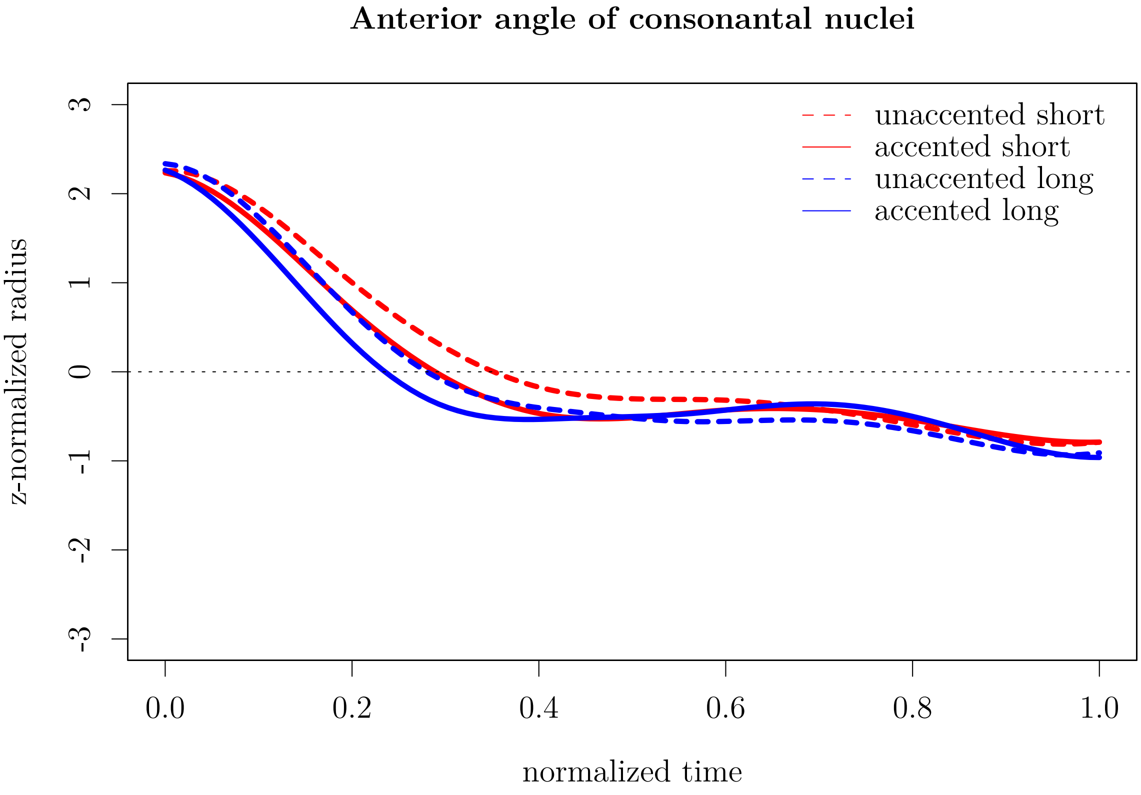


Figure 3.22: Summed effect curves of the tongue movement captured at the anterior angle in target sequences with consonantal nuclei. A greater radius indicates a higher tongue position. Red dashed line: unaccented short nucleus condition ($f_0(t)$); red solid line: accented short nucleus condition ($f_0(t) + f_1(t)$); blue dashed line: unaccented long nucleus condition ($f_0(t) + f_2(t)$); blue solid line: accented long nucleus condition ($f_0(t) + f_1(t) + f_2(t) + f_3(t)$).

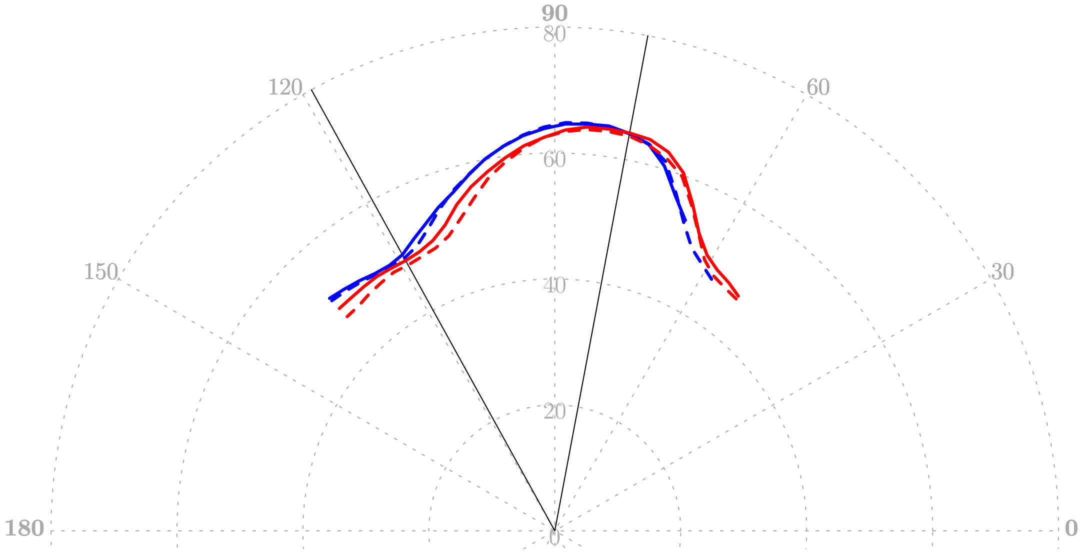


Figure 3.23: Tongue contours captured at the first frame within the acoustically determined consonantal nucleus for Speaker 5, each during the 3rd repetition. The front is to the left. The two angles at which the radius was retrieved are drawn using solid black lines.

Red dashed: Unaccented short condition

Red solid: Accented short condition

Blue dashed: Unaccented long condition

Blue solid: Accented long condition

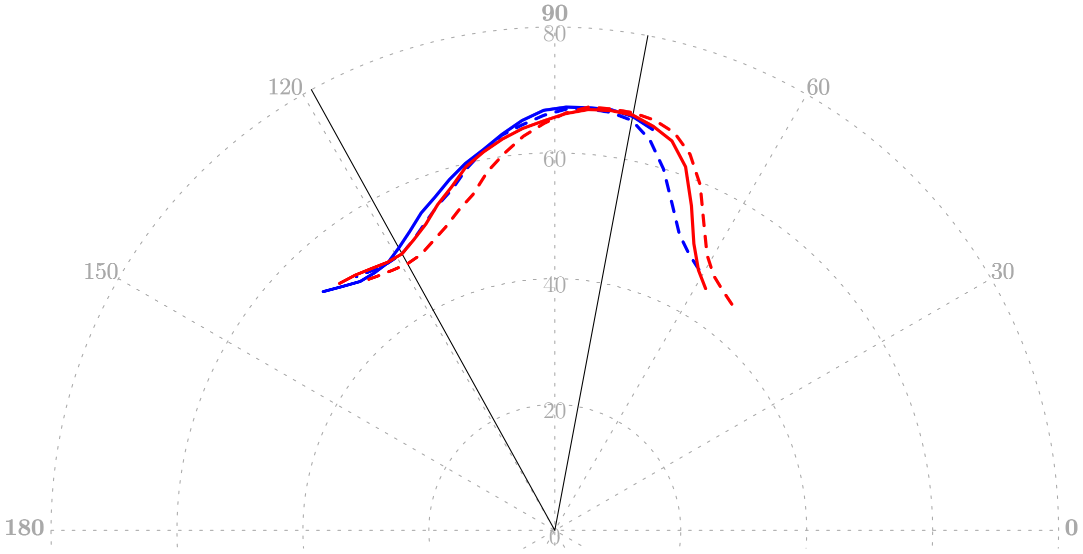


Figure 3.24: Tongue contour captured at the last frame within the acoustically determined consonantal nucleus for Speaker 5, each during the 3rd repetition. The front is to the left. The two angles at which the radius was retrieved are drawn using solid black lines.

Red dashed: Unaccented short condition

Red solid: Accented short condition

Blue dashed: Unaccented long condition

Blue solid: Accented long condition

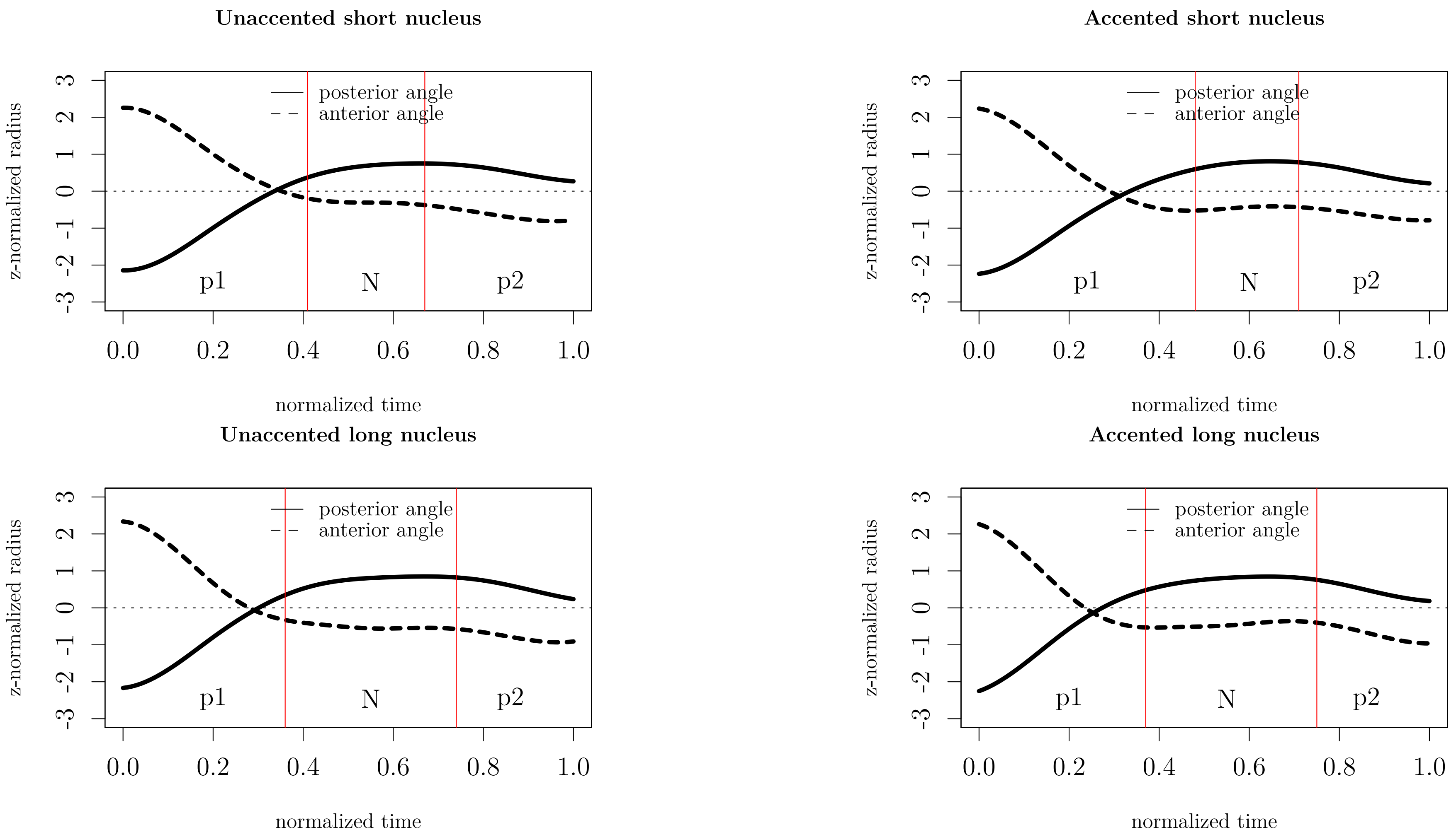


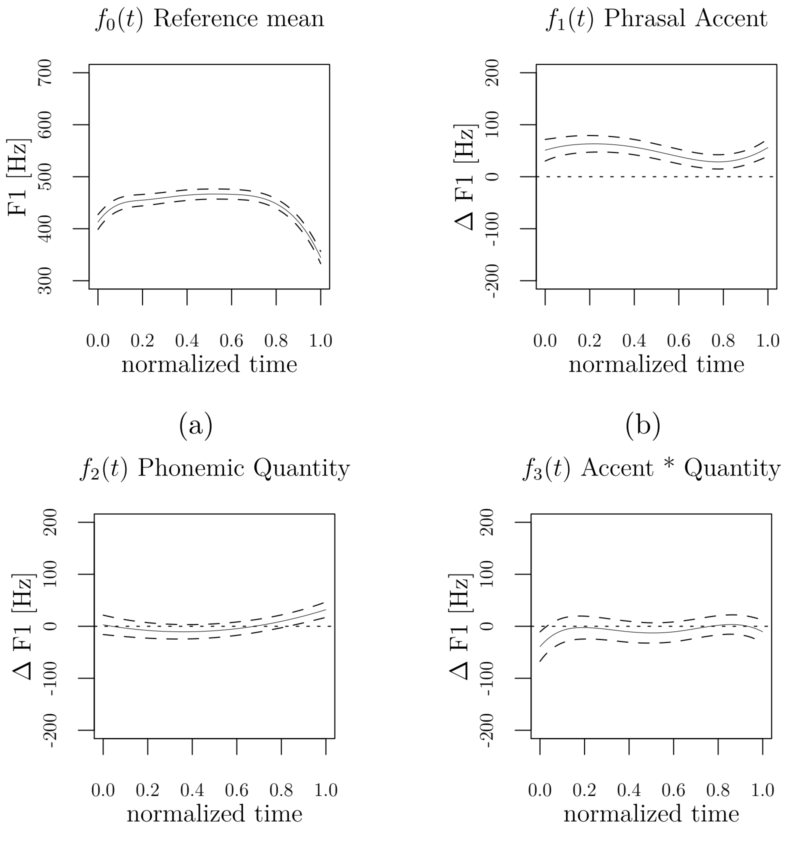
Figure 3.25: Summed effect curves of the tongue movement captured during sequences with consonantal nuclei at posterior and anterior radii separately for each nucleus condition. The red vertical lines represent the boundaries between the acoustically determined segments *p1* (onset consonant /p/), *N* (consonantal nucleus /l/), *p2* (/p/ following the nucleus). A lower radius at the anterior angle indicates tongue lowering, a lower radius at the posterior angle indicates tongue fronting.

Results of the formant analysis for the consonantal nucleus

The formant analysis was carried out during the acoustically segmented nucleus /l/. The time interval differs from the articulatory analysis which was carried out from the beginning of the initial /p/ until the end of the /p/ following the nucleus, but the corresponding acoustically determined segments are indicated in Figure 3.25. FLMM for F1 and F2 was carried out and the model consisted of covariate effects PHRASAL ACCENT and PHONEMIC QUANTITY and their interaction. A dark /l/ generally has a higher F1 and lower F2 than a light /l/ (Sproat & Fujimura, 1993) and according to Lin et al. (2014) a weaker anterior constriction correlates with a smaller difference between F1 and F2. We can also assume that F1 reflects the vertical position of the tongue and F2 the horizontal position (although less consistently). A higher tongue position in the front region, so a stronger anterior constriction should result in a lower F1. A more retracted tongue should result in a lower F2. The Sonority Enhancement Hypothesis predicts a weaker apical constriction and a more retracted tongue back in accented condition. Acoustically, this should be reflected in a higher F1 and a lower F2 in the accented condition. The Hyperarticulation Hypothesis also predicts a more retracted tongue back in the accented condition, but at the same time also a stronger apical constriction. Although both hypotheses predict a more retracted tongue for the accented condition, the articulatory data did not capture a more retracted tongue as an effect of accent. However, the tongue was slightly more retracted for long nuclei compared to short nuclei. At the anterior angle we did not capture the constriction but the region behind the raised tongue tip. In the accented condition a slightly rising movement could be observed. For short nuclei, the tongue was lower in the accented condition compared to the unaccented condition at the beginning of the nucleus, but towards the end the tongue positions were about the same height. The tongue position of long nuclei in accented condition began at about the same position or slightly lower as in the unaccented condition and was higher towards the end of the nucleus. According to the articulatory data F1 should show more change over time for the accented condition. For F2, differences as an effect of phonemic length rather than as an effect of accent are expected.

F1 The reference mean curve, corresponding to the F1 of the short consonantal nucleus /l/, covariate effect curves of PHRASAL ACCENT, PHONEMIC LENGTH, and the interaction effect curve are shown in Figure 3.26. The summed effect curves are shown in Figure 3.27. The reference mean $f_0(t)$ corresponding to the unaccented short /l/ is relatively flat with falling transition at edges, characteristic of the low loci of the bilabial consonants /p/ preceding and following the nucleus (Delattre et al., 1955). Phrasal accent had a significant effect of raising F1, with the effect being larger at the beginning. The covariate effect curve with its confidence band for phrasal accent $f_1(t)$ in Figure 3.26 (b) is above zero throughout the entire segment. When comparing the red dashed line corresponding to the unaccented short nucleus and the red solid line corresponding to the accented short nucleus in Figure 3.27 the solid line is higher, but the difference between the two lines becomes smaller towards the end. Between time points 0.1 and 0.7 the red dashed line is almost parallel to the x-axis and rises minimally. The red solid line however declines minimally. There was no significant effect of phonemic quantity except for at the very end of the nucleus, from about timepoint 0.9 to 1 where the covariate effect curve with its confidence bands $f_2(t)$ is above zero (Figure 3.26 (c)). The red and blue dashed lines in Figure 3.27, corresponding to the unaccented short and long nucleus, respectively, are not significantly different, except for at the very end. The red dashed line starts declining slightly earlier and reaches a lower value. The interaction effect curve $f_3(t)$ in Figure 3.26 (d) indicates that there was no significant interaction effect. This means that the effect of accent on short nuclei is not significantly different from the effect of accent on phonemically long nuclei. Although the solid blue line is lower than the solid red line, the distance between the red dashed line and red solid line and the distance between the blue dashed line and the blue solid line are not significantly different.

Overall, acoustic and articulatory results do not go well together. For both phonemically short and long nuclei, F1 is higher in the accented condition, indicating a wider oral cavity as predicted by the Sonority Enhancement hypothesis. The articulatory data captured at the anterior angle showed a lower tongue position when accented only for short nuclei before time point 0.6 (Figure 3.22). We can assume that the jaw is lower in accented conditions while the tongue moves in the opposite direction to reach the target for /l/. The slightly falling F1 throughout the nucleus in the accented condition might reflect the rising tongue position captured at the anterior angle for the accented condition (Figure 3.22 red solid line roughly between time points 0.4 and 0.7 and blue solid line between time points 0.4 and 0.7).



(a) (b)

(c) (d)

Figure 3.26: Reference mean and covariate effects for F1 in consonantal nuclei.

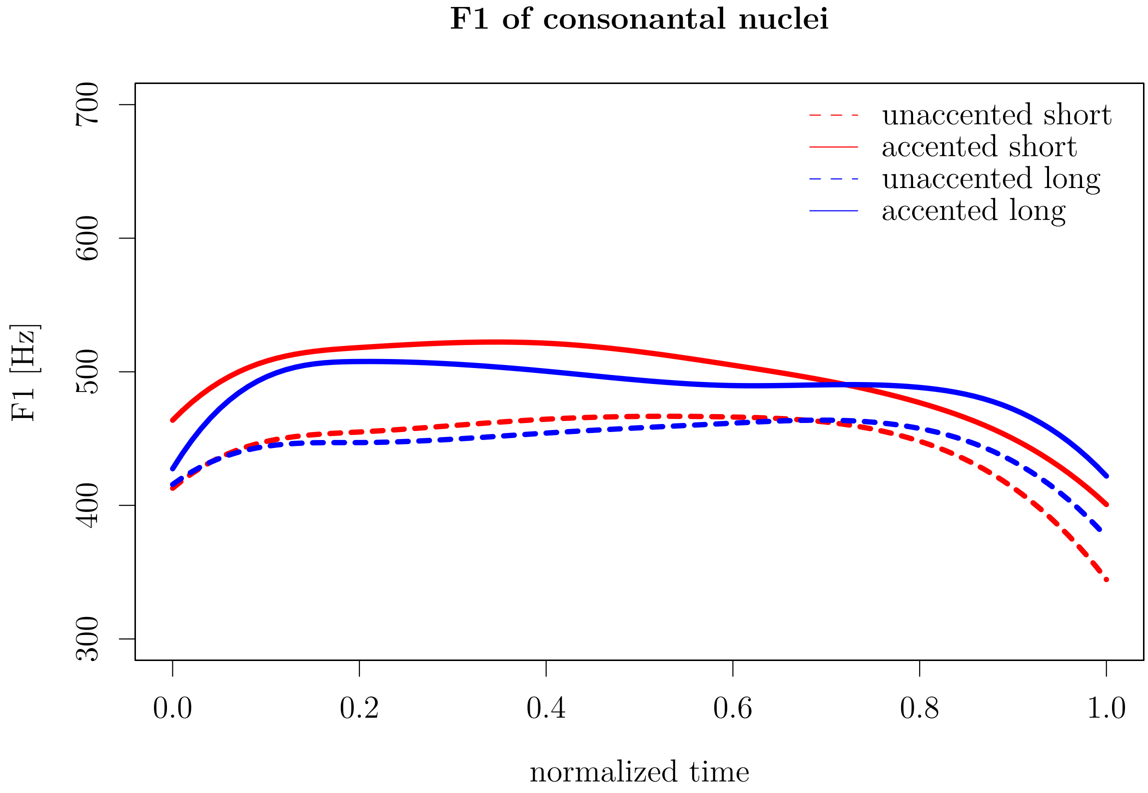
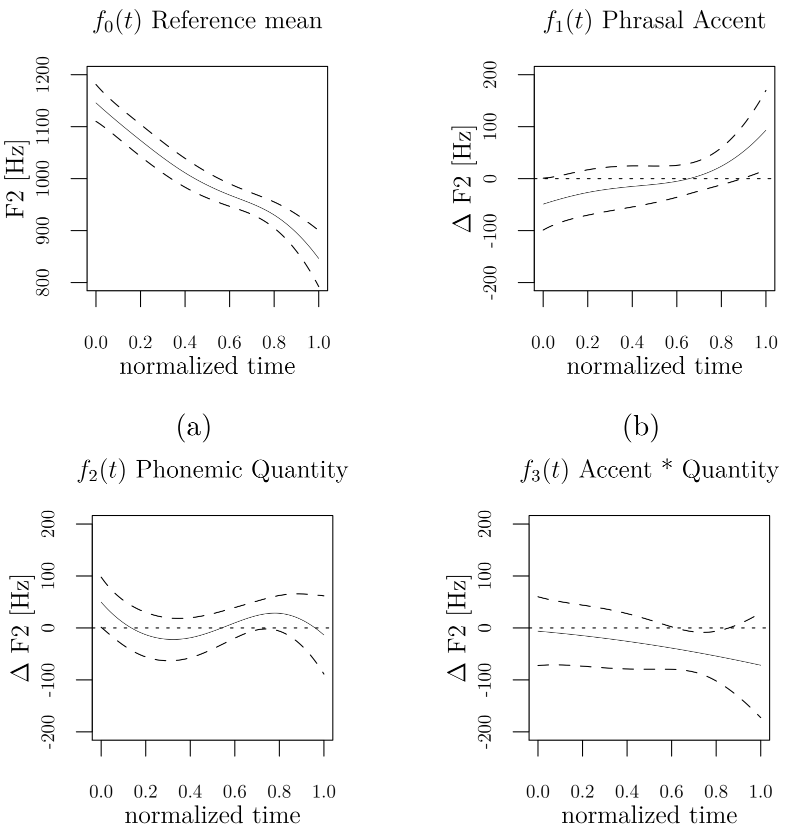


Figure 3.27: Summed effects curves for F1 in consonantal nuclei. Red dashed line: unaccented short nucleus condition ($f_0(t)$); red solid line: accented short nucleus condition ($f_0(t) + f_1(t)$); blue dashed line: unaccented long nucleus condition ($f_0(t) + f_2(t)$); blue solid line: accented long nucleus condition ($f_0(t) + f_1(t) + f_2(t) + f_3(t)$).

F2 The results of the FLMM analysis are shown in Figure 3.28. Summed effect curves are visualized in Figure 3.29. The reference mean curve $f_0(t)$ corresponding to the F2 of the short unaccented consonantal nucleus is a steadily falling slope. A lower F2 indicates a more retracted tongue position, so it suggests that during the /l/ segment a retraction gesture takes place, and the tongue is more retracted at the end of the nucleus. The covariate effect curve for PHRASAL ACCENT has a positive slope but the confidence band overlaps with zero except just before the very end after time point 0.9. So in general PHRASAL ACCENT did not have a significant effect on F2. Because of the positive slope of $f_1(t)$ we can see in Figure 3.29 that the solid red line corresponding to the accented condition is less steep than the dashed red line corresponding to the unaccented condition. PHONEMIC QUANTITY $f_2(t)$ does not have a significant effect on F2 either, although it slightly changes the curve dynamic. The blue dashed line in Figure 3.29 is not a linearly falling slope. Between around time point 0.5 and 0.7 a constant F2 value is maintained. The interaction effect curve $f_3(t)$ shows a significant effect only slightly before time point 0.8, where the interaction effect curve with the confidence bands are below zero. $f_3(t)$ has a negative slope, so its effect goes into the contrary direction to that of $f_1(t)$. Recall that to get the summed effect curve for the long accented nucleus, $f_1(t)$, $f_2(t)$ and $f_3(t)$ have to be added to the reference mean $f_0(t)$. In Figure 3.29 we can see that the blue dashed and solid lines corresponding to the long nuclei run almost parallel until time point 0.8 in contrast to the two red lines.

In sum, F2 for the four conditions did not differ significantly. However, the shape of the F2 curves for short and long nuclei differed, in that for short nuclei the curves were almost linear while for long nuclei it was s-shaped. The acoustic data captured at the posterior angle also did not show significant differences as an effect of phrasal accent. However, long nuclei were more retracted in the accented condition which does not go with the acoustic data.



(a) (b)

(c) (d)

Figure 3.28: Reference mean and covariate effects for F2 in consonantal nuclei.

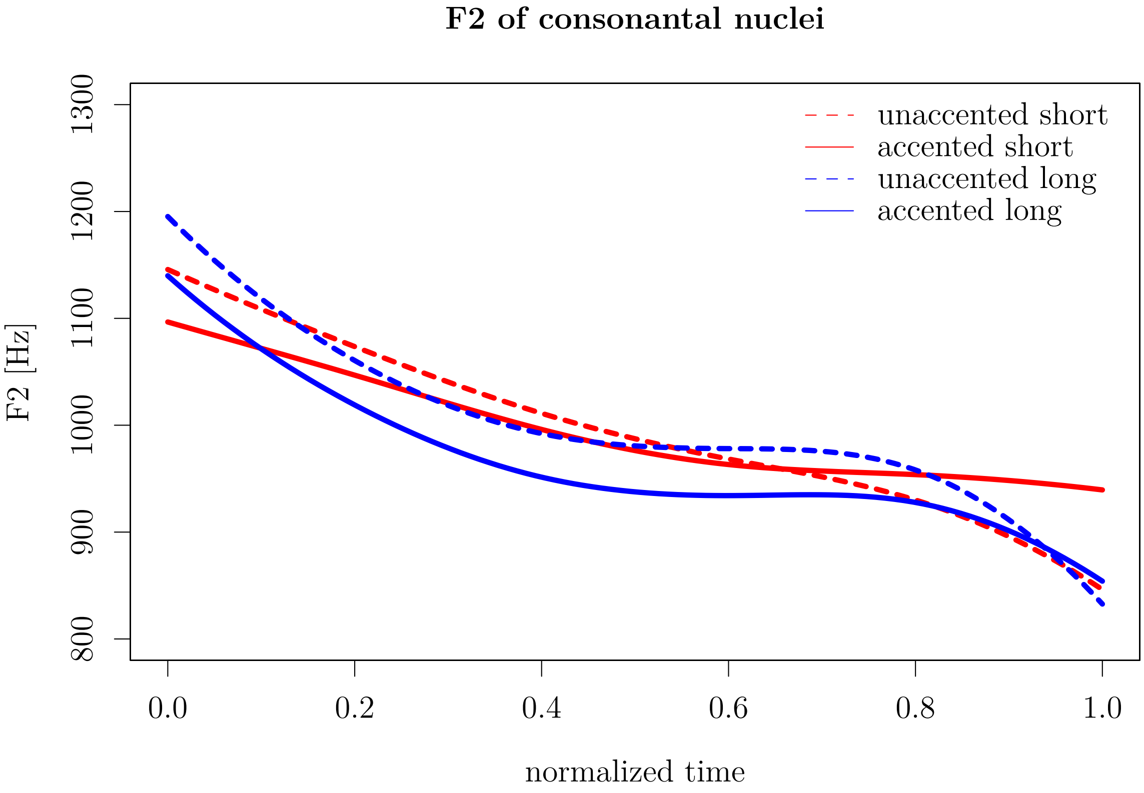


Figure 3.29: Summed effects curves for F2 in consonantal nuclei. Red dashed line: unaccented short nucleus condition ($f_0(t)$); red solid line: accented short nucleus condition ($f_0(t) + f_1(t)$); blue dashed line: unaccented long nucleus condition ($f_0(t) + f_2(t)$); blue solid line: accented long nucleus condition ($f_0(t) + f_1(t) + f_2(t) + f_3(t)$).

3.4 Discussion

Our aim was to investigate how prosodic emphasis is implemented on consonantal and vocalic nuclei, specifically taking into account the predictions made by the Hyperarticulation Hypothesis and the Sonority Enhancement Hypothesis. Therefore we looked at the effect of phrasal accent and phonemic quantity on the articulation and acoustics of vocalic (/e/) and consonantal (/l/) nuclei. The Sonority Enhancement Hypothesis predicts a wider oral cavity in accented contexts, which is achieved by lowering the tongue and jaw, regardless of the segment occupying the nucleus. Acoustically, higher sonority should correlate with a higher F1. The Hyperarticulation Hypothesis predicts enhancement of the segmental features under accentuation. /e/, a mid front vowel, is thus expected to be more fronted, which should acoustically be reflected in a higher F2. Regarding the vertical position of the tongue we suggested that according to the Hyperarticulation Hypothesis no change in height should be observed, as it is a mid vowel. According to the Hyperarticulation Hypothesis /l/ is expected to have a stronger apical constriction and a more retracted tongue back in accented condition. On the one hand, we were interested in these contradictory predictions made by the two hypotheses regarding the anterior region of the oral cavity. We suggested that these contradicting requirements make syllabic consonants less preferred particularly in positions which need to be prosodically emphasized. On the other hand, we also predicted that the 'vocalic' dorsal gesture might function as the carrier of prosodic emphasis.

For /e/, predictions made by the Sonority Enhancement Hypothesis were observed for F1, which was higher in the accented condition and also for long nuclei. The tongue position captured at the anterior angle was not lower, so at this point we can only assume that it was the jaw that contributed to the higher F1. The tongue was significantly higher towards the end of the nucleus for the accented condition, which was not predicted by either of the hypotheses. It is possible that the tongue was raised to compensate for the lower jaw position, or was raised together with the closure movement of the jaw. The Hyperarticulation Hypothesis also predicted a more fronted tongue position and higher F2 in accented condition, which was confirmed. But taking into account that the tongue was already fronted for /i/, which preceded the target sequence, it is interesting that the tongue was more retracted in the unaccented condition and that especially for long nuclei a fronting movement was observed during the nucleus. The more dynamic movement under prosodic emphasis, both in the horizontal as well as vertical direction, can also be interpreted as a consequence of the trough effect which takes place during the initial /p/. The trough effect describes a tongue repositioning or relaxation in a VCV sequence with identical vowels and a consonant which does not require a tongue displacement, for example a bilabial plosive (Fuchs, Hoole, Brunner, & Inoue, 2004; Lindblom, Sussman, Modarresi, & Burlingame, 2002; Vazquez-Alvarez & Hewlett, 2007). In our case the vowels are not identical, but both require an extremely fronted and rather high tongue position, providing an environment enabling the observation of a trough effect. One interpretation of the trough effect given by Lindblom et al. (2002) based on the coproduction model by Fowler and Saltzman (1993) is that it is a result of less overlap of the vowel gestures preceding and following the medial consonant. In prosodically strong conditions, segments are usually less coarticulated, and it can be assumed that the gestures are farther apart. So when the two vowel gestures are less overlapped, the tongue moves away from the target between these two gestures.

For /l/, predictions made by the Sonority Enhancement Hypothesis were observed for F1, but unlike for /e/ a higher F1 could be observed as an effect of phrasal accent only. For /e/ there was additionally a quantity effect, so F1 was higher for long nuclei than for short nuclei. For phonemically short consonantal nuclei the tongue position captured at the anterior angle was lower in the accented condition but only at the beginning of the acoustically segmented nucleus, before time point 0.6. Towards the end the height of the anterior part of the tongue did not differ for the two accent conditions. Unaccented phonemically long nuclei had a lower tongue position than unaccented phonemically short nuclei, but phrasal accent had the effect of raising the anterior tongue position of long nuclei. The anterior part of the tongue of long nuclei was about the same height for the two accent conditions at the beginning of the acoustically segmented nucleus, but towards the end of the nucleus it was higher for the accented condition than for the unaccented condition. What was characteristic for the accented condition in short and long nuclei was the upwards movement of the tongue in the anterior region during the nucleus. This is similar to the movement captured for the vocalic nucleus, but a trough effect cannot be expected. For /l/, a complete reorganization of the tongue shape is required. Possibly, the tongue in the anterior region is first lowered to then allow for the dorsal retraction and apical raising movement. The upwards movement in the accented condition suggests a stronger apical constriction, as predicted by the Hyperarticulation Hypothesis. At the same time the more dynamic movement can be interpreted as prerequisite to allow for a more curled tongue shape and greater lateral opening which enhances sonority.

Both hypotheses predicted a more retracted tongue dorsum for /l/ in the accented condition, which was not confirmed in our data. We hypothesized that /l/ can occupy the nucleus position because it can carry prosody due to the vocalic dorsal gesture. While in the accented condition the retracted tongue position was maintained for a longer time period, the effect of accent was mainly visible in the anterior region.

According to Sproat and Fujimura (1993), the difference between F1 and F2 is an indicator of the degree of darkness of /l/. To be more precise, Lin et al. (2014, p. 26) suggest that “all else being equal, alveolar reduction might be expected to result in a less forward tongue body position and/or greater influence of the dorsal constriction on the acoustic spectrum, either of which could contribute to a lower F2 frequency and increased proximity between F2 and F1”. Our acoustic data do indeed show a smaller distance in the accented condition, but this was mainly due to a higher F1. At the same time the tongue in the anterior region was higher when accented for long nuclei. Recall that we did not measure the apical constriction but the region behind the raised tongue tip, because in many cases the tongue tip was not visible when raised. A generally higher tongue position should still result in a more narrow oral cavity and consequently a lower F1. An explanation for the higher F1 in accented condition could be a lower jaw in accented condition, analogous to the observation for the vocalic nucleus, causing an oral cavity expansion and allowing a greater lateral airflow. Beňuš and Pouplier (2011) examined the jaw position for accented consonantal and vocalic nuclei and showed that the jaw opening of /l/ and /r/ was comparable to vowels. Their results also showed that for /e/ the jaw was lower for long nuclei than for short nuclei, while for /l/ there was no difference. Our F1 measurements corroborate the jaw displacement pattern observed by Beňuš and Pouplier (2011).

To summarize, we have shown that in contrast to the vowel, for which the entire tongue was affected by phrasal accent, prosodic emphasis of the consonantal nucleus was mainly carried by the anterior part of the tongue and not by the dorsal gesture, contrary to our expectations. The vocalic dorsal gesture was clearly present and its duration but not its magnitude was affected by phrasal accent. It seems that for /l/ prosodic emphasis is signalled using F1 and possibly a lower jaw position, and the greater movement of the anterior part of the tongue, but not the vocalic dorsal gesture. A possible explanation for the missing effect of phrasal accent on the magnitude of the dorsal gesture of /l/ is that in the nucleus position the tongue is already so far retracted leaving no room for further retraction. In Figure 3.25, the posterior tongue movement during the acoustical segmented nucleus is flatter for long nuclei than for short nuclei, indicating that the tongue is maintained at the retracted position for a longer duration. However, this claim has to be further investigated by comparing /l/ in nucleus condition to onset and coda conditions. The main difference between /e/ and /l/ is that for /e/ the entire tongue participates in prosodic variation while for /l/ the variation is limited to the anterior part of the tongue.

In the following chapter we will set the syllabic /l/ in relation to /l/ in onset and coda conditions.

Chapter 4

/l/ in different syllable positions

4.1 Introduction

In previous chapters we compared the effect of phrasal accent and phonemic quantity on the nucleus depending on whether it was occupied by a vowel (/e/) or a consonant (/l/). In this chapter, we will compare /l/ in nucleus position with /l/ in onset and coda position in order to investigate how it is affected by phrasal accent depending on which position it occupies within the syllable. As mentioned in Chapter 1, the reason why languages preferably alternate vowels and consonants, especially obstruents, might be to optimize the modulation of the speech signal. The rise and fall of sonority makes it easier to separate a continuous speech signal into meaningful segments. Just like the rising and falling F0 and systematic variation of vowel quality in the nucleus can signal prominence, consonants have been shown to vary systematically depending on syllable position and vicinity to boundaries to signal the position and strength of boundaries. These observations suggest that indeed vowels and consonants have different functions but at the same time the functional differences can be related to syllable positions. In the previous chapters we showed that /l/ in nucleus position can carry prosody. The high F0 is realized on accented syllabic consonants accompanied by the same durational differences as vocalic syllables. Like for vocalic nuclei, the acoustic and articulatory characteristics of consonantal nuclei are affected by phrasal accent. In this chapter we want to investigate whether /l/ in onset and coda position behaves differently compared to when occupying the nucleus position, given the functional differences associated with segments occupying different syllable positions.

4.1.1 Syllable position

It is widely acknowledged that articulatory gestures pattern differently depending on whether they occupy the onset or coda of a syllable, affecting both the intra- and inter-gestural organization (Krakow, 1999). According to Krakow (1999) the intra-gestural organization describes the strength or degree of displacement of a gesture. In general, consonants in final position show a weaker constriction than in initial position (Browman & Goldstein, 1992). This pattern also applies for the apical constriction in laterals, but at the same time it has been observed that the dorsal retraction gesture is more pronounced in coda position (Browman & Goldstein, 1995; Giles & Moll, 1975; Sproat & Fujimura, 1993), resulting in allophonic varieties of /l/ which depend on syllable position. In many varieties of English, the syllable onset is occupied by a clearer /l/, while the coda is occupied by a darker /l/, which can be further vocalized in some contexts (Scobbie & Pouplier, 2010). However, this kind of variation of /l/ is not universal. In Albanian, dark and light /l/ are two distinct phonemes and both varieties can be found word initially as well (cf. Moosmüller, Schmid, and Kasess (2016)). Recasens (2012) compared F2 values of /l/ in onset and coda positions and different vowel contexts for various languages. F2 is sensitive to the dorsal retraction gesture and lower F2 indicates a more retracted tongue (Stevens, 1997), thus a darker /l/. Recasens (2012) showed that languages classified as having a clear /l/ usually show more variation depending on vowel context and syllable position than languages classified as having a dark /l/. Furthermore, the degree of darkness or clearness is also language dependent; some languages have just a moderately clear or moderately dark /l/ compared to other languages. In all languages a trend towards F2 being higher in initial than in final position, thus /l/ being relatively darker in final compared to initial position, was observed. However, in most languages this difference was less than 300 Hz. Only few languages, including American English, exhibited greater differences. Slovak /l/ has been described as dark in all positions. To our knowledge no acoustic analysis exists and the articulatory data is limited to the kinematic analysis of the apical gesture and the temporal coordination (Pouplier & Beňuš, 2011). Pouplier and Beňuš (2011) presented some preliminary results involving the dorsal gesture which were followed up by Beňuš et al. (2014). While we do not know whether the apical constriction is stronger in the onset than in coda, we know that they differ in terms of stiffness and time to peak velocity. /l/ occupying the nucleus position patterns mainly with /l/ occupying the coda position. The dark and light varieties of American English /l/ have been described to differ also in the temporal organization of the apical and dorsal gesture. The inter-gestural organization describes how gestures are timed in relation to each other. Browman and Goldstein (1995) and Sproat and Fujimura (1993) both observed that in syllable onset position the apical and dorsal gesture of /l/ are timed almost synchronously, while in syllable coda the dorsal gesture precedes the apical gesture. Analogously, for nasal consonants it has been observed that the velic gesture begins earlier than the oral gesture in coda position while in onset position the two gestures are closer together (Krakow, 1999). Gick, Campbell, Oh, and Tamburri-Watt (2006) investigated the timing relation between the tongue tip gesture and the dorsal gesture for liquids in different languages but did not find a general pattern. It should be noted, however, that for each language only one or two speakers per language were analysed. Furthermore, depending on the language, the analysed liquid was either /l/, /r/ or a palatal-alveolar lateral. No language observed showed the exact opposite pattern to that observed by Sproat and Fujimura (1993) and Browman and Goldstein (1995) for American English which would be the dorsal gesture preceding the apical gesture in onset position and the dorsal gesture following the apical gesture in coda. However, in Serbo-Croatian the apical and dorsal gesture occurred simultaneously in initial, final and intervocalic positions. The magnitude of the two gestures was not compared in the study by Gick et al. (2006) and although averaged movement trajectories are shown, it seems that with the measure they used, it is not possible to compare the results for different syllable positions. Serbo-Croatian has been described as having a dark /l/ in all positions. This observation suggests that the temporal coordination of the two gestures does not necessarily differ for different syllable positions as suggested by Krakow (1999). The apparently systematic temporal organization of gestures is not limited to gestures within one segment but also between segments. Within the framework of articulatory phonology (Browman & Goldstein, 1989, 1992), a coupled oscillator model predicting the timing patterns of gestures has been developed (Nam et al., 2009). It assumes that a pair of gestures is coupled either in-phase or anti-phase with each other, which means that gestures are initiated either synchronously or sequentially. The model predicts that generally, pairs of consonant gestures are coupled anti-phase with each other. The first consonant in the coda is also coupled anti-phase with the preceding vowel. However, consonants that occupy the syllable onset are each coupled in-phase with the vowel. In case of onset clusters this would result in these consonants being timed synchronously, making them perceptually unrecoverable. Instead, a trading relationship between the anti-phase coupling of the consonants and the in-phase coupling of each consonant with the vowel emerges. It is the midpoint of the entire onset cluster that shows a stable alignment with the vowel. The effect is referred to as the c-center effect (Browman & Goldstein, 1988; Honorof & Browman, 1995; Marin & Pouplier, 2010). Consequently, when consonants are added to the onset, the acoustic duration of the vowel shortens as it is overlapped more by the onset cluster, while adding consonants to the coda should not affect the vowel duration. Temporal coordination has since been used to account for syllabification in several languages when syllable affiliation of consonant sequences was controversial, as for example in Italian (Hermes, Mücke, & Grice, 2013), Moroccan Arabic (Shaw, Gafos, Hoole, & Zeroual, 2009), Georgian and Tashlhiyt Berber (Goldstein, Chitoran, & Selkirk, 2007). However, the c-center effect for onset clusters and a sequential ordering in coda is not universal. For example, Byrd (1995) observed a c-center effect for coda and Marin (2013) found that in Romanian not all onset clusters comply to the c-center organization. Slovak also belongs to the languages which do not show a c-center organization for onset clusters (Pouplier & Beňuš, 2011). Onset consonants in Slovak showed a right-edge alignment for syllables with vocalic as well as consonantal nuclei. Thus, consonants are coordinated sequentially in the syllable onset. However, consonants within the onset cluster are timed differently than the right-most consonant of the onset with the consonantal nucleus. Pouplier and Beňuš (2011) showed that an onset consonant and /l/ or /r/ are timed further apart when the liquids occupy the syllable nucleus than when they are part of the onset cluster. Analogously, /l/ and /r/ and the coda consonant are timed farther apart when /l/ and /r/ occupy the nucleus than when they form the coda cluster.

Differences in kinematics, place and manner of articulation also result in different degrees of overlap between consonants. Byrd (1996) and Byrd and Choi (2010) observed for American English that generally gestures in onset consonant clusters overlapped less than in coda clusters. In Slovak consonants in onset clusters are also timed farther apart than in coda clusters (Pouplier & Beňuš, 2011).

4.1.2 Phrasal accent

In the previous chapter we introduced two hypotheses which predict different realizations of phrasal accent. Both the Hyperarticulation Hypothesis (De Jong, 1995) as well as the Sonority Enhancement Hypothesis (Beckman et al., 1992) are based on the assumption that the syllable consists of a consonantal onset and a vocalic nucleus. Hyperarticulation of a low vowel indeed leads to a lower tongue and jaw position and consequently enhanced sonority. Our main question was whether /l/ in nucleus position would lower the tongue in the anterior region when accentuated to enhance sonority but at the cost of weakening the apical constriction. Our results showed that this was not the case. In accented contexts a more dynamic movement of the tongue was observed. The tongue was first lowered before it was raised again. We attributed this to the reorganization of tongue shape to allow a tongue back retraction and apical constriction and consequently a more curled tongue shape. As already introduced in Section 1.2, the effect of phrasal accent on consonants other than a longer duration is not clear. Fowler (1995, p. 371) suggests that the effect of accent can be described as “an overall increase in global articulatory effort”, which can be interpreted as hyperarticulation. At the same time, the study by Pierrehumbert and Talkin (1992) is often mentioned in which it is suggested that accent makes the rime more vowel-like in contrast to the initial position which makes the syllable onset more consonant-like. However, the study examined only source properties and not oral articulation. Also, if ‘more vowel-like’ refers to a weaker oral constriction our findings for the effect of accent on syllabic /l/ do not support the idea. Bombien (2011) reports unclear effects of phrasal accent on onset consonant clusters. He measured the plateau length of consonants and expected the nucleus adjacent consonants to be more affected by accentuation than the consonants farther apart from the nucleus. However, he found that the plateau duration of the nucleus adjacent consonants were not affected by accent. In general, consonants seem to be affected by boundary strength, rather than phrasal accent (Cho & Keating, 2009). Regarding the effect of phrasal accent on gestural timing Bombien (2011) looked at the effect on onset consonant clusters, and for some of his data he found that there was less overlap in accented conditions. Regarding the temporal coordination of consonants with vowels it was observed for German that onset consonants overlap more with the vocalic nucleus when accented, while coda consonants overlap less with the vocalic nucleus when accented (Peters, 2015).

4.1.3 Research questions

The first aim of this chapter is to better understand the results obtained for the effect of phrasal accent on syllabic /l/ by comparing it to /l/ in onset or coda position. Earlier studies on syllabic consonants in Slovak using EMA showed that /l/ in onset and coda differed kinematically, but /l/ in nucleus patterned mainly like /l/ in coda position (Pouplier & Beňuš, 2011). Studies based on American English proposed that /l/ becomes more vowel like in coda compared to onset by strengthening the dorsal gesture and weakening the apical gesture (Krakow, 1999; Sproat & Fujimura, 1993). However, the cross-linguistic study based on acoustics carried out by Recasens (2012) showed that /l/ does not necessarily vary as much as in American English based on syllable position. We will thus investigate whether in Slovak, too, /l/ in coda position is articulated with a weaker apical gesture and stronger dorsal gesture compared to the onset condition, although indirectly, by comparing the onset and coda /l/ separately with /l/ in nucleus. Our main interest is however, if and how /l/ in nucleus position differs from the onset or coda position. In the previous chapter the dorsal gesture was not involved in signalling phrasal accent. We are thus interested whether a greater variability will be observed as an effect of syllable position for the dorsal gesture, instead of the apical gesture, and whether the lack of variability of the dorsal gesture could be attributed to a ceiling effect in nucleus condition. Syllable position also affects the timing between gestures. In the study by Pouplier and Beňuš (2011) /l/ in nucleus position mainly differed from the onset or coda condition regarding its temporal coordination with the preceding or following consonant. Namely, it was less overlapped when it occupied the nucleus position. So they proposed that /l/ (and also /r/) can function as syllabic nucleus due to the ‘vocalic’ dorsal gesture, which provides the necessary sonorous portion. Although we expected an effect of accent on the ‘vocalic’ dorsal retraction gesture, none was observed. We are thus interested whether the dorsal gesture contributes to prosodic variation in terms of syllable position and the differing temporal coordination. Pouplier and Beňuš (2011) suggested that in nucleus position, the apical gesture should be timed later relative to the dorsal gesture in addition to the reported greater lag relative to the preceding consonant to allow for a longer sonorous interval. Their corpus was not primarily designed to examine the dorsal gesture. However, their preliminary results tend to indicate a stable temporal coordination of the apical and dorsal gestures regardless of onset or nucleus position. We will thus examine the coordination of the two gestures in onset, coda and nucleus position. We will also examine whether phrasal accent has a different effect on /l/ depending on its syllable position. Based on the results reported by Bombien (2011), we expect a weaker effect of phrasal accent on tongue movement of /l/ when it does not occupy the nucleus position. Phrasal accent might also have a different effect on the timing between the gestures of /l/ and the preceding or following /p/. If the cross-linguistic preference for /l/ as a syllabic consonant is because of the dorsal gesture which provides the sonorous segment, this segment should be lengthened in the accented condition, particularly for /l/ in nucleus condition. Therefore, the lag between the release of the initial /p/ and the apical constriction should be lengthened, while the the lag between /p/ and the dorsal gesture should be shortened. This also complies with the observations for German, in that consonantal gestures were less overlapped while the gesture of the onset consonant and the vowel were more overlapped when accented (Peters, 2015). A greater overlap of initial /p/ with the dorsal gesture and less overlap with the apical gesture also implies a greater lag between the dorsal and apical gesture.

4.2 Method

4.2.1 Material

We compare /l/ in nucleus position with both onset coda position. The target words are presented in Table 4.1. For articulatory analysis we will look at tongue movement from the beginning of the acoustically determined segment preceding /l/ to the end of the segment following /l/. Thus for the onset-nucleus comparison we will compare tongue movement in *ple* with *plp*, and for the nucleus-coda comparison in *plp* with *elp*, also shown in Table 4.1. Only short /l/s were compared, because in Slovak phonemic length alternation is observed in nucleus position only. As in the previous chapter, /r/ is not included because its apical gesture was difficult to track in the ultrasound recordings. The target words were elicited for two accent conditions. For further details see the general methods in Section 1.5. As already mentioned, three speakers recorded five repetitions while the other three speakers recorded six repetitions. Two tokens were removed, once *pelpap* in unaccented condition spoken by Speaker 1, and once *pelpap* in accented condition spoken by Speaker 3 because it seemed to be misread as *pepap* or *perpap*. The total number of tokens is thus 196 = 3 target words x 2 accent conditions x (3 speakers x 5 repetitions + 3 speakers x 6 repetitions) - 2 tokens.

Table 4.1: List of target words. The sequence used for articulatory analysis is presented in bold face.

	onset	nucleus	coda
onset - nucleus	**ple**pap	**plp**ap	
nucleus - coda		**plp**ap	p**elp**ap

4.2.2 Measurements

Three measurements will be carried out. First, we will look at the movement of the anterior and posterior part of the tongue. The second measure will be F1 and F2. These two measures are identical to those used in the previous chapter. Contrary to expectations, for /l/ in nucleus position we did not observe a more retracted tongue as an effect of accent. In the anterior region an upward movement of the tongue was observed in the accented condition, but it was not necessarily higher. In this chapter we compare /l/ in nucleus position with /l/ in onset or coda to put the findings of the previous chapter also in relation to the other positions. Thirdly we will look at whether the timing of gestures involved differs for the three syllable positions. This will serve to determine whether and how the temporal coordination of the two gestures differs for the three syllable positions and the two accent conditions. The detailed measurements are described below. It should be kept in mind that the segmental context of /l/ differs for the three syllable positions. In the onset-nucleus comparison the segment following /l/ is different. /l/ in nucleus is followed by /p/, while /l/ in onset is followed by /e/. In the nucleus-coda comparison the segments preceding /l/ differ. /l/ in nucleus position is preceded by /p/ while /l/ in coda position is preceded by /e/.

Tongue movement

First, we will analyse the movement of the anterior and posterior part of the tongue over time. We will compare /l/ in nucleus and onset position, and then separately /l/ in nucleus and coda position. In the onset-nucleus comparison the tongue movement during the acoustically determined *ple*-segment of the target word for the onset condition and *plp*-segment for the nucleus condition will be analysed. In the nucleus-coda comparison the movement during the acoustically determined *plp*-segment for the nucleus condition and *elp*-segment for the coda condition will be analysed. This allows us to also observe the movement into and out of /l/. Tongue contour data are represented in polar coordinates. For each speaker two angles were selected, one at the anterior part and another at the posterior part of the tongue. At each angle, the radius at which it crosses the tongue contour was measured for each ultrasound frame. This results in a curve that represents radial distance of the tongue at the specified angle over time. The same angles were used as for syllabic /l/ in the previous chapter. The posterior angle should capture the tongue retraction movement, while the anterior angle captures the height of the tongue blade and for speakers with a very curled tongue the dip behind the raised tongue tip is captured. As mentioned in the previous chapter, at these angles the tongue was visible throughout the entire target sequence for all tokens.

For statistical analysis time was normalized from 0 to 1 and FLMM (Cederbaum, 2017; Cederbaum et al., 2016; Pouplier et al., 2017) was used (see section 2.2.1 for a brief overview of the method). For each comparison (onset-nucleus and nucleus-coda), tongue movement was compared at two angles, resulting in four model analyses with each model consisting of two covariates: Phrasal accent (0 = unaccented, 1 = accented) and Syllable position (0 = nucleus, 1 = either onset or coda). The overview of covariates and their interaction is given in Table 4.2. The reference mean represents the condition where /l/ occupies the nucleus position and the target word is not accented.

Table 4.2: Overview of covariates for articulatory and acoustic analyses.

$f_0(t)$	Reference mean	unaccented /l/ in nucleus position
$f_1(t)$	Phrasal accent (covariate 1)	0 = unaccented, 1 = accented
$f_2(t)$	Syllable position (covariate 2)	0 = /l/ in nucleus, 1 = onset or coda
$f_3(t)$	Syllable position * Phrasal accent	

F1 and F2

F1 and F2 were extracted for the acoustically segmented /l/. The time normalized F1 and F2 contours in nucleus will be compared separately with the formant contours in onset or coda position, using FLMM. Each of the four models consisted of two covariates, SYLLABLE POSITION and PHRASAL ACCENT, and their interaction. The reference mean corresponds to the formant contour of /l/ in nucleus position of the unaccented target word. The information is also summarized in Table 4.2.

Gestural timing

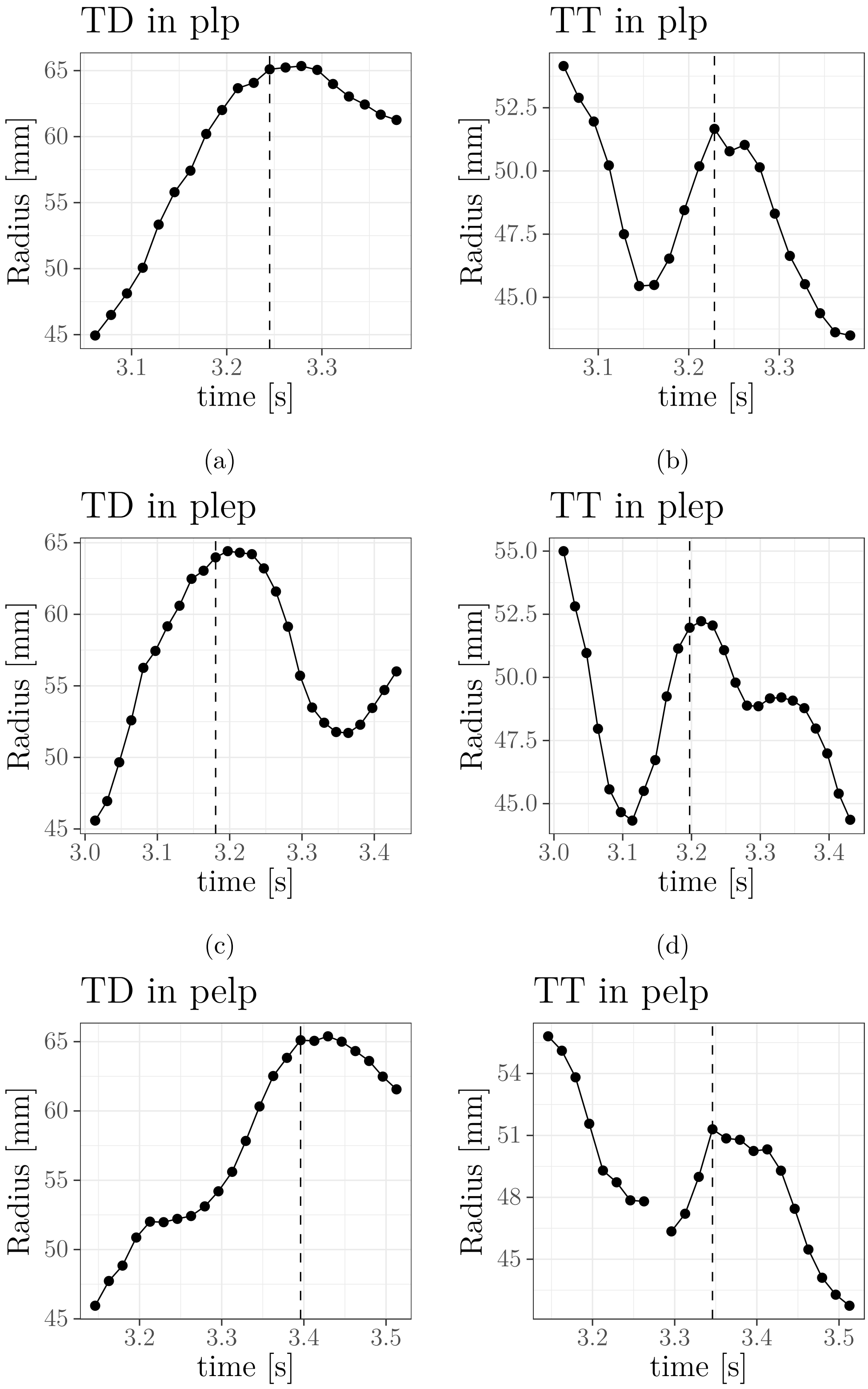
For the timing analysis, we will look at how the dorsal retraction gesture and the apical constriction gesture are timed with each other, but also in relation to the preceding or following consonant. The dorsal retraction gesture is captured at the same posterior angle as for the earlier articulatory analyses. To capture the apical constriction gesture a different angle was chosen than the anterior angle. It is further front and an upward movement of the tongue tip could be captured. The angles at which the data was retrieved are shown in Figure 3.1 of the previous chapter. The angle at which the apical gesture was retrieved is indicated using a red dashed line. For each recording the time points were retrieved at which the gestural peak was achieved. We defined the gestural peak as being achieved when the radius was 0.45mm or slightly less than 3 pixels within the maximal displacement of the gesture within the segment of interest. This measure was chosen based on the methods used by Gick et al. (2006) and Lin (2011). Gick et al. (2006) defined the target to be achieved when the gesture was within one pixel or less than 0.4 mm of the maximal displacement. Lin (2011) measured the distance of the tongue from a region of interested along the palate trace and defined the achievement time point when the aperture was within 3 pixels of the minimum aperture. When the achievement of the target could not be determined, the measurement for that gesture in the specific recording was excluded. Examples of the tongue movement contours with the indicated time point at which the target is achieved are shown in Figure 4.1. The tongue movement contour for the apical gesture during *pelp* (in *pelap*) in Figure 4.1 (f) is not continuous. At one time point between 3.25s and 3.3s the tongue contour was not visible so that the data is missing. In the given examples the gestures are well defined, but this was not the case for all speakers.

Speaker 1 and 6 were excluded completely. For speaker 1, it was not possible to correctly capture the maximal displacement for the apical gesture. When listening to the recordings and looking at the original ultrasound frames, the apical gesture preceded the dorsal gesture. Yet what we thought to be the raising gesture of the tongue tip when looking at the tongue position data for the most anterior angle was the movement out of the constriction. Tongue contour data of speaker 6 had too much noise to reliably extract the displacement data. The total number of tokens used for the timing analysis is 117. Details are listed in Table 4.3.

First we will look at the within-segment coordination. Therefore we analyse the temporal coordination of the dorsal retraction gesture and apical raising gesture. We define the within-segment lag as the achievement of the achievement of the dorsal target minus the achievement of the apical target. A negative lag indicates that the dorsal gesture precedes the apical gesture, and a positive lag indicates that the apical gesture precedes the dorsal gesture. This is the opposite of the measure used by Sproat and Fujimura (1993) and Gick et al. (2006), but correspond to Lin (2011). Secondly, we will look at the timing with respect to the offset or onset of the preceding or following plosive, to examine whether the two gestures of /l/ are timed differently with respect to the plosive, depending on whether they are part of the onset or coda cluster or in nucleus position. To compare /l/ in onset and nucleus position, the duration from the burst of initial /p/ to the achievement of the dorsal retraction gesture or apical constriction gesture was measured. To compare /l/ in nucleus and coda position, the duration from the two gestures to the beginning of the closure for the second /p/ (beginning of the voiceless segment) was measured. In any case, a greater value indicates that the gestures are further apart.

Table 4.3: Number of tokens for each speaker used for the timing analysis.

	Accented			Unaccented		
Speaker	plepap	plpap	pelpap	plepap	plpap	pelpap
2	1	5	2	5	5	5
3	6	6	5	6	5	6
4	5	4	5	5	4	5
5	6	6	6	4	4	6
Total	18	21	18	20	18	22



(a) (b)

(c) (d)

(e) (f)

Figure 4.1: Examples showing the target achievement time points.

4.3 Results

4.3.1 Tongue movement

The first part of our analysis concerns the quality of /l/ in different syllable positions regarding the magnitude of the two gestures. First the results for the onset-nucleus comparison will be presented, followed by the results for the nucleus-coda comparison. Based on the relatively consistent differences across languages, we expect a stronger dorsal gesture and a weaker anterior gesture in coda position compared to the onset condition. Less is known about /l/ in nucleus position. Based on the fact that generally /l/ in coda position is described as more vocalic due to the stronger dorsal constriction and the fact that for Slovak /l/ in nucleus condition patterned more like for coda condition kinematically, we expect /l/ in nucleus to be more similar to the coda condition. Regarding the effect of phrasal accent we expect it to be weaker for the onset and coda condition compared to the nucleus condition. Instead of looking at tongue movement during the acoustically segmented /l/ only, we will compare tongue movement during the preceding and following segment as well, as we did in the previous chapter. This means that for the nucleus context, we will analyse tongue movement during the acoustically segmented *plp* in ***plp****ap*, for the onset context during *ple* in ***ple****pap* and for the coda context during *elp* in *p****elp****ap*. Consequently, for the onset nucleus comparison we will compare *ple* with *plp* and for the the nucleus coda comparison we will compare *plp* with *elp*. For each comparison the results for the posterior angle capturing the retraction angle will be presented first, then the result at the anterior angle.
The reference mean curve for the FLMM analysis corresponds to the unaccented nucleus condition and the first covariate shows the effect of phrasal accent, which is identical to what we already looked at in the previous chapter. Only relevant results will be repeated here.

Analogous to the previous chapter, we will briefly report the acoustic duration of the segments involved in each analysis and visualize how on a time scale from 0 to 1 the target sequence is acoustically divided into the individual segments. This will also enable a better comparison of the acoustic and articulatory data, because the acoustic data is limited to the acoustically defined /l/ segment.

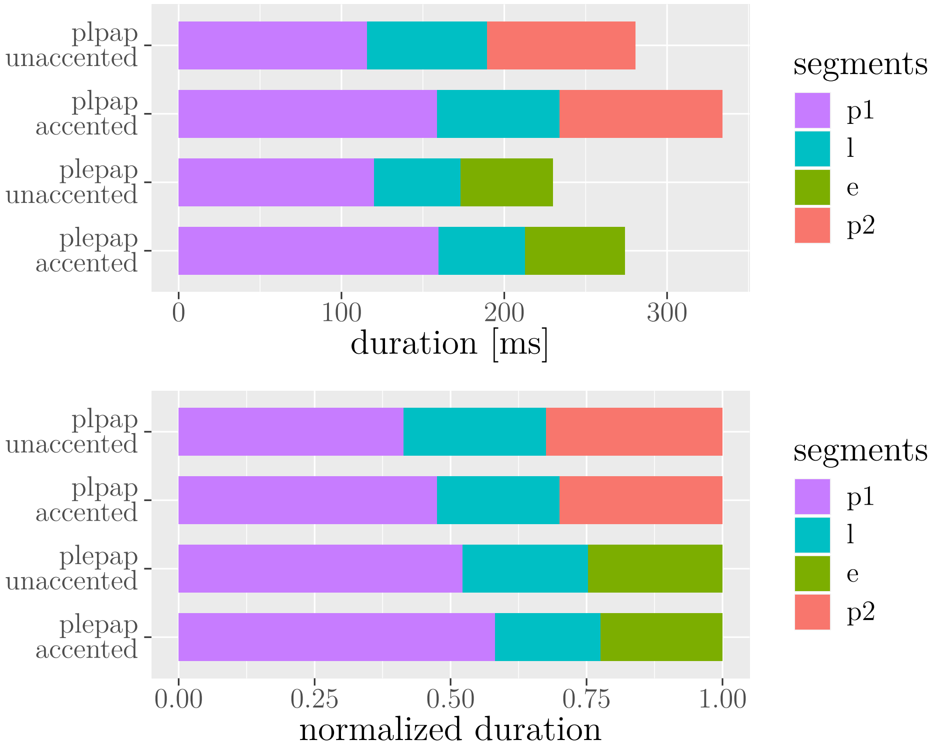


Figure 4.2: Absolute (top) and normalized (bottom) mean durations of *plp* in *plpap* for the nucleus condition and *ple* in *plepap* for the onset condition.

Onset and nucleus position

The mean durations of *p1* (initial /p/), *l* and *e* within the *ple* segment of *plepap* and *p1*, *l* and *p2* (second /p/) within the *plp* segment of *plpap* are shown in the top plot in Figure 4.2 and the corresponding normalized durations are shown in the bottom plot of Figure 4.2. The target word of the nucleus condition is *plpap* and the target sequence consists of *p1* (initial /p/), *l* (nucleus) and *p2*. The target word of the onset condition is *plepap* and the target sequence consists of *p1* (initial /p/), *l* (onset) and *e* (vocalic nucleus). We will first report the absolute duration of individual segments and then interpret the proportional make-up of the time-normalized segment.

Linear mixed effects analysis of the effect of SYLLABLE POSITION of *l* (whether *l* occupied the nucleus or was part of the initial consonant cluster) and PHRASAL ACCENT on the acoustic duration of *p1* revealed that the effect of PHRASAL ACCENT was significant ($\chi^2[1] = 46.6, p < 0.001$). SYLLABLE POSITION of /l/ did not have significant effect on the duration of *p1* ($\chi^2[1] = 0.2, p > 0.1$) and neither did the interaction effect of PHRASAL ACCENT and SYLLABLE POSITION ($\chi^2[2] = 0.3, p > 0.1$). The mean duration of initial *p1* was $161.7ms \pm 11.1$ when accented. In unaccented condition *p1* was $41.2ms \pm 5.5$ shorter than in accented condition.

Regarding the acoustic duration of *l* linear mixed effects analysis confirmed that SYLLABLE POSITION had a significant effect ($\chi^2[1] = 44.2, p < 0.001$), while PHRASAL ACCENT ($\chi^2[1] = 0.1, p > 0.1$) and the interaction of SYLLABLE POSITION and PHRASAL ACCENT ($\chi^2[2] = 0.2, p > 0.1$) did not have a significant effect. Mean acoustic duration of /l/ in *plepap* was $53ms \pm 6.3$ and in *plpap* it was $22.3ms \pm 4.1$ longer.

For the third segment the effect of *phrasal accent* was tested separately for *p2* in *plpap* and *e* in *plepap*. *p2* was affected significantly by phrasal accent ($\chi^2[1] = 14.9, p < 0.001$) while *e* was not ($\chi^2[1] = 3.7, p > 0.05$). The acoustic duration of *p2* in accented condition was $100.2ms \pm 7.8$ and in unaccented condition it was $8.9ms \pm 2.1$ shorter. The mean duration of *e* was $59.2ms \pm 3.3$. Thus, *e* is much shorter than *p2*.

The acoustic duration of *l* and *e* in *plepap* (/l/ in onset) were both shorter than the acoustic duration of *l* and *p2* in *plpap* (/l/ in nucleus), but the syllable affiliation of /l/ did not have an effect on *p1*. The segments *l* and *e* take up a smaller portion of the normalized time sequence than *l* and *p2*. Consequently, *p1* takes up a greater portion for the onset condition than for the nucleus condition. PHRASAL ACCENT had a significant effect on *p1* and *p2*. However, the effect on *p2* was much smaller. When time normalized, *p1* is longer in the accented than in the unaccented condition and consequently all other segments are shorter in the accented condition.

Posterior angle We now begin with the results for the tongue movement captured at the posterior angle for the onset-nucleus comparison. We compared the tongue movement during acoustically segmented *ple* in *plepap* for the onset condition and *plp* in *plpap* for the nucleus condition. In Figure 4.3 the data used for the statistical analysis is visualized. Time is normalized from 0 to 1 and the radius used for the FLMM analysis is z-normalized for each speaker. A higher radius indicates a more retracted tongue. Recall that the vowel preceding the target sequence is /i/, a front vowel, so the retraction gesture is very prominent.

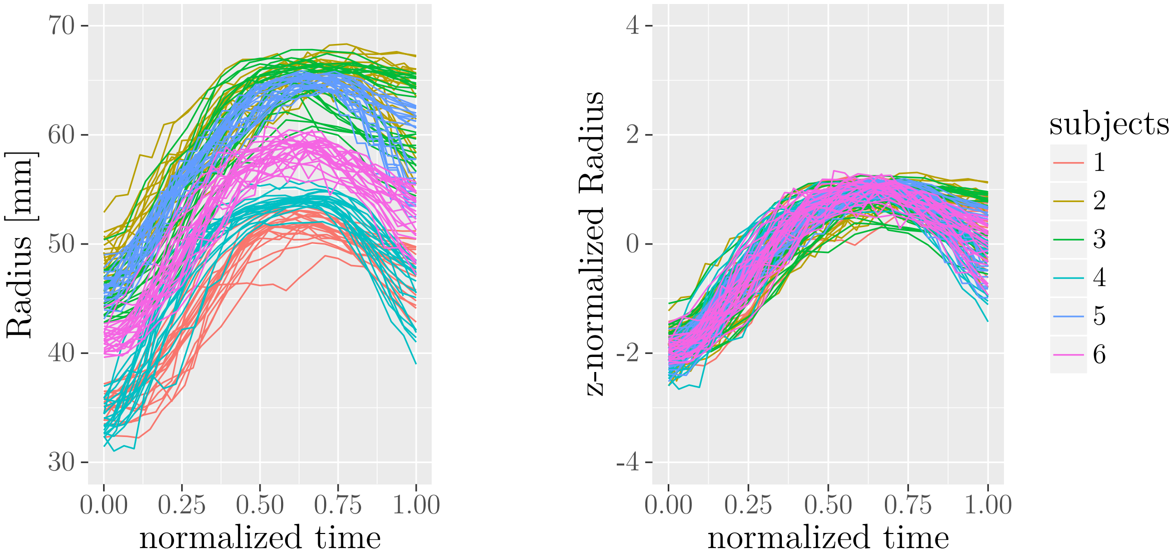


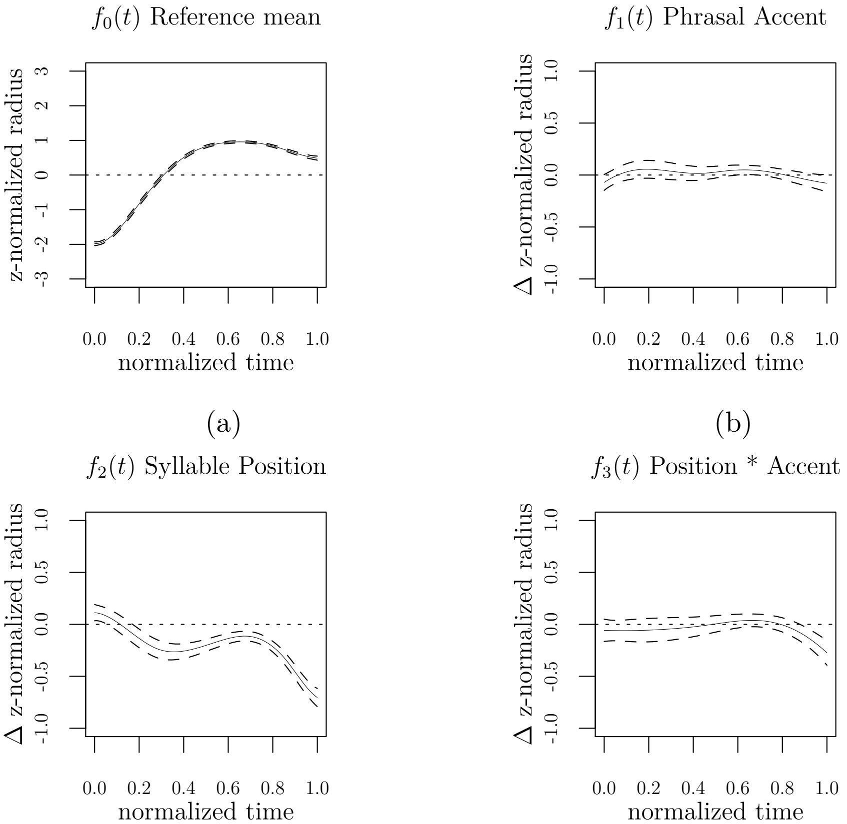
Figure 4.3: Radius over normalized time measured at the posterior angle, capturing tongue retraction for the onset-nucleus comparison. For the onset condition the radius during acoustically segmented *ple* in ***ple****pap* is shown and for the nucleus condition the *plp* sequence in ***plp****ap* is shown. On the left the radius over normalized time, on the right z-normalized radius. A higher radius indicates a more retracted tongue.

To compare tongue movement data in the *plp* segment of *plpap* in which /l/ occupies the syllable nucleus with the tongue movement in the *ple* segment of *plepap* in which /l/ occupies the onset, FLMM was used. The model consisted of two covariates, PHRASAL ACCENT and SYLLABLE POSITION as well as their interaction. The reference mean corresponds to the tongue movement in unaccented condition in which /l/ occupies the nucleus position (see also Table 4.2 as a reference). The reference mean curve $f_0(t)$, covariate effect curves for PHRASAL ACCENT $f_1(t)$ and SYLLABLE POSITION $f_2(t)$, and the interaction effect curve $f_3(t)$ are shown in Figure 4.4. The effects are significant when the confidence bands do not include zero. In the previous chapter where we looked at /l/ in nucleus position, we had already shown that PHRASAL ACCENT did not have a significant effect on tongue retraction. The covariate effect curve for SYLLABLE POSITION $f_2(t)$ has a significant lowering effect after time point 0.2. This means that the tongue is less retracted for /l/ when it occupies the onset position than in the nucleus position. This is in accordance with our prediction that the tongue is generally less retracted for the onset condition. The effect is particularly strong after time point 0.8. This is during the acoustically segmented vowel /e/ for the tokens in which /l/ occupies the onset as can be seen in the bottom plot of Figure 4.2. /l/ in nucleus position is followed by /p/ for which the tongue position is not defined, so the tongue moves towards the following vowel /a/, for which the tongue does not need to be fronted as much as for /e/.

The tongue movement captured at the posterior angle for the onset condition is also not affected by PHRASAL ACCENT except after time point 0.9. This is during the vowel /e/, which in accented condition is more fronted than in unaccented condition.

The summed effect curves are visualized in Figure 4.5. Red lines represent the nucleus condition and blue lines the onset condition. The unaccented condition is visualized using dashed lines and for the accented condition solid lines are used. The blue lines raise more slowly compared to the red lines, but it does not mean that the tongue movement for the nucleus condition is indeed faster. Recall that within the time normalized segment the initial /p/ occupied a longer segment for the onset condition although SYLLABLE POSITION did not have a significant effect on its absolute duration. If time were not normalized, the red lines would be stretched along the time axis, and the rising slope would become less steep.

In Figure 4.10 the summed effect curves are shown separately for each condition with their corresponding acoustically determined segment boundaries. To summarize, the tongue was slightly less retracted when /l/ occupied the onset position, as predicted. There was no effect of phrasal accent on the magnitude of the retraction gesture.



(a) (b)

(c) (d)

Figure 4.4: Reference mean and covariate effect curves of the tongue movement captured at the posterior angle for the onset-nucleus comparison. A higher radius indicates a more retracted tongue.

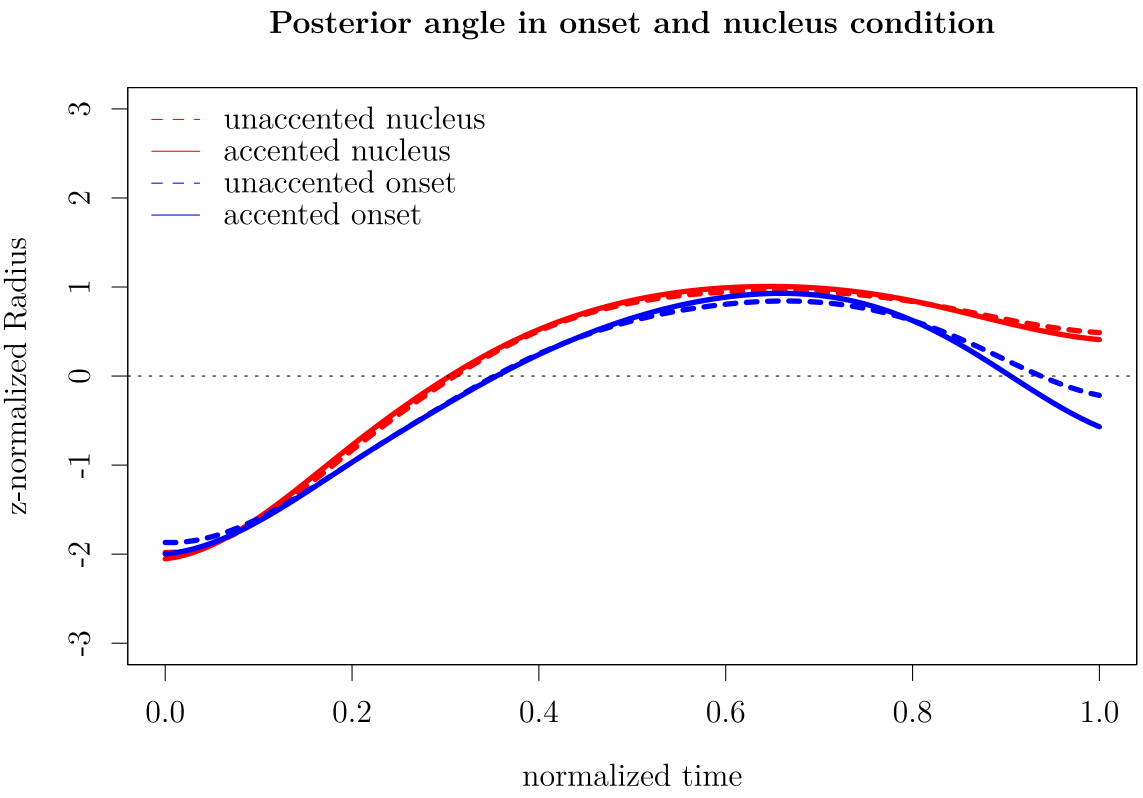


Figure 4.5: Summed effect curves of the tongue retraction movement captured at the posterior angle during the *plp* sequence of *plpap* and the *ple* sequence of *plepap*. A higher radius indicates a more retracted tongue position. The red dashed line corresponds to the unaccented nucleus condition which is also the reference mean curve $f_0(t)$. The red solid line corresponding to the accented nucleus condition is obtained by adding the covariate effect curve $f_1(t)$ for PHRASAL ACCENT to the reference mean curve. The blue dashed line corresponds to the unaccented onset condition and is obtained by adding the covariate effect curve for SYLLABLE POSITION $f_2(t)$ to the reference mean curve. The blue solid line representing the accented onset condition is obtained by adding the covariate effect curves for PHRASAL ACCENT and SYLLABLE POSITION and the interaction effect curve to the reference mean $(f_0(t) + f_1(t) + f_2(t) + f_3(t))$.

Anterior angle We now continue with the results for tongue movement data as captured at the anterior angle. Recall that this angle does not capture the very tip of the tongue at the apical constriction. For speakers with a curled tongue for /l/ it captures the lowered tongue blade behind the raised tongue tip. The tongue movement data during the target segment as it was used for the statistical analysis is shown in Figure 4.6. Time is normalized from 0 to 1 and the radius was z-normalized as shown on the right panel. A higher radius indicates a higher tongue position. The vowel preceding the target word is /i/ which requires a high tongue position. During the initial /p/, the tongue is lowered. It might be unintuitive, but at this angle the tongue is lowered during /l/ to create a curled tongue shape with a higher tongue tip and a more retracted tongue back. For /e/ in *plepap* the tongue is then raised again, although not as much as for /i/. Figure 4.7 shows the tongue contours at three time points within the *ple* sequence with the angle at which the tongue position for the front region was measured for speaker 4. The tongue contour at the very beginning of the the target sequence, just after the preceding high vowel /i/ of the carrier phrase is shown in orange. With the tongue tip being at the left, we see that a very high tongue position is captured. The blue contour represents the tongue at the first frame within the acoustically segmented /l/. It shows a raised tongue tip and a retracted dorsal part of the tongue, but at the angle where the tongue position was captured the radius is lower than at the beginning of the target sequence. The green contour shows the tongue at the last frame of /e/. The tongue position is higher than for /l/ but not as high as at the beginning of the target sequence.

The model for the FLMM statistics consisted of two covariates, SYLLABLE POSITION and PHRASAL ACCENT, and their interaction. The reference mean corresponds to the tongue movement in unaccented condition where /l/ occupies the nucleus position (see also Table 4.2). Figure 4.8 shows the reference mean, covariate and interaction effects. The summed effect curves are shown in Figure 4.9. There is a significant effect for PHRASAL ACCENT. As already described in the previous chapter, in the accented condition the tongue is lowered more during the initial consonant and has a lower position at the beginning of the acoustically segmented /l/. However, the tongue has a more dynamic movement in accented condition so that towards the end the tongue for the two accent conditions has the same height.

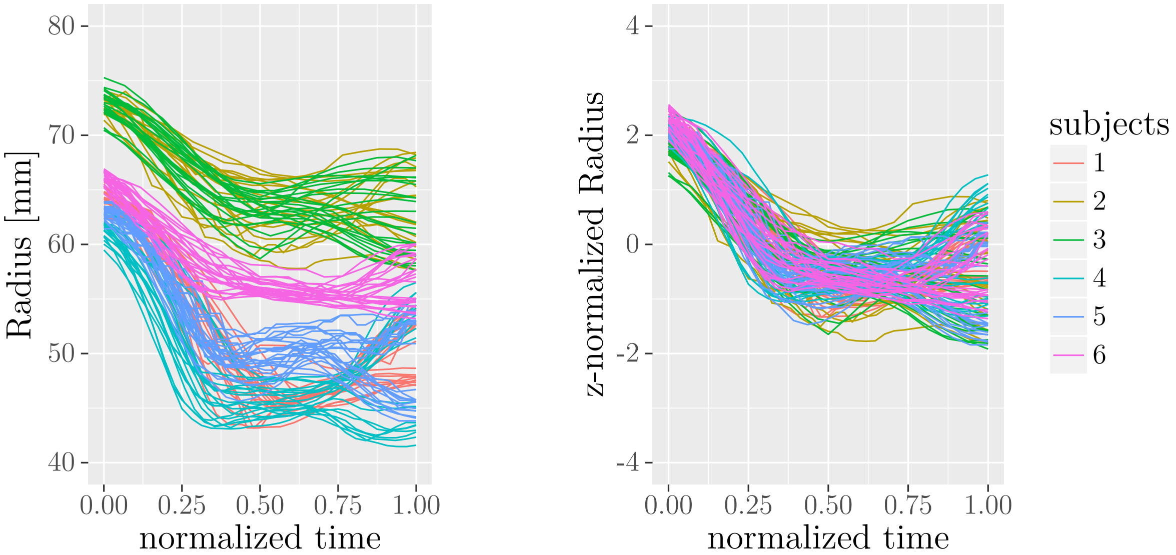


Figure 4.6: Radius over normalized time measured at the anterior angle, capturing tongue height for the onset-nucleus comparison. For the onset condition the radius during acoustically segmented *ple* in ***ple**pap* is shown and for the nucleus condition the *plp* sequence in ***plp**ap* is shown. On the left the radius over normalized time, on the right z-normalized radius. A higher radius indicates a higher tongue

The covariate effect curve for SYLLABLE POSITION $f_2(t)$ has a significant raising effect. This means that the tongue is higher for /l/ in onset condition. After time point 0.7 the effect becomes even greater, because the tongue rises for /e/. In Figure 4.9 the blue dashed line corresponding to the unaccented onset condition is higher than the red dashed line, corresponding to the unaccented nucleus condition.

The interaction effect curve $f_3(t)$ (Figure 4.8 (d)) has a significant effect between around time point 0.1 and 0.4. This has an effect on the initial falling slope and also works in the opposite direction to $f_1(t)$. In Figure 4.9 we can see that the difference between the blue dashed and solid lines is smaller than between the red dashed and solid lines. The blue lines are less steep than the red lines and the blue solid line reaches its lowest position later than the red solid line, but these might also be an artefact of time normalization. There is a trend

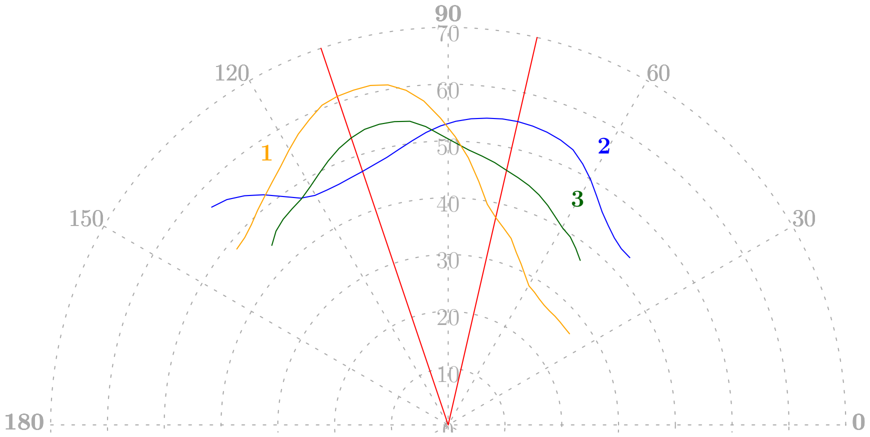
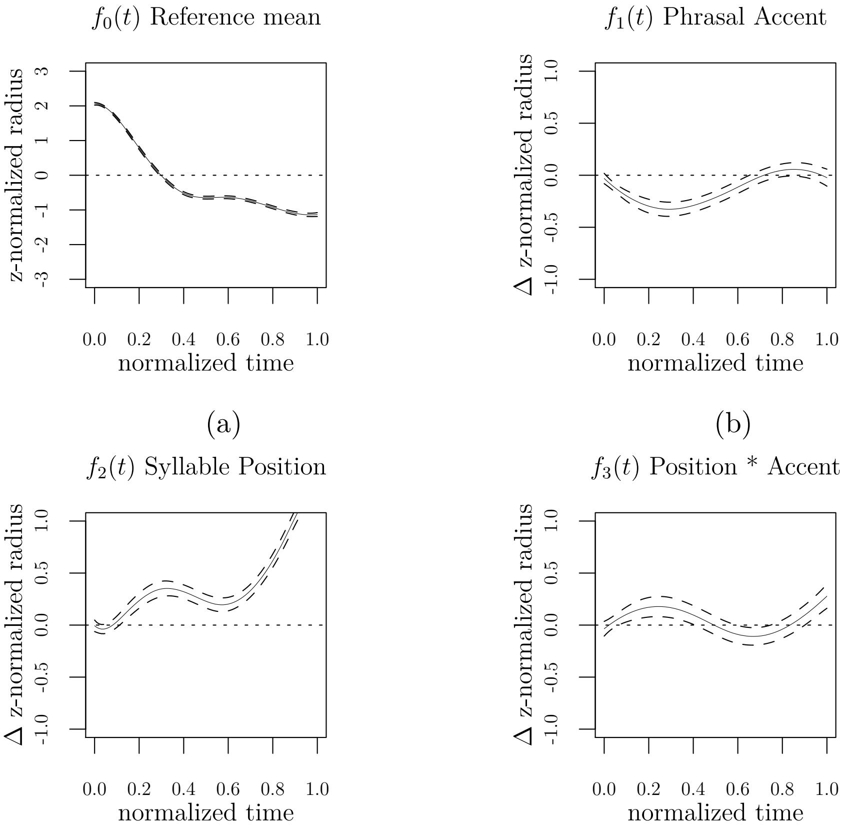


Figure 4.7: Tongue contours of speaker 4 at three time points during the target sequence in *plepap*. The tongue tip is to the left. The tongue contour at the first frame in *plepap* is shown in orange, at the first frame of the acoustically segmented /l/ is shown in blue and at the last frame of /e/ is shown in green. The angles at which the radius of the tongue position were measured are drawn in red.

level effect around time point 0.7 and a significant effect after time point 0.9. These effects can be better understood when looking at the summed effect curves in Figure 4.9. The time point at which the blue lines cross is slightly later than the time point at which the red lines cross. And after time point 0.9 the blue solid line is higher than the blue dashed line, while the two red lines are overlapping. This shows that /e/ is affected by accent. The effect of $f_3(t)$ opposes the effect of $f_1(t)$, although at different magnitude. In general, it can be said that $f_3(t)$ weakens the effect of PHRASAL ACCENT $f_1(t)$.

In sum, the tongue front was higher for /l/ in onset condition than in nucleus condition. When accented, the tongue front was lower but the effect was smaller for the onset condition. The results confirm our predictions in that the tongue position during /l/ in onset was higher than during /l/ in nucleus and the effect of accent was greater for the nucleus condition.



(a) (b)

(c) (d)

Figure 4.8: Reference mean and covariate effect curves of the tongue movement captured at the anterior angle for the *ple*-sequence in *plepap* (onset condition) and the *plp*- sequence in *plpap* (nucleus condition). A higher radius indicates a higher tongue position.

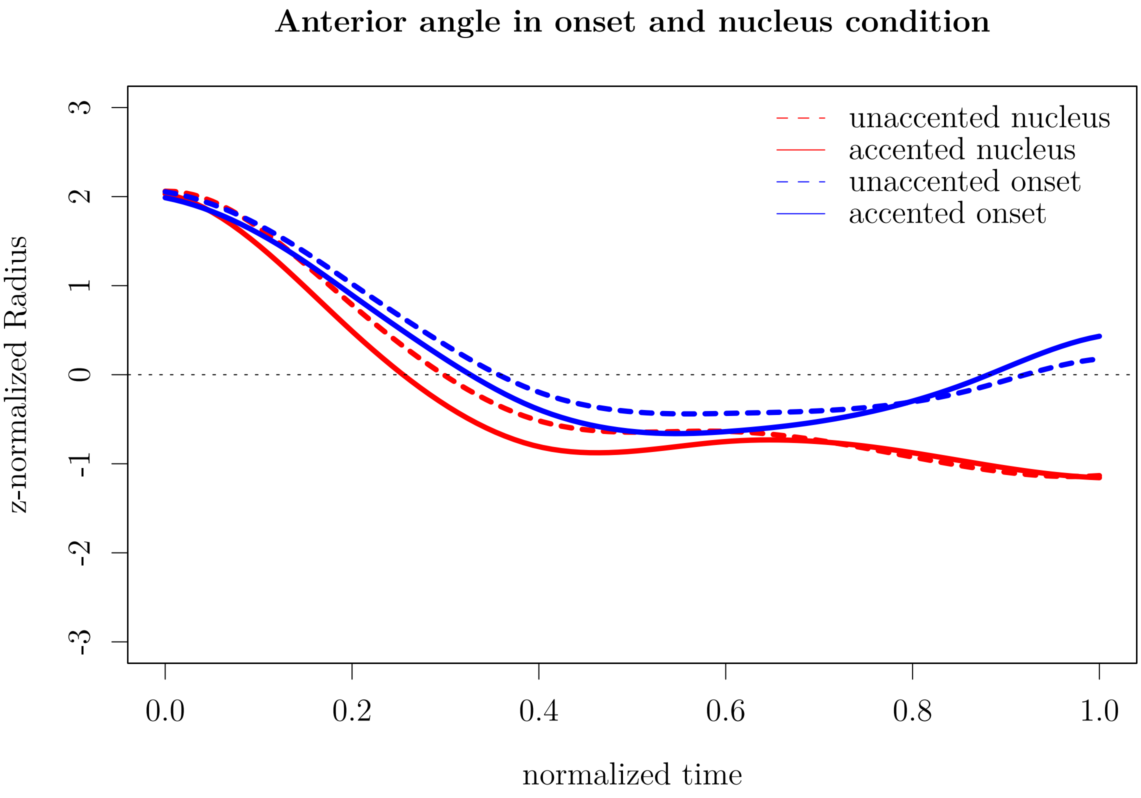


Figure 4.9: Summed effect curves of the tongue movement captured at the anterior angle for the *ple*-sequence of *plepap* (onset condition) and the *plp*- sequence of *plpap* (nucleus condition). A higher radius indicates a higher tongue position. The red dashed line, also the reference mean $f_0(t)$, corresponds to the unaccented nucleus condition, the red solid line to the accented nucleus condition ($f_0(t) + f_1(t)$), the blue dashed line to the unaccented onset condition ($f_0(t)+f_2(t)$) and the blue solid line to the accented onset condition ($f_0(t)+f_1(t)+f_2(t) + f_3(t)$).

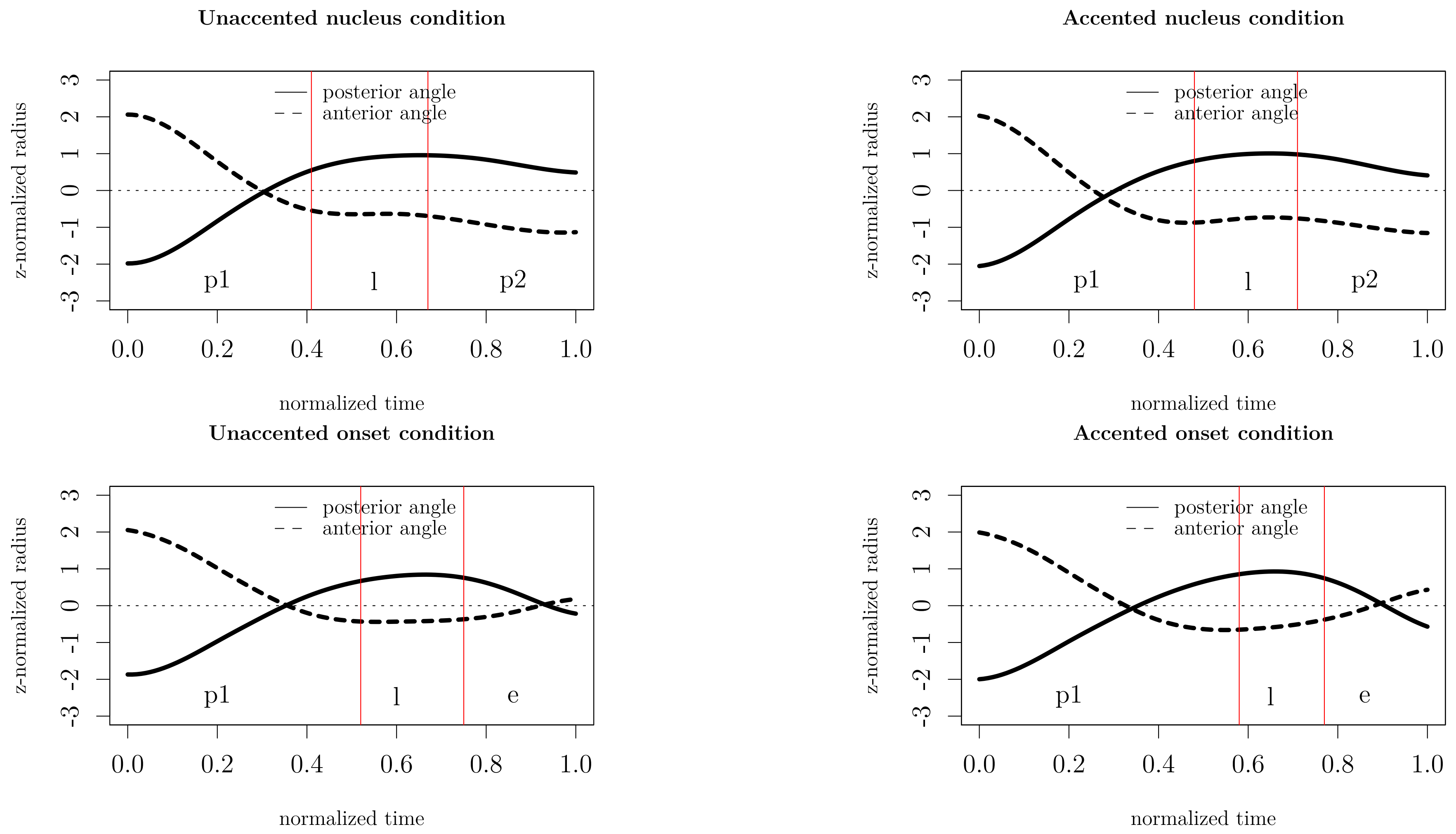


Figure 4.10: Summed effect curves of the tongue movement captured at the posterior and anterior radii for the onset-nucleus comparison. The red vertical lines represent the boundaries between the acoustically determined segments *p1* (onset consonant /p/), *l*, and *e* (when /l/ is part of the onset cluster) or *p2* (/p/ following /l/ in nucleus position). A lower radius at the anterior angle indicates tongue lowering, a lower radius at the posterior angle indicates tongue fronting..

Nucleus and coda position

Consistent to the results presented for the onset-nucleus comparison, we will first look at the acoustic duration of the target sequence in ***plp**ap* and *p**elp**ap* for the nucleus-coda comparison. Figure 4.11 shows how the target sequence is divided in each segment for the different conditions. The target word for the nucleus condition is *plpap* and the target sequence consists of *p1*, *l* and *p2*. The target word for the coda condition is *pelpap* and the target sequence consists of *e*, *l* and *p2*. The absolute duration of individual segments will be reported first. We will then relate these to the time-normalized segments. The absolute and normalized durations are visualized in Figure 4.11.

We have already shown that the acoustic duration of *p1* in *plpap* was affected significantly by PHRASAL ACCENT. The mean duration of *p1* in accented condition was $161.7ms \pm 11.1$ when accented and $41.2ms \pm 5.5$ shorter when unaccented. The acoustic duration of *e* was also affected significantly by PHRASAL ACCENT ($\chi^2[1] = 5.61.98, p < 0.05$). The mean duration of *e* in *pelpap* in accented condition was $68.4ms \pm 5.8$ and in unaccented condition it was $5.5ms \pm 2.3$ shorter. This is unexpected, as generally the acoustic duration of short nuclei was not affected by PHRASAL ACCENT (see Chapter 2 Section 2.3.2 and the previous section for the onset-nucleus comparison for the effect of PHRASAL ACCENT on *e* in *plepap*). One reason for this observation could be that it was generally difficult to define an acoustic boundary between *e* and *l*.

To test for the effect of phrasal accent and position of *l* within the syllable on the acoustic duration of each segment, a linear mixed effects analysis with PHRASAL ACCENT and SYLLABLE POSITION as fixed factors was conducted. The acoustic duration of *l* was significantly affected by SYLLABLE POSITION ($\chi^2[1] = 17.2, p < 0.001$), but not by PHRASAL ACCENT ($\chi^2[1] = 1.98, p > 0.1$). Compared to the mean acoustic duration of *l* in *pelpap* ($68.3ms \pm 8.1$) where it occupies the coda position, *l* in nucleus position was $7.9ms \pm 2.8$ longer. Acoustic duration of *p2* was affected by both SYLLABLE POSITION($\chi^2[1] = 7.5, p < 0.01$) and PHRASAL ACCENT ($\chi^2[1] = 27.1, p < 0.001$), but the interaction of SYLLABLE POSITION and PHRASAL ACCENT did not have a significant effect ($\chi^2[1] = 5 * 10^{-4}, p > 0.1$). For the accented condition the mean duration of *p2* in *pelpap* was $95.6ms \pm 7.7$ and in *plpap* it was $4.5ms \pm 2.3$ longer. In unaccented condition, it was $9.0ms \pm 2.3$ shorter.

The greatest difference was between the duration of *p1* and *e* which we did not com-

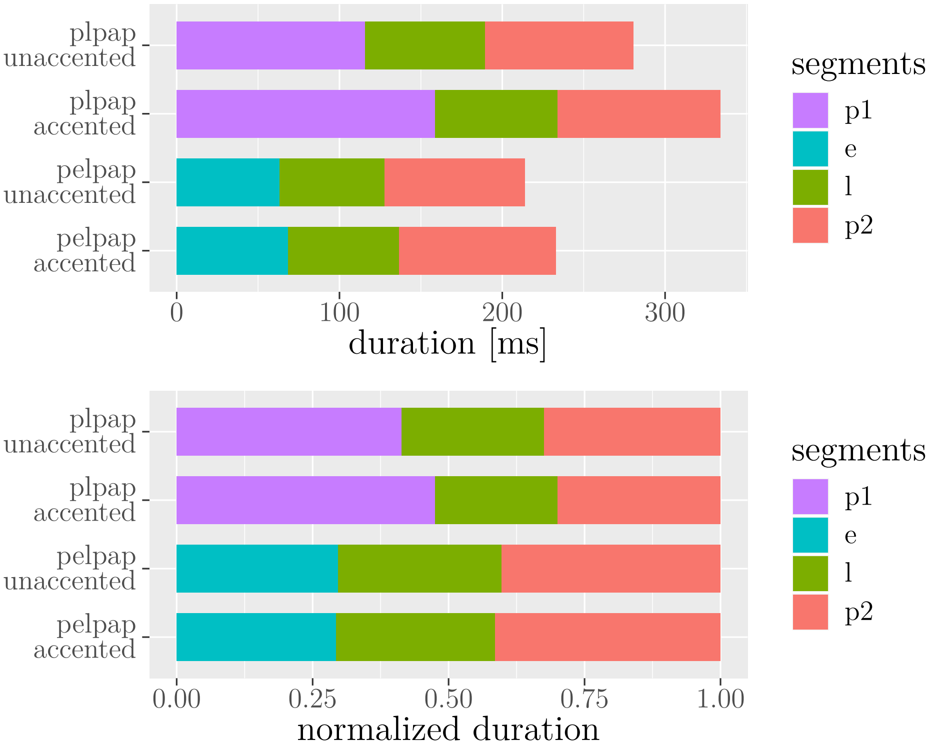


Figure 4.11: Absolute (top) and normalized (bottom) mean durations of *plp* in *plpap* for the nucleus condition and *elp* in *pelpap* for the coda condition.

pare directly, but both occupy the initial position of the target sequence. Their duration contributes to the main difference in how the time normalized target segment is divided into individual segments. *p1* takes more time than *e*. Although /l/ in nucleus condition was slightly longer than in coda condition, it takes up proportionally less time. And *p2* was slightly longer in the nucleus condition than in the coda condition but when the target sequences are time normalized *p2* is shorter for the nucleus condition than for the coda condition.

Posterior angle Again, we begin with the tongue movement captured at the posterior angle. For the statistical analysis we compared the tongue movement during the *plp* segment in ***plp**ap* (nucleus condition) with the *elp* segment in *p**elp**ap* (coda condition). The data used for the analysis is shown in Figure 4.12. A higher radius indicates a more retracted tongue. Time is normalized from 0 to 1 and radius is z-normalized. At the beginning we can see that the contours are split in two groups for each speaker. During the initial /p/ in *plpap* for the nucleus condition the tongue is less retracted because of the preceding /i/. For the coda condition the target segment begins at /e/ for which the tongue is already more retracted.

The results of the FLMM analysis are shown in Figure 4.13. Figure 4.14 shows the summed effect curves and Figure 4.18 shows the summed effect curves separately for each condition but with their corresponding acoustically determined segment boundaries. The model consisted of two covariates, PHRASAL ACCENT and SYLLABLE POSITION, and their interaction (see also Table 4.2). The reference mean $f_0(t)$ corresponds to the tongue retraction movement in unaccented condition where /l/ occupies the nucleus. PHRASAL ACCENT did not have a significant effect. As expected the covariate effect curve for SYLLABLE POSITION $f_2(t)$ shows a significant effect at the beginning of the target sequence. In coda condition the target sequence begins with /e/ for which the tongue is less retracted compared to the nucleus condition. The segment of the nucleus condition begins with /p/ which has no specification for tongue position, thus the fronted tongue position of the preceding vowel /i/ can be observed. After time point 0.6 SYLLABLE POSITION has a significant lowering effect. However, recall that due to the difference in the acoustic duration of *p1* and *e*, the acoustic /l/ segment in coda condition is placed earlier within the time normalized segment. In Figure 4.14 we can see that the peak of the blue dashed line (as well as the

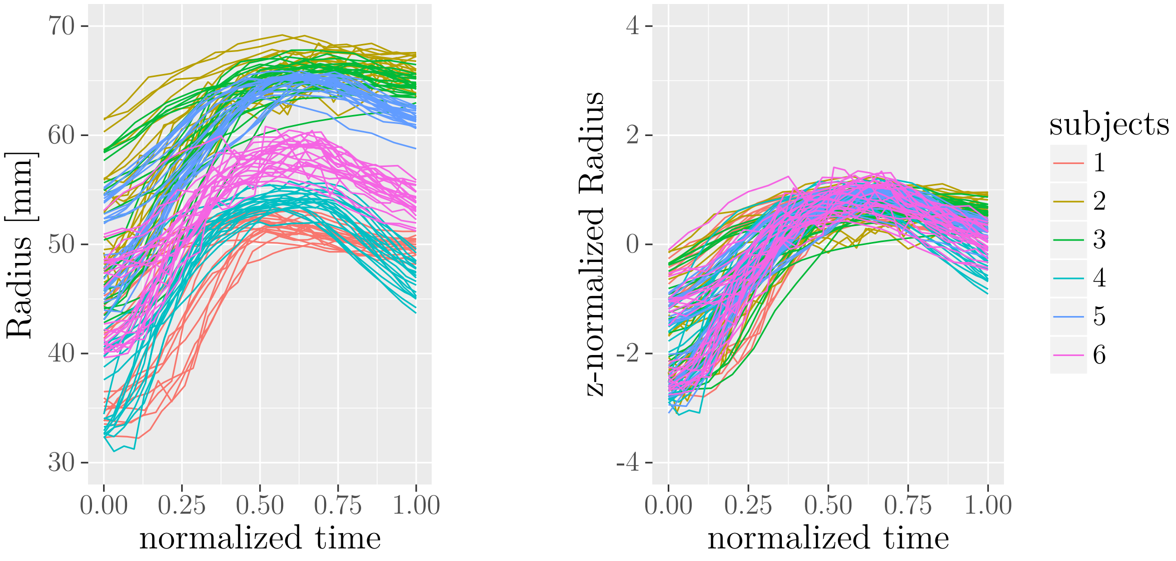


Figure 4.12: Radius over normalized time measured at the posterior angle, capturing tongue retraction for the nucleus-coda comparison. A greater radius indicates a more retracted tongue position. For the nucleus condition the radius during acoustically segmented *plp* in ***plp****ap* is shown and for the coda condition the *elp* segment in ***p******elp****ap* is shown. On the left the radius over normalized time, on the right z-normalized radius.

blue solid line) is earlier than for the red lines. This is just due to our choice of the target sequence and should not be interpreted in the sense of gestural timing. The interaction of SYLLABLE POSITION and PHRASAL ACCENT $f_3(t)$ has significant lowering effect from the beginning until around time point 0.5. In Figure 4.14 we can see that at the beginning the blue solid line corresponding to the accented coda condition is lower than the blue dashed line corresponding to the unaccented coda condition. The main difference is during the nucleus /e/ but it persists also during the beginning of the acoustically segmented coda /l/. The effect is larger at the beginning, which is during the vowel.

In sum, while the peak of the blue lines has a slightly lower radius than the peak of the red lines the difference is most probably only at trend level. Phrasal accent has only a trend level effect on /l/ in nucleus position. However, the interaction effect of phrasal accent and syllable position has a significant effect, so that when /l/ occupies the coda, /e/, the segment preceding /l/ is more fronted in accented condition, and also during /l/ the tongue is less retracted. Phrasal accent thus has an opposite effect on /l/ in nucleus and in coda condition, but the difference can be partially attributed to the segmental context.

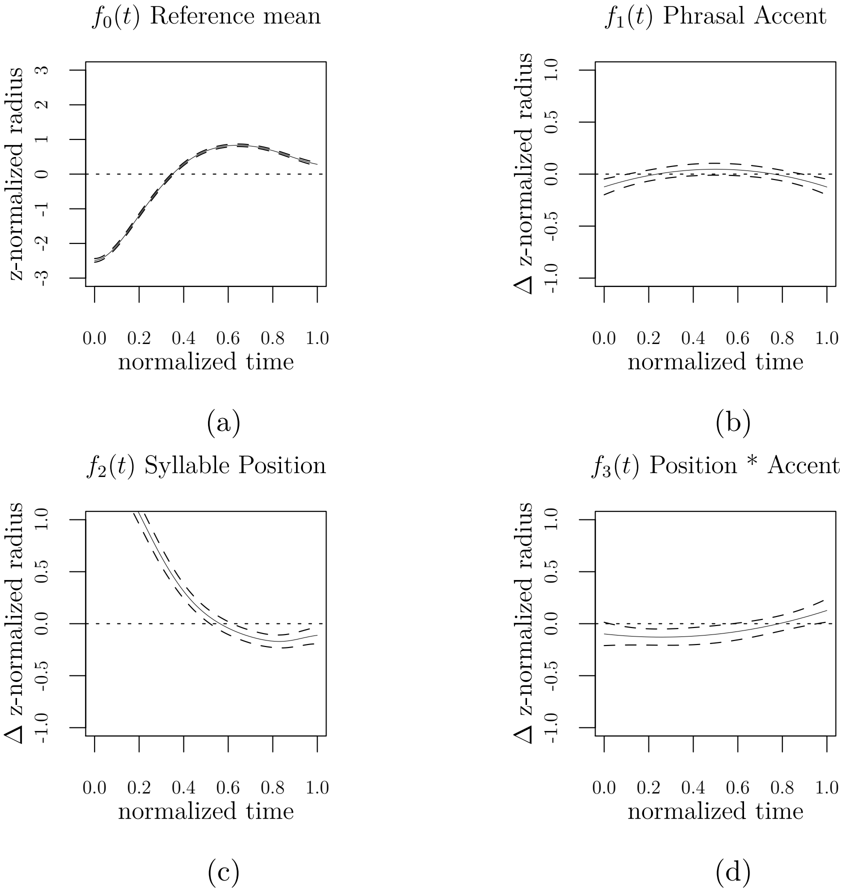


Figure 4.13: Reference mean and covariate effects of the tongue movement captured at the posterior angle for the nucleus-coda comparison: *plp* sequence of *plpap* and *elp* sequence of *pelpap*. A greater radius indicates a more retracted tongue.

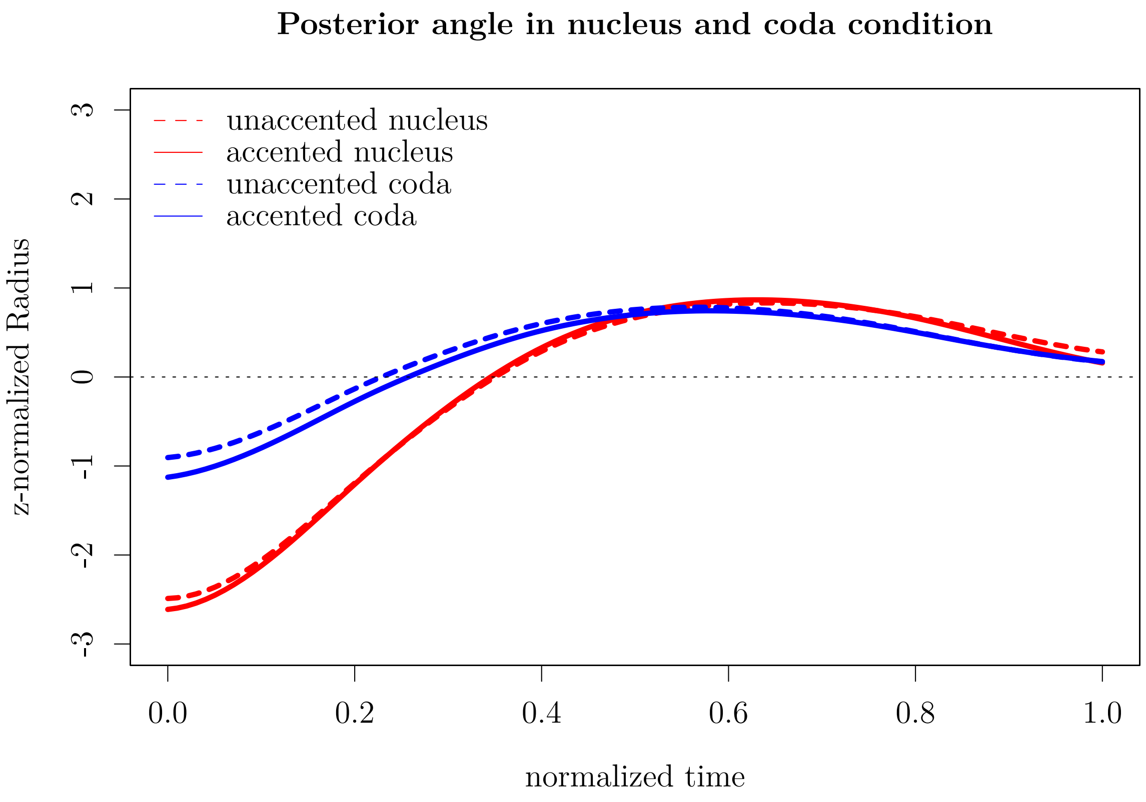


Figure 4.14: Summed effect curves of the tongue movement captured at the posterior angle in *plp* sequence of *plpap* and *elp* sequence of *pelpap*. A greater radius indicates a more retracted tongue. The red dashed line, also the reference mean $f_0(t)$, corresponds to the unaccented nucleus condition, the red solid line to the accented nucleus condition ($f_0(t) + f_1(t)$), the blue dashed line to the unaccented coda condition ($f_0(t) + f_2(t)$) and the blue solid line to the accented coda condition ($f_0(t) + f_1(t) + f_2(t) + f_3(t)$).

Anterior angle Finally the results for the tongue movement for the nucleus-coda comparison captured at the anterior angle will be presented. Remember that at this angle it is not the apical constriction gesture for /l/ that is captured but rather a low tongue due to tongue blade curling. For the statistical analysis tongue movement during the acoustically segmented *plp* in ***plp**ap* (nucleus condition) and *elp* in ***pelp**ap* (coda condition) will be compared. In Figure 4.15 the data used for the statistical analysis is visualized. For each speaker the tongue contours begin at two different heights. In *plp* for the nucleus condition the tongue position at the beginning is higher because the target word is preceded by /i/, while in *elp* the tongue has already moved to a lower position for /e/ during the initial /p/.

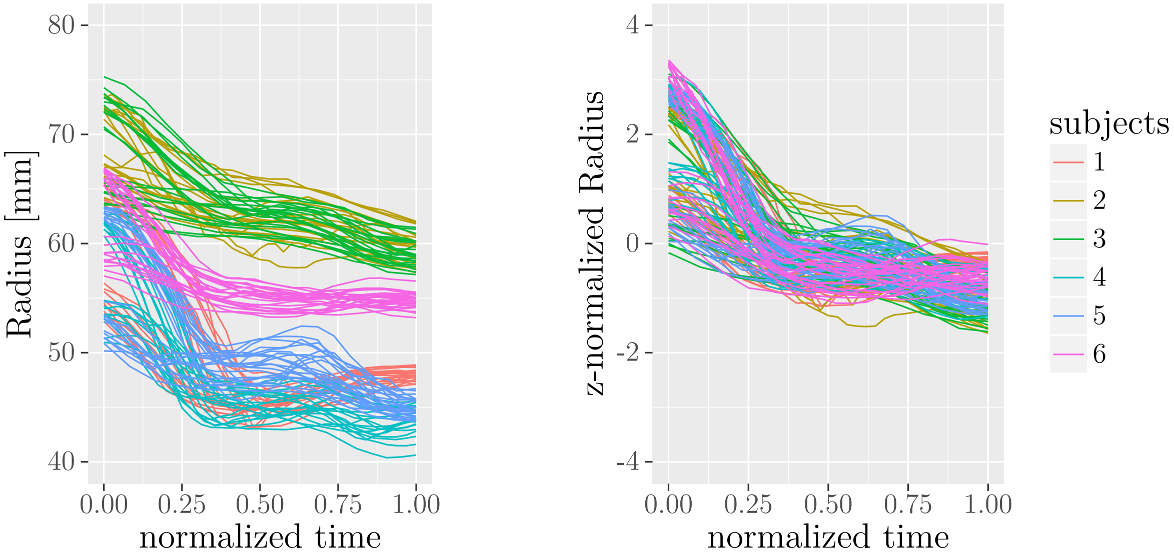


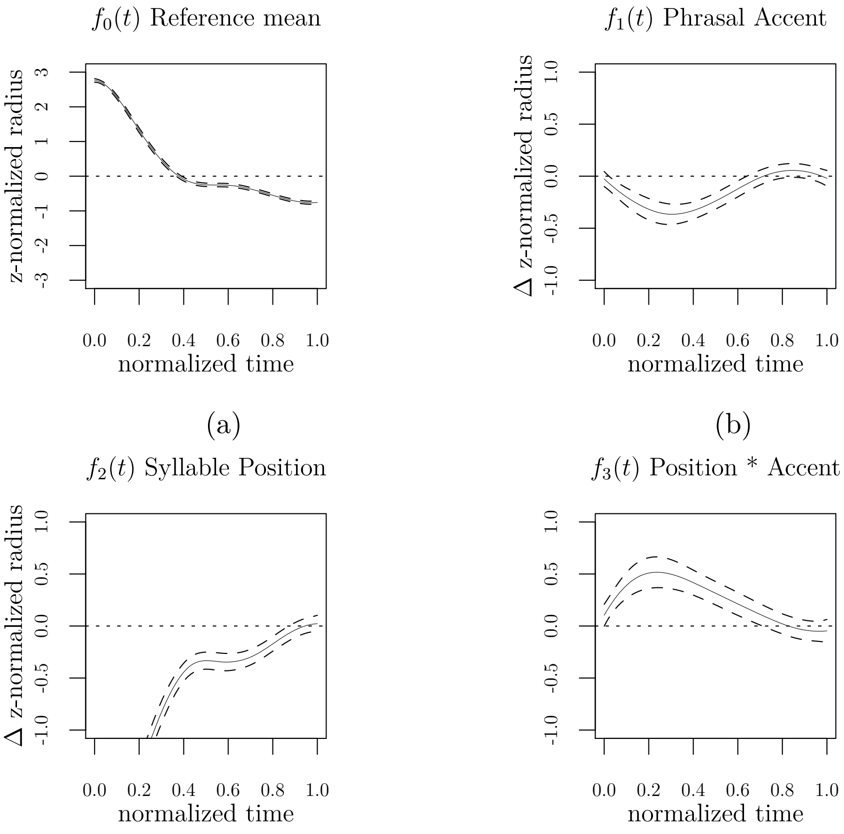
Figure 4.15: Radius measured at the anterior angle, capturing tongue height for the nucleus-coda comparison. For the nucleus condition the radius during acoustically segmented *plp* in ***plp**ap* is shown and for the coda condition the *elp* segment in ***pelp**ap* is shown. On the left the radius over normalized time, on the right z-normalized radius. A greater radius indicates a higher tongue position.

For the statistical analysis FLMM was used. The model consisted of two covariates, PHRASAL ACCENT and SYLLABLE POSITION, as well as their interaction. The reference mean corresponded to the unaccented condition where /l/ occupied the syllable nucleus (see also Table 4.2). The reference mean curve and covariate effect curves as well as the interaction effect curve are shown in Figure 4.16 and the summed effect curves are shown in Figure 4.17. As discussed in the previous chapter there is a significant effect of PHRASAL ACCENT from the beginning until around time point 0.6. Between these time points the tongue is significantly lower in accented condition compared to the unaccented condition.

SYLLABLE POSITION has a significant effect almost throughout the entire segment with the exception after time point 0.9. The strongest effect is at the beginning until around time point 0.4. Recall that we are comparing the *plp* segment for the nucleus condition and the *elp* segment for the coda condition. In the nucleus condition the tongue is still high during the initial /p/ because of the vowel /i/ preceding the target word. In the coda condition however, the tongue is lower for /e/. Between time point 0.4 and 0.9 the tongue position is also lower for the coda condition.

The interaction effect curve of PHRASAL ACCENT and SYLLABLE POSITION ($f_3(t)$) has an effect in the opposite direction of the covariate effect curve PHRASAL ACCENT ($f_1(t)$). As a result, while PHRASAL ACCENT had a significant lowering effect on tongue movement in nucleus condition, the effect is not existent for coda condition. In Figure 4.17 we can see that the red solid line is lower than the red dashed line until around time point 0.6. However, the blue lines are almost overlapping, with the solid blue line being slightly higher than the blue dashed line. Thus accent has an effect in the opposite direction for the two syllable positions. It is not possible to tell whether $f_1(t)$ and $f_3(t)$ cancel each other out completely so the difference between the dashed and solid blue lines is significant or only at trend level.

To summarize, the tongue at the anterior angle is higher when /l/ occupies the nucleus than when it occupies the coda. /l/ occupying the nucleus was affected by phrasal accent so that the tongue reached a lower point but moved upwards during the acoustically segmented /l/. In coda condition, the tongue front was slightly higher when accented than when it was not accented, which we did not expect.



(a) (b)

(c) (d)

Figure 4.16: Reference mean and covariate effects of the tongue movement captured at the anterior angle for the nucleus-coda comparison: *plp* sequence of *plpap* for the nucleus condition and *elp* sequence of *pelpap* for the coda condition. A greater radius indicates a higher tongue position.

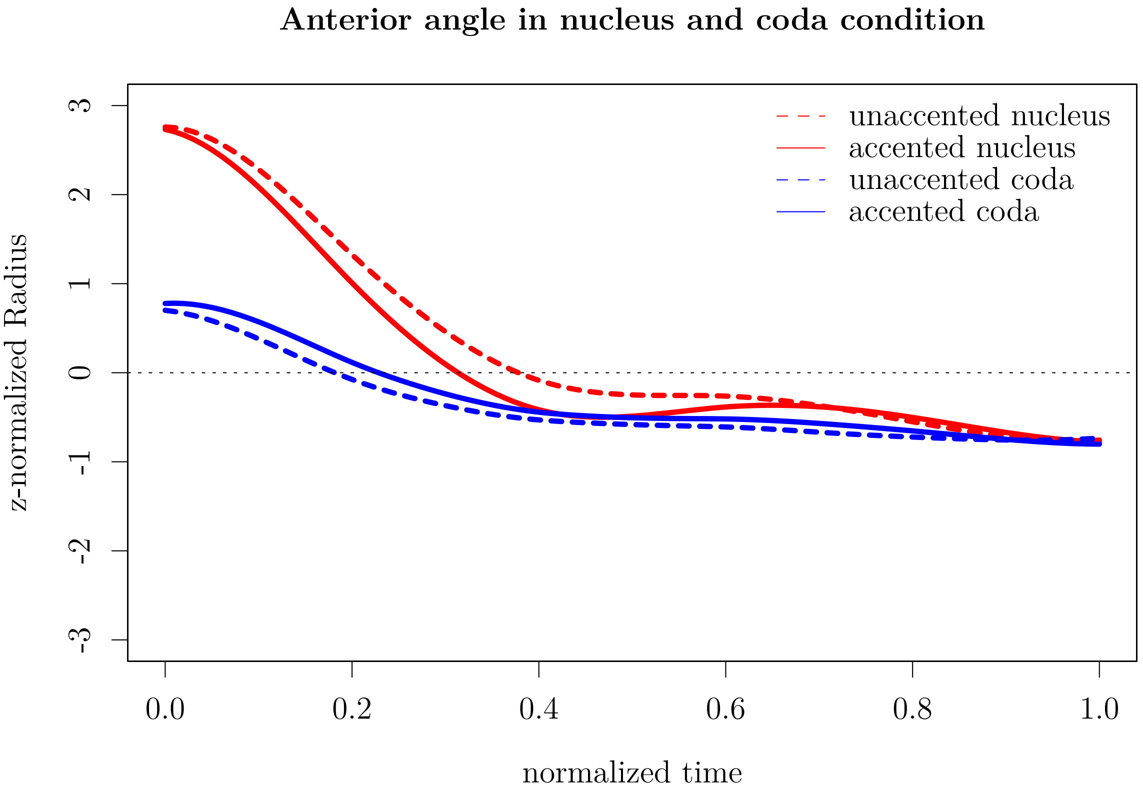


Figure 4.17: Summed effect curves of tongue movement in *plp* sequence of *plpap* and *elp* sequence of *pelpap* for the nucleus-coda comparison. A greater radius indicates a higher tongue position. The red dashed line, also the reference mean $f_0(t)$, corresponds to the unaccented nucleus condition, the red solid line to the accented nucleus condition ($f_0(t) + f_1(t)$), the blue dashed line to the unaccented coda condition ($f_0(t) + f_2(t)$) and the blue solid line to the accented coda condition ($f_0(t) + f_1(t) + f_2(t) + f_3(t)$).

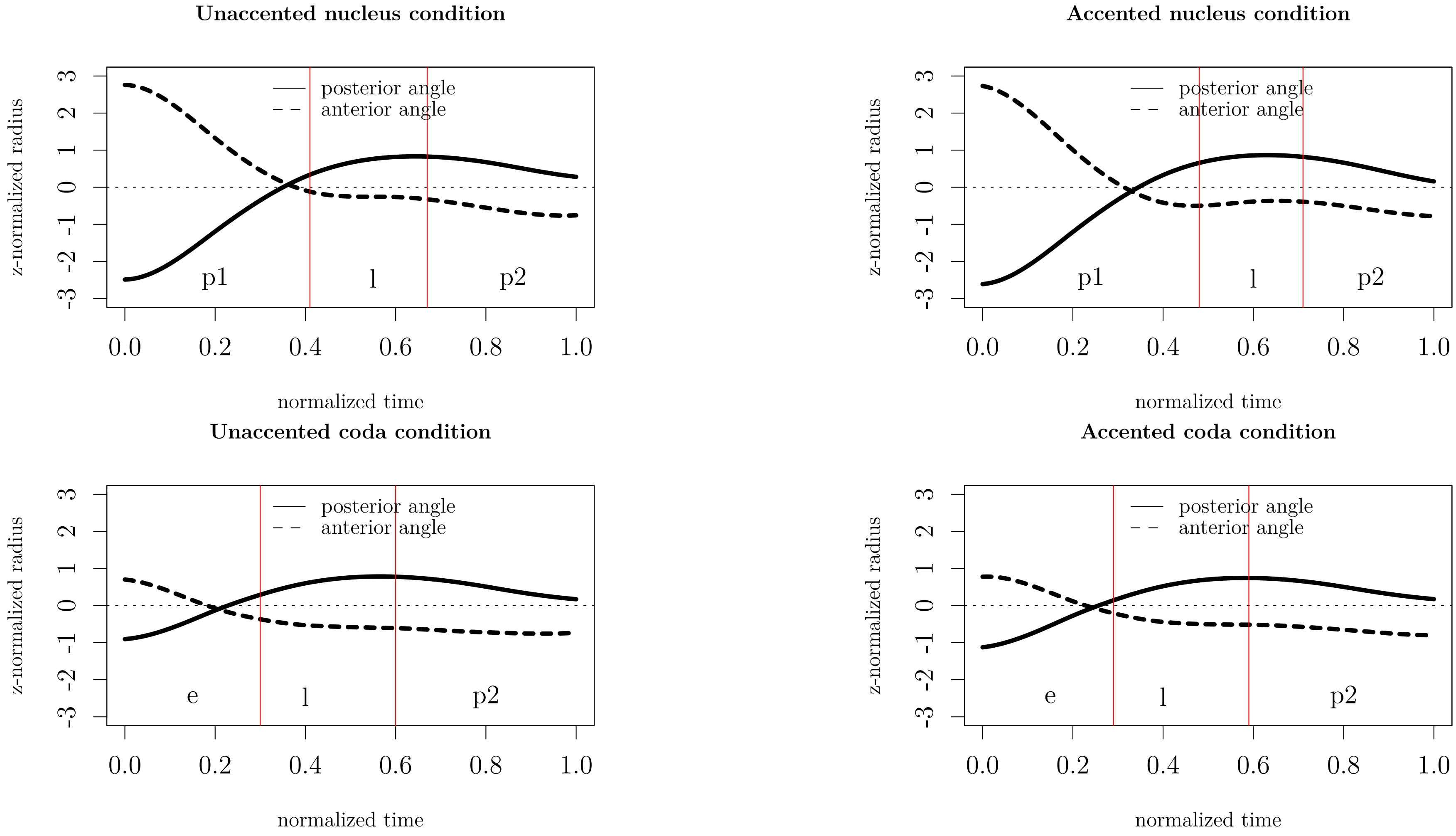


Figure 4.18: Summed effect curves of the tongue movement captured at the posterior and anterior radii for the nucleus-coda comparison. The red vertical lines represent the boundaries between the acoustically determined segments *p1* (if the following /l/ occupies the nucleus condition) or *e* (when /l/ occupies the coda condition), *l*, and *p2*. For the contours captured at the anterior angle a higher radius indicates a higher tongue position, at the posterior angle a higher radius indicates a more retracted tongue.

Summary of the articulatory data

We will recapitulate the predictions and findings of the articulatory data before we continue with the acoustic results.

Based on literature on American English in which the coda /l/ is usually described as more vowel like and the findings by Pouplier and Beňuš (2011) where /l/ in nucleus position patterned kinematically more with the coda position we expected to find greater differences in the onset-nucleus comparison than in the nucleus-coda comparison. In addition, we expected the effect of phrasal accent to differ depending on whether /l/ occupied the nucleus or not. Our predictions were confirmed for the most part. In onset condition /l/ was less retracted than in nucleus condition while the extent of retraction did not differ between the nucleus and coda condition. At the anterior angle we observed that the tongue was higher in the onset condition compared to the nucleus condition while in coda condition it was lower than in the nucleus condition. An effect of phrasal accent at the posterior angel capturing the tongue retraction was observed for the coda condition only. When /l/ occupied the coda the tongue was less retracted in the accented condition. The difference could have been conditioned by contextual differences. A stronger coarticulatory effect of the vocalic nucleus /e/ in accented condition could have caused a more fronted tongue also during the coda /l/. At the anterior angle the effect of phrasal accent was strongest for the nucleus condition in the form of a more dynamic movement. In the coda position the tongue movement did not differ for the two accent conditions, while in onset the tongue was lower when accented.

4.3.2 F1 and F2

We will now continue with the acoustic data. For the acoustic analysis we compare the formant trajectories during acoustically segmented /l/ only. This corresponds to the *l*-segments indicated in Figures 4.10 and 4.18 of the articulatory results. Like for the articulatory analysis, we will make two separate comparisons. First we will compare /l/ in nucleus position with the onset position, and then the nucleus position with the coda position. For each comparison, F1 and F2 during /l/ was extracted and FLMM was used to compare the curves, which resulted in four models (two models per comparison). Each model consisted of two covariates, PHRASAL ACCENT and SYLLABLE POSITION, as well as their interaction (see also

Table 4.2). First the results of the onset-nucleus comparison will be presented, then the results of the nucleus-coda comparison. When presenting the results, we will also refer to the articulatory results. A smaller distance between F1 and F2, and particularly a lower F2 are associated with a darker /l/. F1 is associated with the openness of the oral cavity. A higher F1 indicates a lower tongue and jaw position. A lower F2 generally indicates a more retracted tongue.

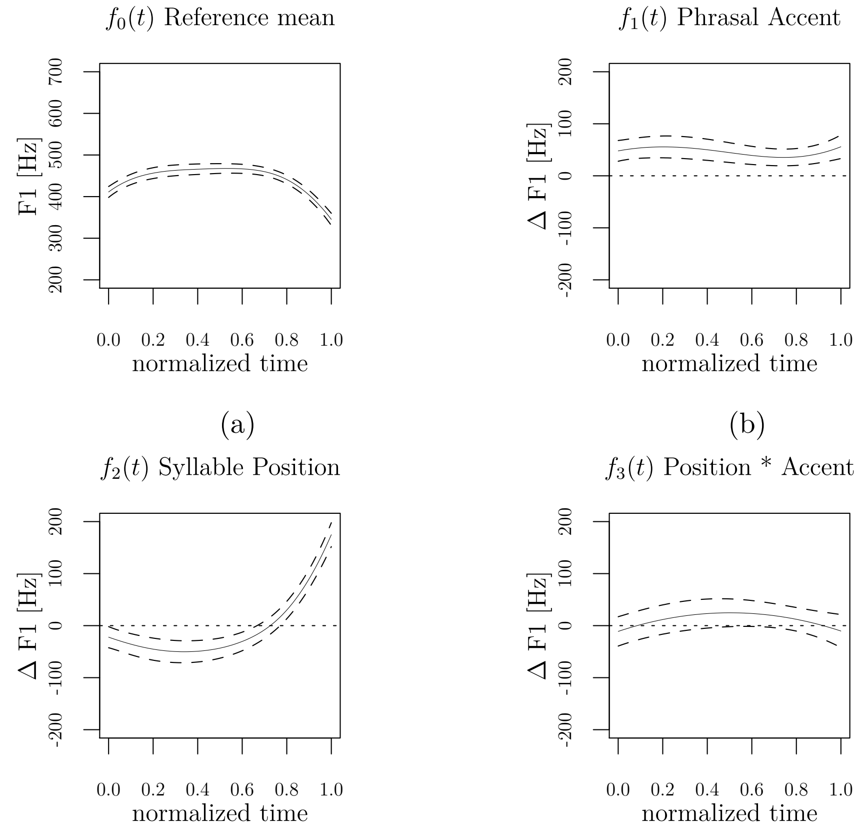
Onset and Nucleus position

F1 We begin with the results for F1. The time normalized F1 contour of the acoustically segmented /l/ in *plpap* (nucleus condition) and *plepap* (onset condition) in two accent conditions were analysed using FLMM. The model consisted of two covariates, PHRASAL ACCENT and SYLLABLE POSITION, and their interaction. The reference mean curve, covariate effect curves and the interaction effect curve are shown in Figure 4.19 and the summed effect curves in Figure 4.20. The reference mean $f_0(t)$ (Figure 4.19 (a)) corresponds to F1 of /l/ in nucleus position in unaccented condition. PHRASAL ACCENT $f_1(t)$ had a significant effect throughout the entire /l/ segment (Figure 4.19 (b)). By comparing the red dashed line for the unaccented nucleus condition and the red solid line for the accented nucleus condition in Figure 4.20 we can see that when accented, F1 is higher, suggesting a more open oral cavity. As already discussed in the previous chapter the articulatory data captured at the anterior angle showed that the tongue was indeed lower when the target word was accented, but only at the beginning of /l/ (see results for F1 in Section 3.3.2). F1 not only reflects tongue height but also jaw position, so we can assume that the jaw is lower in accented condition.

SYLLABLE POSITION $f_2(t)$ (Figure 4.25 (c)) had a significant effect throughout the entire segment except for at around time point 0.75 where it crosses zero. F1 is lower for the onset condition until around time point 0.75, after which it rises rapidly. This indicates a more narrow oral cavity until time point 0.75 in onset condition. The articulatory data also showed that the tongue was higher in onset condition compared to nucleus condition during the /l/ segment ($f_2(t)$ in Figure 4.8). In the onset condition the tongue is not only higher during /l/ but also rises further towards the end of the target sequence. Recall that /l/ in nucleus condition was flanked by voiceless bilabial plosives on both sides, while /l/ in onset condition was followed by /e/. So /l/ in nucleus position is followed by a complete closure of the oral cavity and consequently a lower F1, while a more open oral cavity is expected after /l/ in onset position and consequently a higher F1. The tongue rises towards the end of /l/ in onset position, because at the angle at which tongue height was measured the tongue has a higher position in /e/ than in /l/. At this angle we were observing a low tongue position during /l/ which is reached by tongue curling. It is likely that the jaw is lower during /e/, which also contributes to a rising F1 at the end of /l/ in the onset condition.

The interaction effect of SYLLABLE POSITION and PHRASAL ACCENT $f_3(t)$ (Figure 4.25 (d)) is not significant, but indicates a trend that between time point 0.4 and 0.7 the effect of PHRASAL ACCENT is stronger on /l/ in onset condition. In Figure 4.20 the summed effect curves are shown. The reference mean curve, corresponding to the unaccented nucleus condition is visualized using a red dashed line. The summed effect curve corresponding to the accented nucleus condition, obtained by adding the effect of PHRASAL ACCENT on the reference mean, is visualized using a red solid line. When we compare the distance between the two red lines with the distance between blue dashed line, the distance between the red lines from around time point 0.4 to 0.7 is slightly smaller than between the two blue lines. This only partially reflects the articulatory data. As visualized in Figure 4.9, the tongue is generally lower for the nucleus condition than for the onset condition. A lower tongue contributes to a wider oral cavity which is reflected in a higher F1. However, the effect of phrasal accent was weaker on the tongue movement in the onset condition. Possibly phrasal accent has a stronger effect on the jaw for /e/ and consequently the jaw is also lower during /l/ in onset condition for the accented condition.

In sum, F1 was higher for /l/ in nucleus condition than in onset condition. A higher F1 indicates a wider oral cavity. This is consistent with the articulatory data observed at the anterior angle, where the tongue was lower for /l/ in nucleus condition. In accented condition F1 is higher for both, the nucleus as well as onset condition. This matches the articulatory data as well, as we observed a lower tongue at the anterior angle in accented condition for both, /l/ in nucleus as well as in onset. However, for the articulatory data the effect of accent was stronger for the nucleus condition than for the onset condition, while for the acoustic data the effect of accent did not differ significantly as an effect of syllable position.



(a) (b)

(c) (d)

Figure 4.19: Reference mean and covariate effects for F1 for the onset-nucleus comparison.

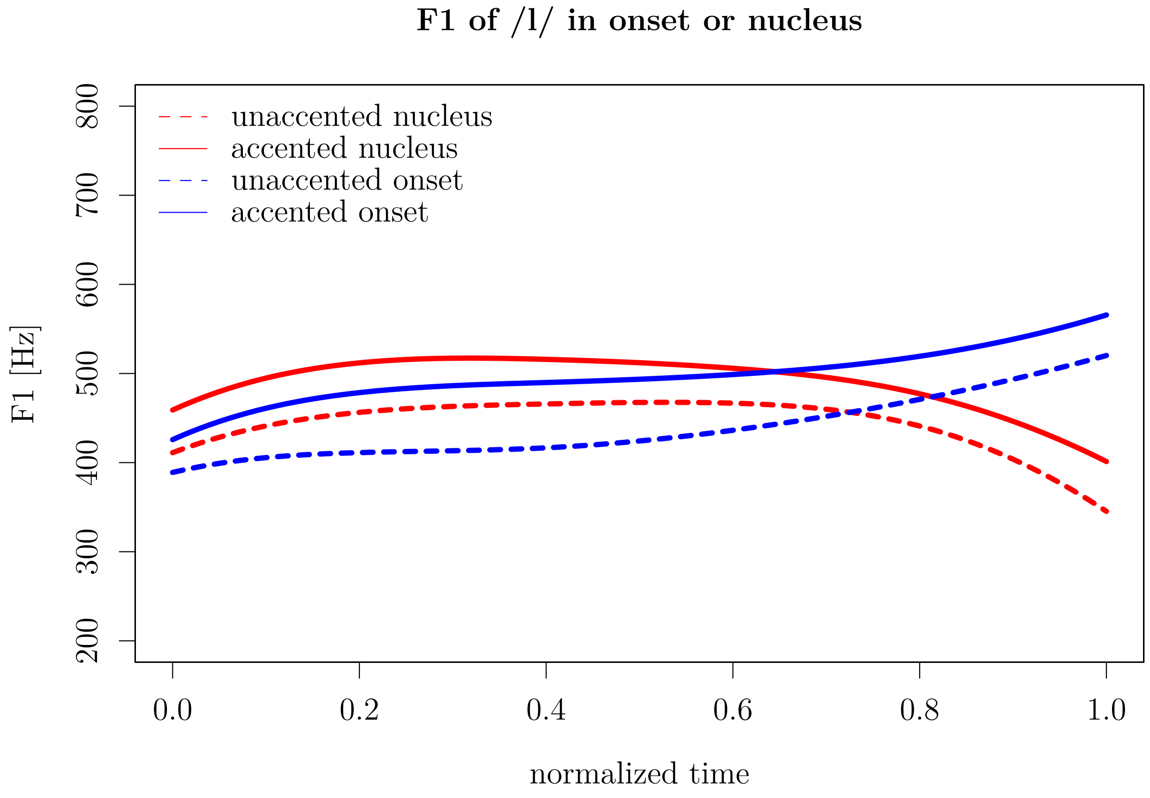


Figure 4.20: Summed effect curves for F1 for the onset-nucleus comparison. The red dashed line, also the reference mean $f_0(t)$, corresponds to the unaccented nucleus condition, the red solid line to the accented nucleus condition ($f_0(t) + f_1(t)$), the blue dashed line to the unaccented onset condition ($f_0(t) + f_2(t)$) and the blue solid line to the accented onset condition ($f_0(t) + f_1(t) + f_2(t) + f_3(t)$).

F2 We continue with the results for F2 of /l/ in onset or nucleus positions. The reference mean, the covariate effect curves and the interaction effect curves are presented in Figure 4.21 and the summed effect curves in Figure 4.22. The reference mean $f_0(t)$ (Figure 4.21 (a)) corresponds to the unaccented nucleus condition. Recall that for /l/ a lower F2 value indicates generally a more retracted tongue back. In the unaccented nucleus condition F2 is a linearly falling slope. Although the main retraction movement captured at the posterior angle could be observed during the acoustically segmented /p/, there was also some retraction movement during /l/. For F2, PHRASAL ACCENT $f_1(t)$ (Figure 4.21 (b)) had a significant effect only at the very end. As a result, in accented condition, F2 did not fall linearly during the entire /l/ segment, but slowed down and kept a stable position at the end, as can be seen when comparing the red dashed and solid lines in Figure 4.22.

SYLLABLE POSITION $f_2(t)$ (Figure 4.21 (c)) had a significant rising effect from around time point 0.6 until the end. F2 of /l/ in the onset condition begins at roughly the same height as for the nucleus condition. However, in the onset condition F2 begins to rise at around time point 0.4, so it does not reach F2 values as low as for the nucleus condition. The articulatory data of the posterior part of the tongue also showed that in the onset condition the tongue was less retracted during the entire segment. The acoustic data suggests that the movement towards /e/, which has a more fronted tongue position, begins already at around time point 0.4. In contrast, /l/ in nucleus position is followed by a bilabial consonant which does not require tongue fronting.

The interaction of SYLLABLE POSITION and PHRASAL ACCENT $f_3(t)$ (Figure 4.21 (d)) was also significant. While in the nucleus condition, F2 did not differ as an effect of PHRASAL ACCENT, in the onset condition F2 is significantly lower at the beginning until around time point 0.3. This suggests that for the onset conditions, at the beginning, the tongue is more retracted when accented. The summed effect curves of the tongue movement captured at the posterior angle show a difference for the two accent conditions for the onset condition (comparing the two blue lines in Figure 4.5), but the effect was not significant. It is possible that the effect was not captured at the chosen angle.

To summarize, for F2 the most prominent effect was the effect of syllable position, which can be attributed to the different segmental contexts. In the unaccented onset condition the F2 value does not reach a F2 value as low as in the other conditions. Phrasal accent had an

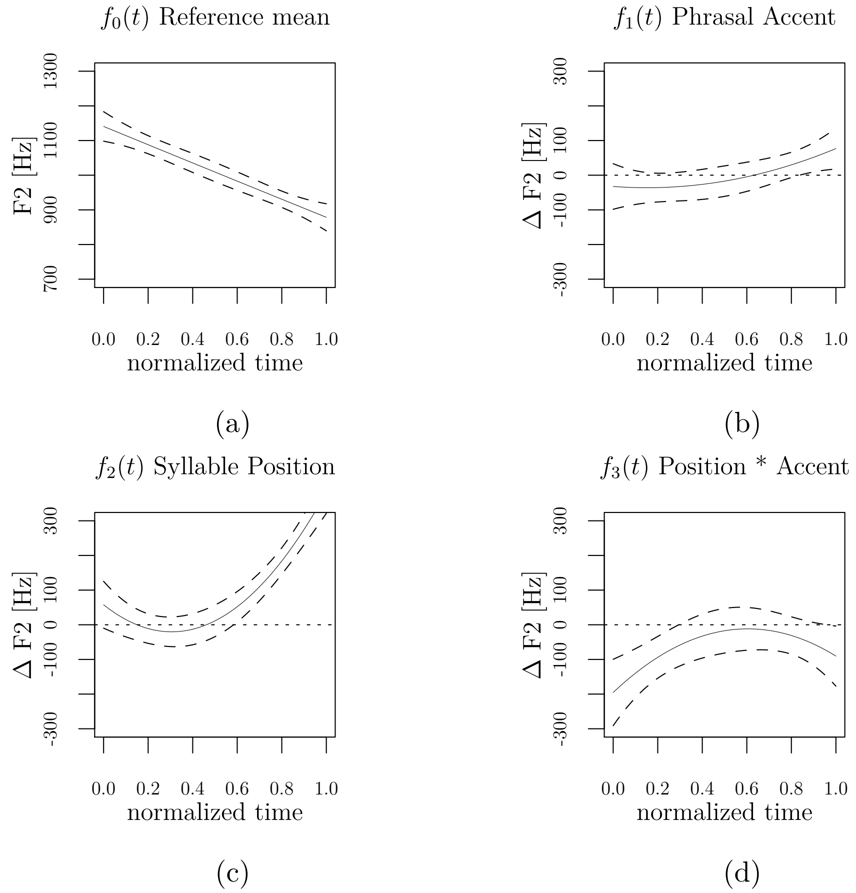


Figure 4.21: Reference mean and covariate effects for F2 during /l/ in *plpap* (nucleus condition) and *plepap* (onset condition).

effect on the F2 of /l/ in the onset condition, but not in the nucleus condition.

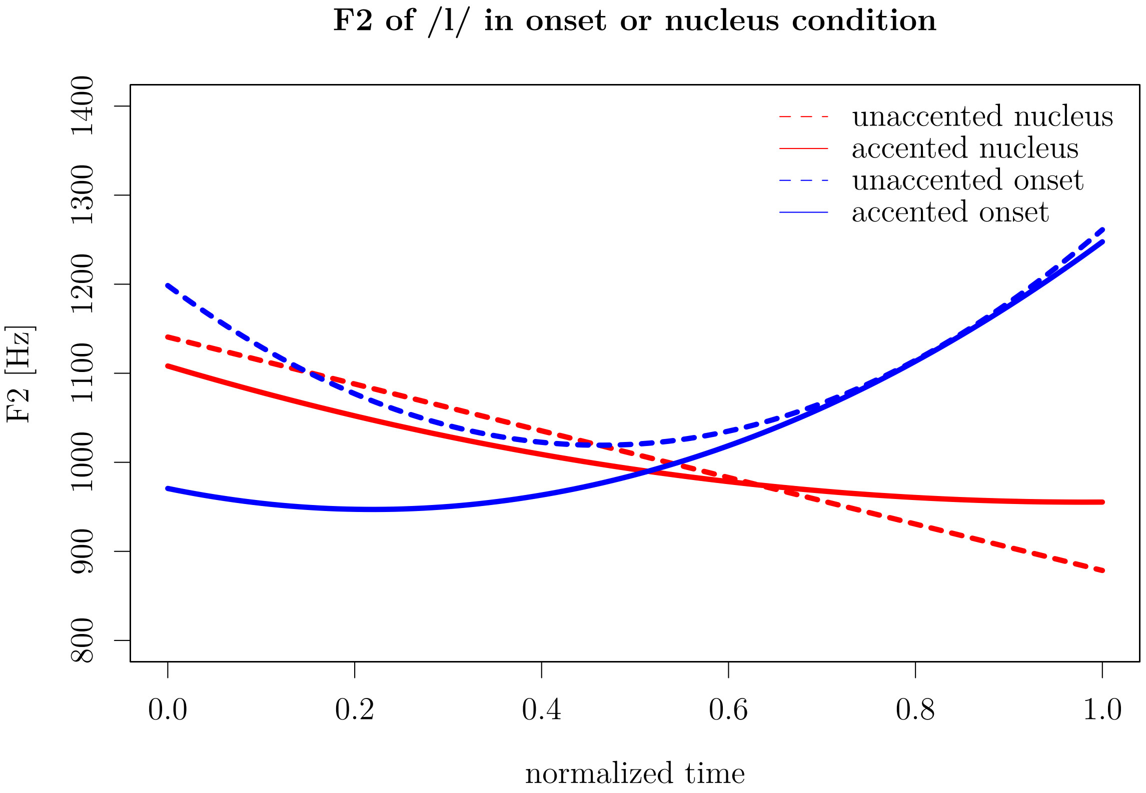


Figure 4.22: Summed effect curves of F2 during /l/ in nucleus and onset positions. The red dashed line, also the reference mean $f_0(t)$, corresponds to the unaccented nucleus condition, the red solid line to the accented nucleus condition ($f_0(t) + f_1(t)$), the blue dashed line to the unaccented onset condition ($f_0(t) + f_2(t)$) and the blue solid line to the accented onset condition ($f_0(t) + f_1(t) + f_2(t) + f_3(t)$).

Nucleus and Coda position

Now results of the nucleus-coda comparison will be presented. F1 and F2 during the acoustically segmented /l/ in *plpap* (nucleus condition) and *pelpap* (coda condition) in two accent conditions will be compared. The corresponding articulatory data can be seen in Figure 4.18, during the interval labelled *l*. Time normalized data of F1 and F2 during acoustically determined /l/ was used to carry out the FLMM analyses. The model consisted of two covariates, SYLLABLE POSITION and PHRASAL ACCENT, and their interaction effect. The reference mean corresponds to the unaccented nucleus condition. What should be kept in mind is that the segmental context at the left edge is different for the two syllable positions (for the onset-nucleus comparison it was the right edge that differed). /l/ in nucleus context is preceded by /p/, while /l/ in coda context is preceded by /e/.

F1 In general F1 is correlated with the openness of the oral cavity. A lower jaw and weaker apical constriction contribute to a higher F1.

The reference mean curve, covariate effect curves and interaction effect curve are shown in Figure 4.23 and the summed effect curves are shown in Figure 4.24. The reference mean $f_0(t)$ (Figure 4.23 (a)) is identical to the reference mean curve discussed in the onset-nucleus comparison. It corresponds to the unaccented nucleus condition. The covariate effect of PHRASAL ACCENT $f_1(t)$ (Figure 4.23 (b)) is significantly above zero. Thus, when accented, F1 is higher during the entire /l/-segment in nucleus position. The difference is quite stable, and as can be seen in Figure 4.24 the red dashed line corresponding to the unaccented nucleus condition and the red solid line corresponding to the accented nucleus condition run almost parallel to each other.

SYLLABLE POSITION $f_2(t)$ (Figure 4.23 (c)) has a significant effect during the first half of the /l/ segment. This can be attributed to the differing preceding contexts for the two syllable positions. When /l/ occupies the nucleus condition, F1 points to the low locus for the preceding /p/. However, when /l/ occupies the coda position the preceding segment is /e/, which has a higher F1.

The interaction of PHRASAL ACCENT and SYLLABLE POSITION $f_3(t)$ (Figure 4.23 (d)) has a significant effect at the very end (after around time point 0.8) of /l/. PHRASAL ACCENT has an effect on F1 so that it is higher in accented condition than in unaccented condition, and for /l/ in coda position this effect is even larger at the end of the segment. It can be seen clearly looking at the summed effect curves in Figure 4.24 that after time point 0.8 the difference between the two blue lines which correspond to the coda condition is larger than the difference between the two red lines which correspond to the nucleus condition. The interaction effect $f_3(t)$ is significant also around time point 0.3 which works in the direction opposite to the effect of PHRASAL ACCENT $f_1(t)$. In Figure 4.24 we can see that around time point 0.3 the difference between the blue dashed line and the blue solid line is much smaller than between the two red lines.

To summarize, at the beginning of the /l/ segment F1 was higher for the coda condition than for the nucleus condition. In accented condition F1 was higher for /l/ in nucleus as well as coda condition. This only partly matches the articulatory data. At the anterior angle, the tongue was indeed for the most part lower in coda condition than in nucleus condition. However, a lower tongue position for the accented condition was observed only for /l/ occupying the nucleus at the beginning. For /l/ in coda, the effect of accent was minimal, but actually in the opposite direction, so that the tongue was higher in the accented condition. This was reflected in the acoustic results only at the beginning until time point 0.4 and for the accented condition after time point 0.8. We can assume that F1 is mainly affected by the jaw.

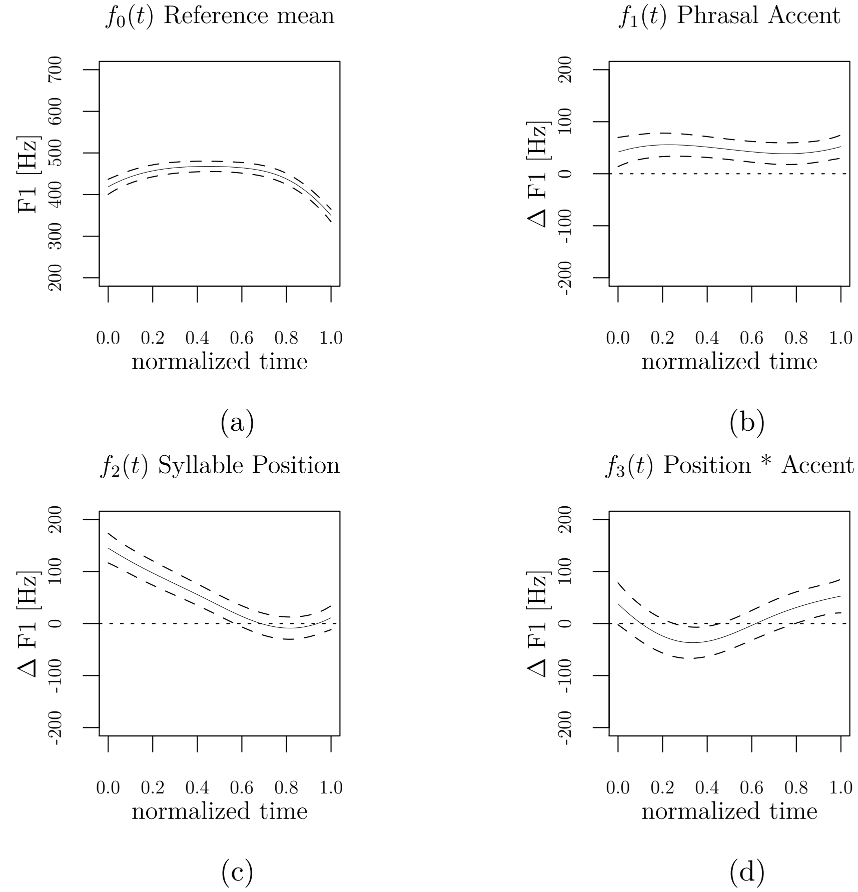


Figure 4.23: Reference mean and covariate effects for F1 during /l/ of *plpap* (nucleus condition) and *pelpap* (coda condition).

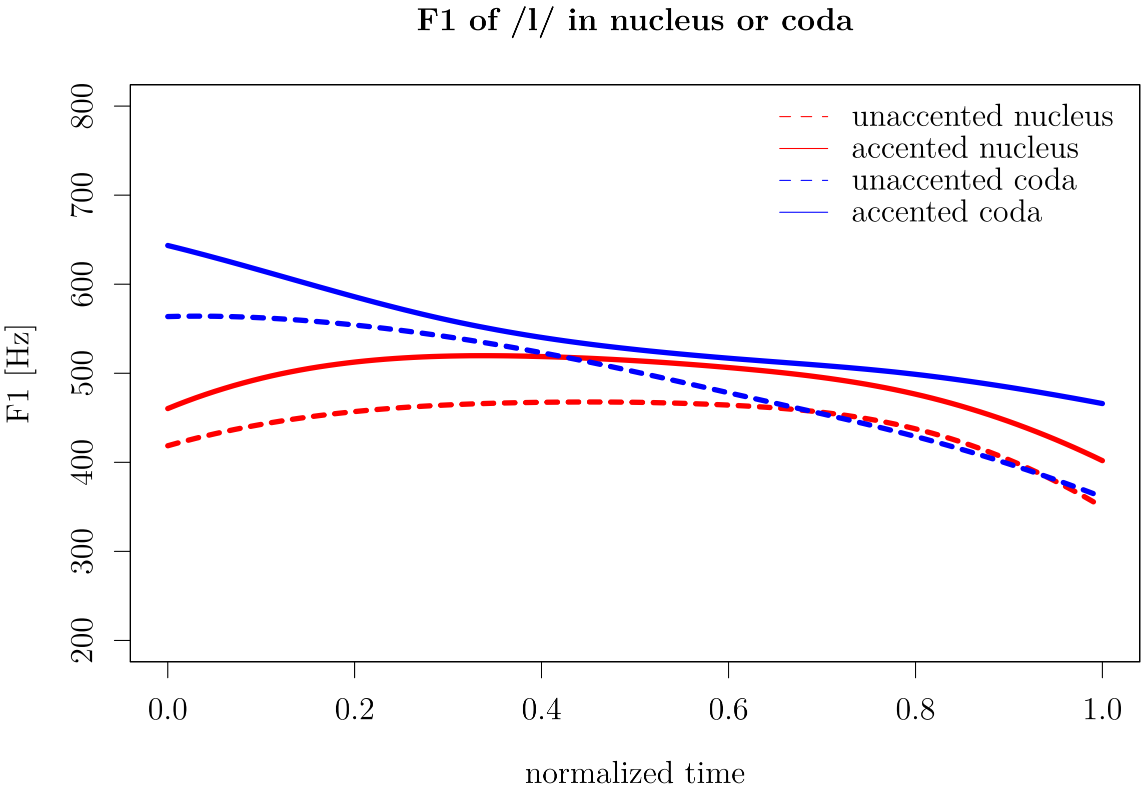


Figure 4.24: Summed effect curves of F1 during /l/ in nucleus and coda positions. The red dashed line, also the reference mean $f_0(t)$, corresponds to the unaccented nucleus condition, the red solid line to the accented nucleus condition ($f_0(t) + f_1(t)$), the blue dashed line to the unaccented coda condition ($f_0(t) + f_2(t)$) and the blue solid line to the accented coda condition ($f_0(t) + f_1(t) + f_2(t) + f_3(t)$).

F2 The reference mean curve, covariate effect curves and interaction effect curve of F2 for the nucleus-coda comparison are shown in Figure 4.25. A lower F2 indicates a more retracted tongue. The reference mean curve $f_0(t)$ (Figure 4.25 (a)) corresponds to the F2 of /l/ in nucleus position in unaccented condition. The effect of PHRASAL ACCENT $f_1(t)$ (Figure 4.25 (b)) is not significant which also reflects the articulatory data.

SYLLABLE POSITION $f_2(t)$ (Figure 4.25 (c)) has a significant effect at the beginning until around time point 0.2 and between time points 0.5 and 0.8. The effect can be better understood when we look at the summed effect curves in Figure 4.26. The red dashed line corresponds to the reference mean, the unaccented nucleus condition. The covariate effect curve for SYLLABLE POSITION $f_2(t)$ begins above zero, indicating that F2 in coda condition is higher at the beginning. The difference is significant until around time point 0.2 but the covariate effect curve continues to fall until around time point 0.6 and then raises again. F2 is significantly lower in coda condition from around time point 0.5 to 0.8. When we look at the summed effect curves in Figure 4.26 and compare the blue dashed line, corresponding to the unaccented onset condition, with the red dashed line, we see that the blue dashed line begins at a higher F2 and has a steeper falling slope than the red dashed line. It crosses the red dashed line at around time point 0.3 and keeps declining until around time point 0.6. It then stays at that position until the end, while the red dashed line keeps declining. It suggests that at the beginning of the acoustically segmented /l/ the tongue is more retracted for /l/ in nucleus condition but a more retracted tongue position is reached later during /l/ in coda position. The articulatory data captured at the posterior angle did not show that in coda position the tongue was more retracted. However, due to time normalization it is difficult to tell whether the target for the dorsal constriction is reached earlier for the coda condition. We will look at the timing data in Section 4.3.3.

The interaction effect curve $f_3(t)$ (Figure 4.25 (d)) reveals that PHRASAL ACCENT has a significantly different effect on /l/ in coda condition than on /l/ in nucleus condition. The effect of PHRASAL ACCENT $f_1(t)$ is not significant. For /l/ in the coda position, F2 is higher when accented until around time point 0.8. Only at the very end F2 is lower when accented. The acoustic data suggest that in the coda condition the tongue is less retracted at the beginning of /l/ compared to the nucleus condition. This makes sense considering the fact that /l/ in the coda condition is preceded by /e/, for which a more fronted tongue position is required than for /l/, while in the nucleus condition /l/ is preceded by /p/, for which no specific tongue position is required and during which the target position for /l/ can already be reached. Furthermore, for /l/ in coda position it can be assumed that the effect of the preceding vowel is stronger in accented condition resulting in a less retracted tongue and consequently a higher F2. After time point 0.9 $f_3(t)$ has a significant lowering effect.

In sum, a significant difference for the two syllable positions was observed for F2. In coda condition, F2 is significantly higher than in nucleus condition, which is probably due to the contextual difference. In coda condition /l/ is preceded by /e/, which has a higher F2, while in nucleus condition /l/ is preceded by /p/ for which lower loci are characteristic. Later during the acoustically segmented /l/ F2 is lower for the unaccented coda condition compared to the unaccented nucleus condition. However, we did not observe a more retracted tongue for the coda condition at the posterior angle. While phrasal accent did not affect F2 in nucleus condition, F2 was significantly higher in coda condition until time point 0.7. At the posterior angle we also observed a less retracted tongue for the accented coda condition compared to the unaccented coda condition.

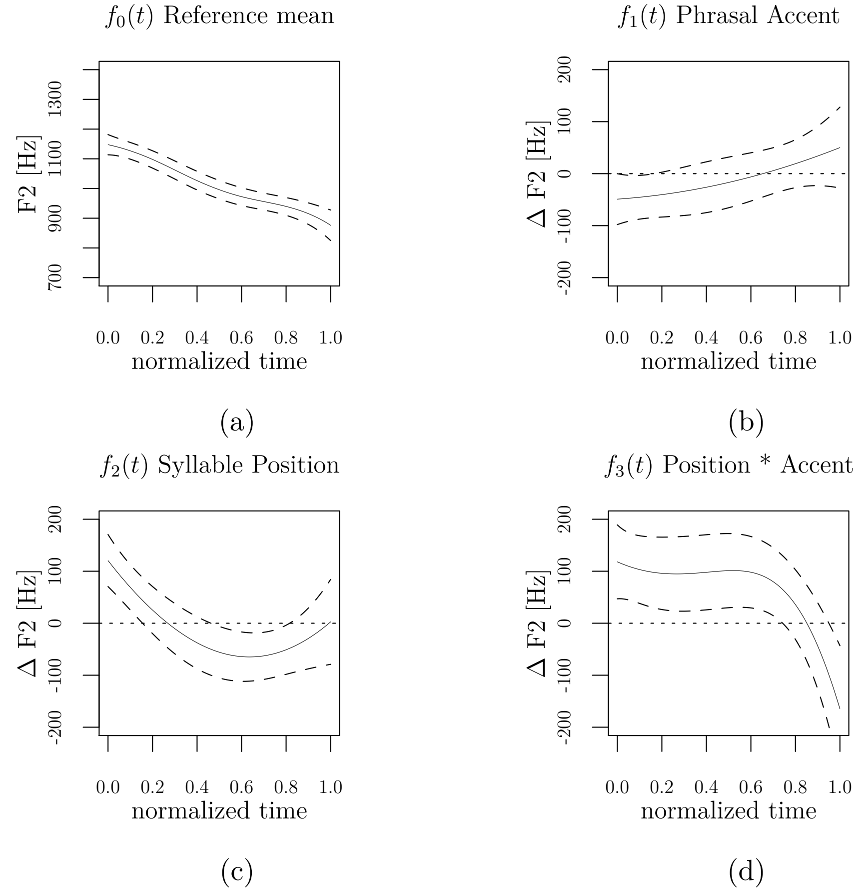


Figure 4.25: Reference mean and covariate effects for F2 during /l/ in *plpap* (nucleus condition) and *pelpap* (coda condition).

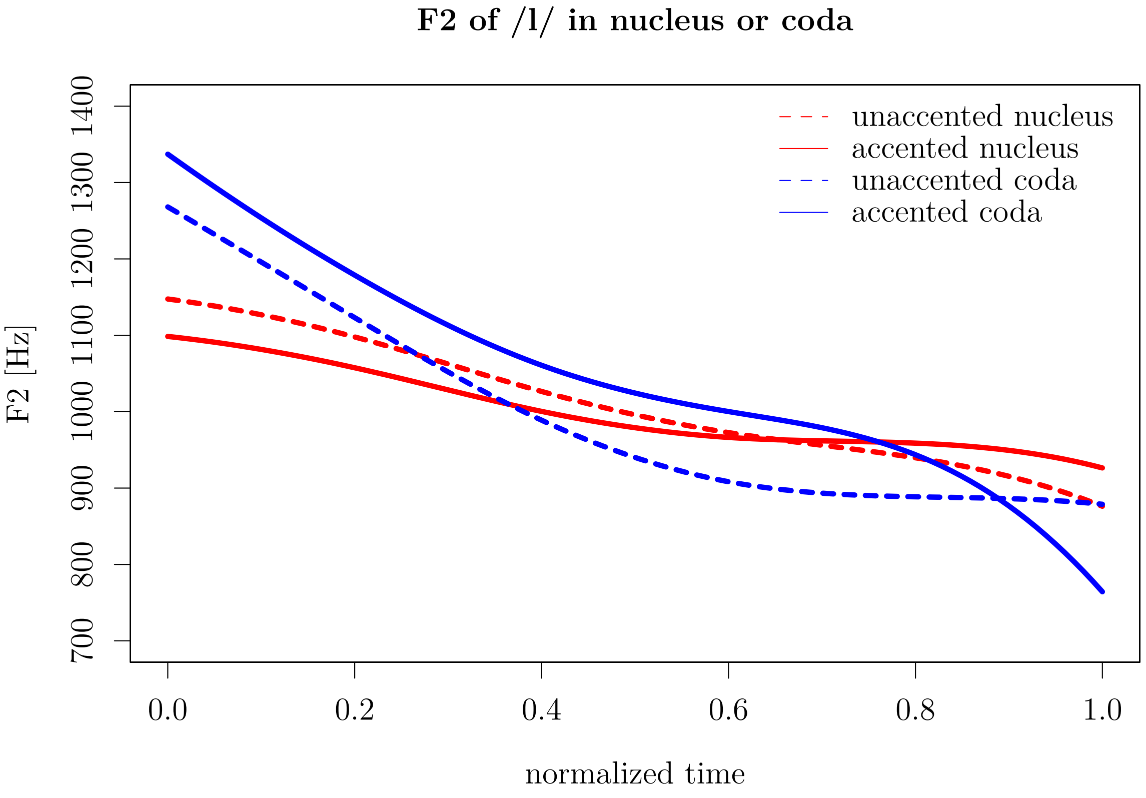


Figure 4.26: Summed effect curves for F2 during /l/ in nucleus and coda positions. The red dashed line, also the reference mean $f_0(t)$, corresponds to the unaccented nucleus condition, the red solid line to the accented nucleus condition $(f_0(t) + f_1(t))$, the blue dashed line to the unaccented coda condition $(f_0(t) + f_2(t))$ and the blue solid line to the accented coda condition $(f_0(t) + f_1(t) + f_2(t) + f_3(t))$.

Summary of the acoustic data

In contrast to the articulatory data which was measured also during the acoustically defined segments flanking /l/, the acoustic data were measured only during /l/.

Most differences in F1 and F2 between the different syllable positions can be related to the different segmental contexts. Our acoustic data showed a clear effect of phrasal accent on F1, which can be related to jaw height. Regardless of the syllable position /l/ occupied, F1 was higher in the accented condition than in the unaccented, indicating a wider oral cavity in the accented condition. Interestingly, for the most part the effect did not differ depending on the syllable position. F2 of /l/ in nucleus position was not affected significantly by phrasal accent. However, if /l/ occupied the onset position F2 was significantly lower when accented, while for /l/ in the coda condition F2 was significantly higher when accented.

4.3.3 Timing

Two aspects of temporal organization will be addressed. First, results of the within-segment temporal organization of the apical and dorsal gesture of /l/ will be presented. Then the between-segment temporal organization will be presented where temporal coordination of the two gestures of /l/ with respect to the preceding or following consonant was examined.

Within-segment timing

Pouplier and Beňuš (2011) reported that their preliminary results did not show an effect of syllable position on the temporal coordination of the apical and dorsal gesture of /l/. We wanted to know additionally, whether the gestural coordination is affected by phrasal accent. As outlined in Section 4.1.3 we expect that the lag between the dorsal and apical gesture of /l/ in nucleus is greater when accented to allow for a longer sonorous segment.

To test the effect of syllable position and phrasal accent on the temporal lag between the achievement of the apical constriction and dorsal retraction, linear mixed effects analysis was performed using the lme4-package (Bates et al., 2014) in R. The data is visualized in Figure 4.27. There is a lot of variation, but in most cases the lag values are positive, which, unexpectedly, means that the apical constriction precedes the dorsal retraction extremum. In Figure 4.1 some examples of the movement contours are shown. In the case of the

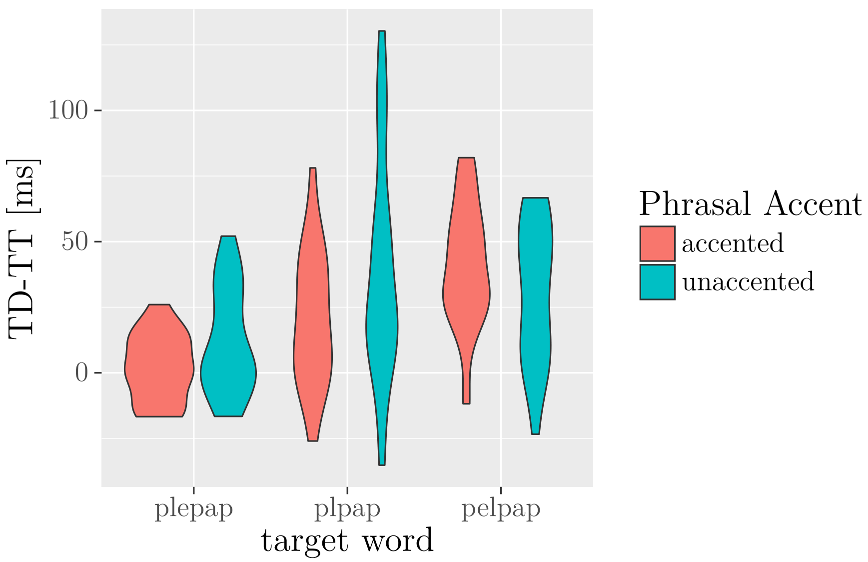


Figure 4.27: Temporal lag between the achievement of the apical constriction and the dorsal retraction of /l/ in *plepap*, *plpap* and *pelpap* which correspond to the onset, nucleus and coda position, respectively. The data for the accented condition are shown in red and for the unaccented condition in blue. A positive value means that the apical gesture precedes the dorsal gesture.

nucleus condition thus, during *plp* (shown in Figure 4.1 (a) and (b)), the target of the dorsal retraction gesture is reached slightly later than the target of the apical constriction. However, the movement of the dorsal gesture approaches the target more slowly. The movement of the apical gesture could be initiated later and still reach the target earlier than the dorsal gesture. In any case, the smaller the lag value, the earlier the dorsal target is achieved relative to the apical target.

PHRASAL ACCENT (two levels: accented, unaccented) and SYLLABLE POSITION (three levels: onset - *plepap*, nucleus - *plpap*, coda - *pelpap*) were fixed factors, with REPETITION and SPEAKER (with by-SPEAKER intercept and slope for PHONEMIC QUANTITY and NUCLEUS TYPE) as random factors. P-values were then obtained by a likelihood-ratio-test in which the final model was compared to a model with or without the effect in question.

The final model consisted of PHRASAL ACCENT (two levels: accented, unaccented), SYL-

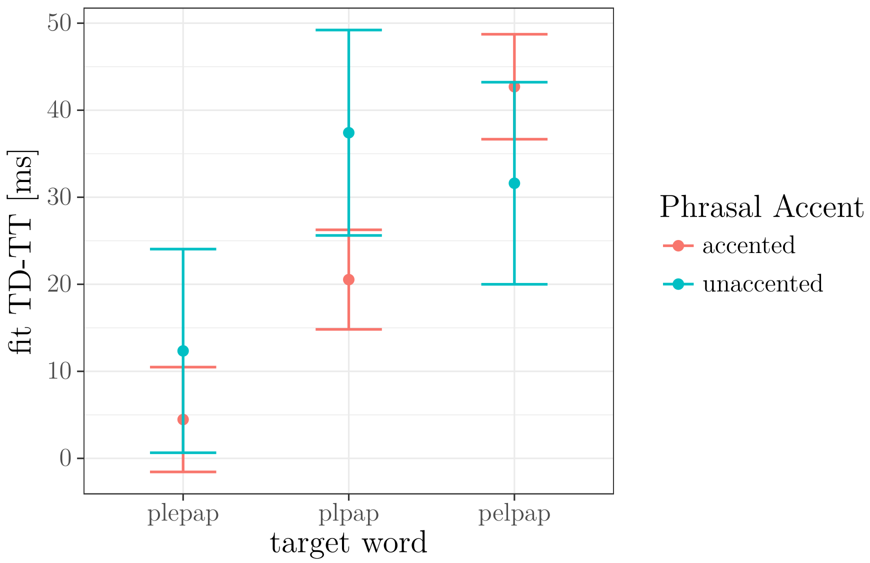


Figure 4.28: Visualized effects of PHRASAL ACCENT and SYLLABLE POSITION on the temporal lag between the apical constriction and the dorsal retraction achievement of /l/. The plot shows the mean and standard error for each condition. The data for the accented condition are shown in red and for the unaccented condition in blue. A positive value means that the apical gesture precedes the dorsal gesture.

LABLE POSITION (three levels: onset, nucleus, coda) and their interaction as fixed effects and REPETITION and SPEAKER (with by-SPEAKER intercept and slope for PHRASAL ACCENT) as random factors. Figure 4.28 shows the mean and standard error of the temporal lag between the apical constriction and dorsal retraction achievement for each condition as fitted by the linear mixed effect model. For /l/ in the three target words *plepap*, *plpap* and *pelpap* where it occupies the onset, nucleus or coda, respectively, the accented condition is visualized in red and the unaccented in blue. Effects were extracted using the effects package (Fox, 2003) in *R*.

There was no significant effect of PHRASAL ACCENT ($\chi^2[1] = 0.25, p > 0.1$) but the effect of SYLLABLE POSITION ($\chi^2[2] = 29.34, p < 0.001$) and the interaction of PHRASAL ACCENT and SYLLABLE POSITION were significant ($\chi^2[2] = 8.73, p < 0.05$). The post-hoc Tukey test revealed that in accented condition, the temporal lag in onset and nucleus condition were not significantly different and in the unaccented condition the difference between the nucleus and coda position was not significant. So in accented condition, the lag for /l/ in coda was significantly greater than the lag in onset and nucleus condition. The target of the dorsal gesture was reached later relative to the achievement of the target of the apical gesture in all conditions, but the lag was significantly greater in coda condition. In unaccented condition the nucleus patterned with the coda and the two conditions differed significantly from the onset condition. The lag for the nucleus and coda condition was significantly greater than for the onset condition. In accented condition the temporal lag between the apical constriction achievement and dorsal retraction achievement was $4.5ms \pm 6.0$ in onset condition, $20.5ms \pm 5.7$ in nucleus condition and $42.7ms \pm 6.0$ in coda condition. In unaccented condition the temporal lags were $12.3ms \pm 11.7$ in onset, $37.41ms \pm 11.8$ in nucleus and $31.6ms \pm 11.6$ in coda condition. Although in the accented condition the onset and coda condition were not significantly different according to the post-hoc Tukey test, the results visualized in Figure 4.28 (accented condition is presented in red) show that the pattern for the nucleus condition is between the patterns of the onset and coda condition. The small number of tokens for the timing analysis could explain why in the post-hoc test the difference between the onset and nucleus in accented condition was not significant.

Surprisingly, in Slovak the apical constriction preceded the target achievement of the dorsal gesture in all conditions. In addition, we hypothesized that the dorsal gesture would precede the apical gesture to a greater extent in nucleus than in onset condition, which was not verified.

Between-segment timing

To compare the temporal coordination of the two /l/-gestures with the preceding or following consonant, four separate analyses were performed. The nucleus condition was compared separately with the onset and coda condition. For the onset-nucleus comparison the timing of the achievement of the apical or dorsal gesture was measured with respect to the burst of initial /p/ (end of the voiceless segment). A positive value means that the target of the dorsal or apical gesture of /l/ are achieved after the burst. For the nucleus-coda comparison the timing of the achievement of the two gestures was measured with respect to the acoustically defined beginning of the second /p/. Here, a positive value means that the target of the dorsal or apical gesture of /l/ are achieved before the labial closure of /p/. In each analysis PHRASAL ACCENT (two levels: accented, unaccented) and SYLLABLE POSITION (two levels: either onset - *plepap* and nucleus - *plpap*, or nucleus and coda - *pelpap*) were fixed factors, with REPETITION and SPEAKER (with by-SPEAKER intercept and slope for PHONEMIC QUANTITY and NUCLEUS TYPE) as random factors. P-values were obtained by a likelihood-ratio-test in which the final model was compared to a model with or without the effect in question. The results are also visualized in Figure 4.29.

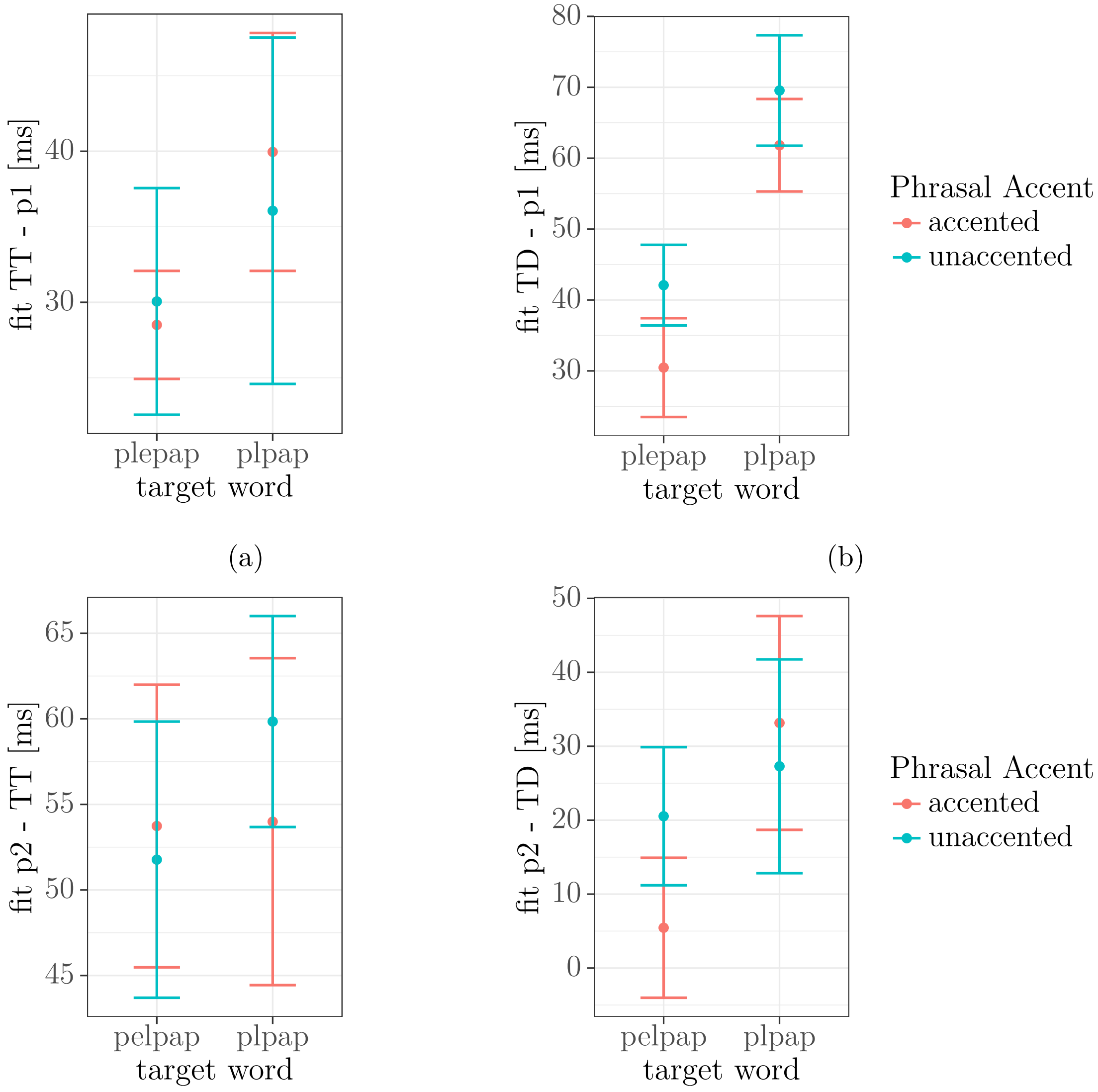
In the onset-nucleus comparison the temporal lag between the burst of the initial /p/ and the achievement of the apical constriction of /l/ was not significantly affected by PHRASAL ACCENT ($\chi^2[1] = 0.04, p > 0.1$) or SYLLABLE POSITION ($\chi^2[1] = 1.25, p > 0.1$). The mean lag was $32.8ms \pm 20.3$.

The temporal lag between the initial /p/ and achievement of the dorsal retraction extremum for the onset-nucleus comparison was significantly affected by SYLLABLE POSITION ($\chi^2[1] = 9.99, p < 0.01$) but not by PHRASAL ACCENT ($\chi^2[1] = 1.34, p > 0.1$). The interaction of PHRASAL ACCENT and SYLLABLE POSITION was also not significant ($\chi^2[1] = 0.29, p > 0.1$). The mean lag in *plepap* (onset condition) was $35.5ms \pm 15.7$ and in *plpap* (nucleus condition) was $64.6ms \pm 25.6$.

In the nucleus-coda comparison the temporal lag between the achievement of the apical constriction and the acoustically determined beginning of following /p/ was also not affected significantly by PHRASAL ACCENT ($\chi^2[1] = 0.09, p > 0.1$) or SYLLABLE POSITION ($\chi^2[1] = 0.31, p > 0.1$). The mean lag was $53.2ms \pm 22.0$.

When comparing the nucleus condition with the coda condition, PHRASAL ACCENT ($\chi^2[1] = 0.82, p > 0.1$) and SYLLABLE POSITION ($\chi^2[1] = 3.38, p > 0.05$) did not have an significant effect on the temporal lag between the achievement of the dorsal retraction extremum and beginning of the following /p/. However the interaction of PHRASAL ACCENT and SYLLABLE POSITION did have a significant effect ($\chi^2[1] = 5.01, p < 0.05$). Post-hoc Tukey tests revealed a significant difference for the nucleus and coda condition when the target word was accented ($z = 0.02, p < 0.05$). In accented condition the lag between the achievement of the dorsal constriction for /l/ and the closure achievement for the following /p/ was greater when /l/ was in nucleus condition than when it was in coda condition. In the accented condition the temporal lag between the dorsal retraction extremum and /p/ in *pelpap* was $5.4ms \pm 9.5$ and in *plpap* it was $33.2ms \pm 14.5$. In the unaccented condition the lag was $20.5ms \pm 9.3$ in *pelpap* and $27.3ms \pm 14.5$ in *plpap*.

In sum, the temporal coordination of the apical gesture with the preceding or following consonant did not differ depending on whether /l/ occupied the nucleus or not, so the results reported by Pouplier and Beňuš (2011) could not be replicated. They observed a greater lag between the apical gesture and the preceding or following consonant when /l/ occupied the syllable nucleus. But in Figure 4.29 (a) for the accented condition the lag between the initial /p/ and the apical constriction is greater when /l/ occupies the nucleus in *plpap* than the onset in *plepap*. The difference between the mean values is slightly above 10 ms, which is similar to the observations made by Pouplier and Beňuš (2011). However, probably due to a great variability, our results are not significant. The lag between the initial /p/ and the dorsal gesture of /l/ was significantly smaller for the onset condition compared to the nucleus condition. The target of the dorsal retraction gesture was reached significantly earlier in *plepap*. This is in accordance to the acoustic results for F2 in onset and coda condition. When accented, the lag between the achievement of the dorsal target and the lip closure of the following /p/ was greater in *plpap* (nucleus condition) than *pelpap* (coda condition).



(a)

(b)

(c)

(d)

Figure 4.29: Visualized effects of PHRASAL ACCENT and SYLLABLE POSITION of /l/ on the temporal lag in consonant sequences. The first row shows the lag between the acoustically determined end of initial /p/ (*p1* and apical constriction (*TT*) on the left and dorsal retraction achievement (*TD*) of /l/ in onset (*plepap*) and nucleus position (*plpap*) on the right. In the second row, the lag data for the coda (*pelpap*) and nucleus condition are visualized. The left plot shows the lag between the apical constriction and acoustic onset of the following consonant (*p2*) and the right plot shows the lag between the dorsal retraction achievement and *p2*. The plots show the mean and standard error for /l/ in onset and nucleus position.

4.4 Discussion

The aim of this chapter was to examine whether /l/ behaves differently depending on whether it occupies the nucleus, a position usually occupied by vowels, compared to when it occupies the onset or coda. Therefore we looked at articulatory and acoustic data of /l/ and how they are affected by phrasal accent depending on their syllable position.

Based mainly on findings for American English (Krakow, 1999; Sproat & Fujimura, 1993), we hypothesized that the movement magnitude and the temporal coordination of the apical and dorsal gesture of Slovak /l/ would differ depending on whether /l/ occupies the onset, nucleus, or coda. In general, regarding the magnitude of gestures, consonants in coda position have a weaker constriction than in onset position (Byrd, 1996). This is also true for the apical gesture of /l/ in American English. The 'vocalic' dorsal constriction of /l/, however, is usually stronger. A tendency for /l/ to have a lower F2 in coda, indicating a more retracted tongue, compared to onset has also been observed for many languages (Recasens, 2012). For /l/ occupying the nucleus, we hypothesized that it would be more similar to the coda condition, because the coda is often referred to as being more vowel-like and because Pouplier and Beňuš (2011) found in their study on Slovak that the apical gesture of /l/ in nucleus patterned kinematically more with the apical gesture in coda than onset. In our study we compared the onset and coda condition separately with the nucleus condition and they were not compared with each other directly. Our prediction for the effect of phrasal accent on consonants was less clearly defined. In general, consonants have been reported to be affected more by boundary than accent (Cho & Keating, 2009). In our target words, /l/ was never word initial or word final, thus never boundary adjacent. Since vowels are often said to carry the prosodic information, we initially hypothesized that /l/ in nucleus position, by taking over the vowel function, and particularly the 'vocalic' dorsal gesture would be affected most by phrasal accent. Yet, we already saw in the previous chapter, in which we compared the consonantal and vocalic nuclei that there was no significant effect of accent on the dorsal gesture of /l/, which left us with the question whether this was due to a ceiling effect.

Indeed, at the posterior angle the radius of /l/ in onset condition was smaller than in nucleus condition, indicating a weaker dorsal retraction in the onset condition. The magnitude of the retraction gesture for the coda condition was about the same as for the nucleus condition. However, in the accented condition a stronger coarticulatory effect of the preceding vowel /e/ on coda /l/ could be observed, so that during the acoustically determined vowel segment and beginning of the coda /l/ segment the tongue was less retracted. The retraction gesture for /l/ in onset or nucleus positions was not affected by accent. At the anterior angle a higher tongue position for the onset condition compared to the nucleus condition was observed, while in the coda condition the tongue was lower than in the nucleus condition. Phrasal accent had a weaker effect on the anterior region of the tongue for both onset and coda conditions compared to the nucleus condition. For both the onset-nucleus comparison and the nucleus-coda comparison, the interaction effect curve $f_3(t)$ went in the opposite direction to the covariate effect curve $f_1(t)$ for PHRASAL ACCENT, consequently levelling out the effect of accent more or less. For the coda condition the summed effect curves for the two accent conditions are both very flat lacking the more dynamic movement observed for the accented nucleus condition, which we considered to be the repositioning of the tongue followed by a rising movement of the anterior region.

In contrast to the coda condition, a greater difference for the two accent conditions could still be observed in the onset condition. Tongue rising occurred during the second half of the target sequence regardless of the accent condition. We attributed this to the /e/ following the onset /l/, as the /e/ requires a higher tongue position in the region observed. However, in the accented condition the anterior part of the tongue reaches a lower position at the beginning of the acoustically determined /l/-segment which cannot be attributed to a greater coarticulation effect and can thus be interpreted as stronger laterality. Although the tongue height measured at the anterior angle differed more between the onset and nucleus conditions than the nucleus and coda conditions, the effect of accent was more similar between the former two, in that the tongue reached a lower radius when accented and as a result greater overall movement was observed for the accented condition compared to the unaccented condition. This is surprising, because /l/ in onset condition was not boundary adjacent. As mentioned in Section 4.1.2, German nucleus adjacent consonants in onset clusters were found to not be affected by accent (Bombien, 2011), and for English it has been shown that consonants are affected more by boundary than phrasal accent (Cho & Keating, 2009). It is possible that some speakers inserted a boundary before the target word when accented, and the boundary affected not only the word-initial /p/ but also /l/. In some cases the initial bilabial stop was very long, which could be interpreted as a pause, however even by visual inspection of the video recordings of the lip movement it could not be established with certainty, since it was the lip closure that was prolonged.

The acoustic data were analysed for the /l/-segments only, but the subtle differences for syllable position observed in the articulatory data were lost. The formant data showed different dynamics for the different syllable positions caused by coarticulatory effects. Acoustically, we related F1 to tongue height in the anterior region and jaw height. At the end of the onset-nucleus comparison and at the beginning of the nucleus-coda comparison the differences were the greatest, with F1 being higher for the onset and coda conditions compared to the nucleus condition, which can be attributed to the lower jaw position for /e/ following or preceding /l/ in onset or coda position, respectively. Otherwise, the F1 data suggest a lower jaw and tongue height for the nucleus condition compared to the onset condition, and about the same height for the nucleus and coda conditions. In all positions, F1 was higher when accented compared to unaccented. The articulatory results measured at the anterior angle showed a greater accent effect for the nucleus compared to onset or coda. This was replicated in the acoustic results only partially for the coda condition where around time point 0.3 the interaction effect $f_3(t)$ was significant, resulting in a smaller difference between the two accent conditions for /l/ in coda compared to nucleus. For the onset-nucleus comparison, no significant interaction effect was observed but there was a trend level effect indicating the opposite to the articulatory data. The accent effect on F1 was at trend level greater for the onset condition than nucleus condition.

F2 has been associated with the retraction gesture. Again, at the end of the onset-nucleus comparison and at the beginning of the nucleus-coda comparison, a higher F2 for onset and coda compared to the nucleus was observed due to coarticulation with the vowel /e/. Other than that, we observed effects on F2 that could not be simply matched with the articulatory data measured at the posterior angle. The very low F2 at the beginning of the accented onset condition compared to the unaccented onset condition, and for the unaccented coda condition compared to the unaccented nucleus condition, must be a consequence of other articulatory differences. While in general, a lower F2 indicates a darker /l/, which usually correlates with a more retracted tongue dorsum, it has also been associated with less dorsopalatal contact and the length of the cavity behind the primary constriction (Recasens, 2004). When accented, F2 for coda /l/ was higher than when unaccented, which goes hand in hand with the less retracted tongue measured at the posterior angle. A detailed token by token analysis was not carried out, so that it is not possible to bring together the acoustic and articulatory results with certainty. To visualise some possible realizations of /l/ in the three different syllable contexts and to demonstrate the shortcomings of the measures taken, the tongue contours during the acoustically determined /l/ segments of one repetition of the onset, nucleus and coda condition, each by the same speaker, are shown in Figure 4.30. The tongue contours captured during the nucleus condition are shown in red, for the onset in green and for the coda condition in blue. For each condition contours for each recorded frame during the acoustically determined /l/ are shown and not just at one time point. All three target words were accented. At the posterior angle the red contours corresponding to the nucleus condition have the highest radius, although the differences between the three conditions are minimal. This was also shown in the statistical results. At the anterior angle, it is the nucleus condition for which the lowest radius is observed, while at the measured angle the radius for the onset and coda condition are overlapping. While at this angle the tongue position behind the raised tongue tip was captured for /l/ in nucleus (red) and onset (green) condition, for the coda condition an angle slightly further back would have been more appropriate. Obviously, reducing the articulatory data to only two points limits what is captured in the analysis.

Articulatory phonology predicts different coordination of gestures depending on their position within the syllable. Temporal coordination between consonant gestures and of consonant gestures in relation to the vocalic nucleus have been shown to vary systematically for onset and coda, however it is not clear how two gestures in the nucleus, as is the case for syllabic /l/, would behave, and the framework of articulatory phonology does not provide any general prediction. However, Pouplier and Beňuš (2011) reported no effect of syllable position on the temporal coordination of the two gestures of /l/ for Slovak, based on preliminary measurements. In our data, the temporal lag between the two gestures differed significantly depending on whether /l/ occupied the onset or the nucleus. The two gestures were timed closer together in the onset condition compared to the coda condition, which corresponds to what is known for American English (Krakow, 1999). Contrary to American English, in

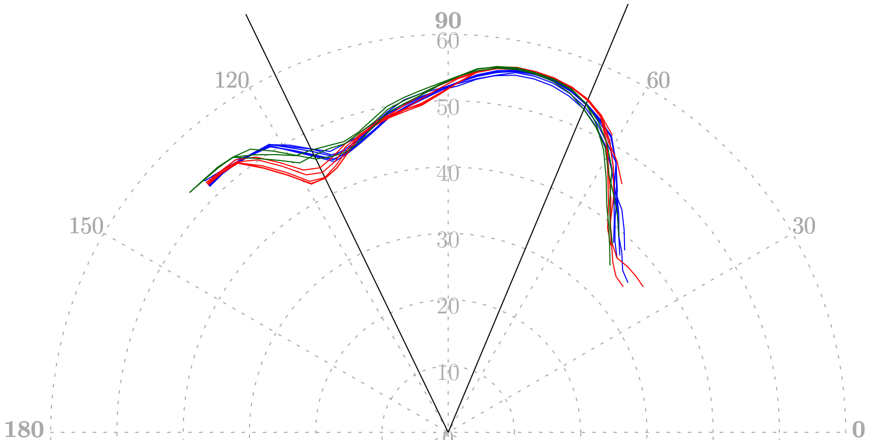


Figure 4.30: Example of tongue contours captured during one instance of each of the acoustically segmented /l/ in nucleus condition (red), onset condition (green) and coda condition (blue) by the same speaker. Tongue tip is to the left. The target words (*plpap, plepap, pelpap*) were in accented condition. The contours of the same colour were extracted from the same token but at different time points within the acoustically determined /l/.

which the dorsal gesture precedes the apical gesture, the apical gesture in our data seems to precede the dorsal gesture. However, the dorsal gesture approaches the target more slowly than the apical gesture and, based on visual inspection, the dorsal gesture seems to be generally initiated earlier than the apical gesture. With our method of determining the target achievement the time point of achievement also seemed to be in the middle or at the end of the target plateau. Such a case is visualized in Figure 4.31. The tongue movement captured at the posterior movement slows down at around 2.55s, and what follows looks like a plateau. However, some higher radius is captured later in the segment, after 2.75s. The target was defined to be achieved when the radius was within 0.45mm within the maximal displacement, comparable to the measure used by Gick et al. (2006) and Lin (2011). This is a very small value, also considering the variation when tracking the tongue contour. The nucleus condition patterned between the onset and coda condition. In the accented condition the nucleus was more similar to the onset condition while in the unaccented condition it was more similar to the coda condition.

The results for the temporal coordination of the two gestures of /l/ with either preceding or following /p/ revealed that if there was a significant effect of SYLLABLE POSITION, the lag between /p/ and the gesture in question was greater when /l/ occupied the nucleus than when it occupied the onset or coda. This is in accordance with the effect observed by Pouplier and Beňuš (2011). However, in our case the significant differences were observed for the dorsal gesture, while Pouplier and Beňuš (2011) measured the temporal coordination for the apical gesture only. Our results showed only a tendency in the same direction for the apical gesture. The results for the temporal coordination of the dorsal gesture with the release of the initial /p/ for the onset and nucleus conditions showed a significantly smaller lag for the onset condition. Thus, in the onset condition the target of the dorsal constriction is achieved earlier than in the nucleus condition. This is also consistent with the F2 data, where in the onset condition the lowest F2 values were observed at the beginning of the /l/ segment, while in the nucleus condition it was at the end. A possible explanation for the small lag in the onset condition is that the dorsal gesture is shorter, because the following vowel /e/ requires a more fronted tongue position. When looking at the summed effect curves for the dorsal gesture in Figure 4.10, it seems that for the accented onset condition it is the plateau that is shorter. The acoustically determined duration of the initial /p/ (or *p1* in Figure 4.10) in *plpap* (nucleus condition) and *plepap* (onset condition) did not differ for the two target words, but the target sequence *ple* was shorter than *plp* due to durational differences of the remaining segments. Yet, it does not seem that the dorsal gesture for /l/, which begins during the *p1*-segment, is faster or begins earlier. The difference in temporal lag between /p/ and /l/ for the onset and nucleus condition do not seem to be caused by a greater overlap of the two, but rather the difference following segments, namely /p/ without specified tongue configuration and /e/ with conflicting tongue configuration in the dorsal region.

The lag between the achievement of the dorsal constriction target and the closure of the following consonant /p/ for the nucleus and coda condition was significantly different only when accented. In that case, the lag was greater for the nucleus condition. The release of the /p/ closure for the onset condition or the beginning of the /p/ closure for the coda condition and the achievement of the dorsal gesture a closer together when accented. It is not clear whether the dorsal gesture plays a crucial role in the temporal organization of gestures into syllables, or whether the differences emerge in order to still provide for the retraction despite the opposing tongue configurations in the context of a front vowel, analogously to stress shift in stress clash situations (e.g. Beckman, Swora, Rauschenberg, and Jong (1990), Prieto (2005)), a scenario not foreseen in common coarticulation theories.

The interpretation of the results in this chapter proved to be difficult, because /l/ in onset and coda was always followed or preceded by a vowel, while /l/ in nucleus condition was never next to a vowel. A slightly better comparability between the onset or coda condition with the nucleus condition could have been obtained by choosing a different vowel for the onset and coda condition. By changing the target word for the onset condition to *plapap* and for the coda condition to *pilpap*, for example, the vocalic context of /l/ skipping the preceding or following /p/ would be the same. Regarding the timing data, using the beginning of the dorsal gesture or the peak velocity might have been more reliable. Nevertheless, the apical constriction remains difficult to capture using ultrasound.

Regardless of the above mentioned difficulties, we can still discuss the results concerning our initial questions. We initially expected that the dorsal gesture functions as a vowel by signalling prosodic emphasis and enables a temporal coordination with consonants in onset and coda similar to a vocalic nucleus. We already saw in the previous chapter that this was

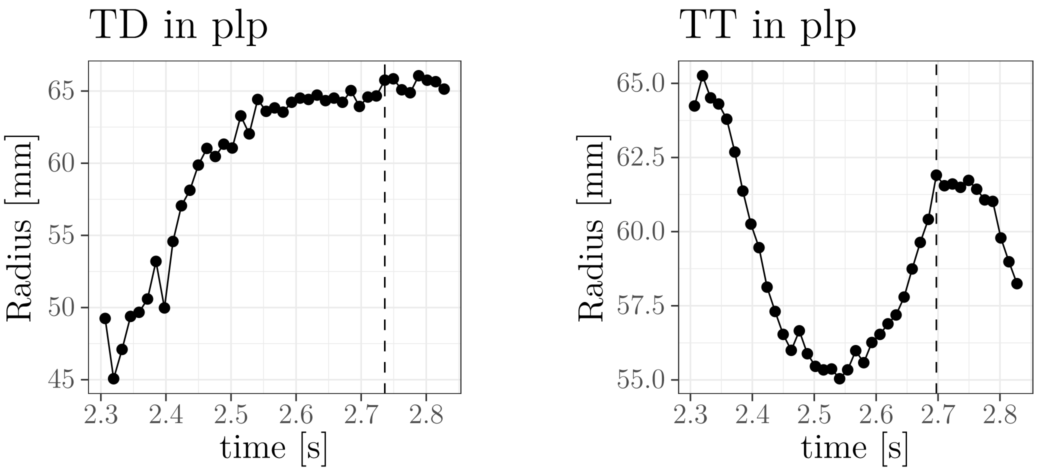


Figure 4.31: Example showing tongue movement contour captured at the posterior and apical angles.

not the case. It was the anterior part of the tongue that mainly changed under prosodic emphasis, functioning like a vowel in that respect. The dorsal gesture in American English /l/ differs depending on its position within a syllable, and consequently allophonic variation can be observed. For Southern British English, where similar to American English /l/ in coda is much darker than in onset, Scobbie and Pouplier (2010) observed that in ambisyllabic position the dorsal gesture was stable in that it remained retracted, thus coda-like, even when the apical gesture was resyllabified and became onset-like. In these cases the dorsal gesture differs systematically depending on its syllable affiliation. However in Slovak, the differences depending on the prosodic context, both syllable affiliation and phrasal accent, on the dorsal gesture were rather subtle and its main function appears to be providing the dark quality of /l/. The main differences in terms of magnitude of the gesture depending on syllable position and phrasal accent were observed at the anterior angle.

In addition to the difference in magnitude or constriction degree between syllable positions, articulatory phonology also predicts systematic differences in temporal coordination. The relation of /l/ to the preceding or following consonant for the onset and coda condition could not be compared directly. For the onset-nucleus comparison we measured the temporal lag between the release of the preceding stop and the achievement of the gestural targets of

/l/, while for the nucleus-coda comparison the lag between the achievement of /l/ gestures and the beginning of the following stop closure was measured. The temporal lag between the dorsal gesture and the preceding or following consonant differed for the onset-nucleus comparison, as well as the nucleus-coda comparison. For the tongue tip gesture no differences were observed. However, as already mentioned, the differences did not seem to be caused due to more overlap with preceding or following /p/ when /l/ occupied the onset or coda, but rather due to the presence of a vowel with conflicting tongue configuration in the dorsal region.

Smith and Lammert (2013) analysed vocalized and unvocalised /l/ production in American English and observed “two models of production of the consonantal element of /l/. One is the prototypical tongue tip closure at the alveolar ridge. The other is tongue curling through blade lowering, which is present whether or not tongue tip closure takes place” (p.3233). Consequently, they question whether the actual goal of the consonantal gesture for /l/ is “a non-constriction-based consonantal task” (p.3233). As also noted by Recasens (2016) regarding the tongue body, “the specific tongue body configuration for alveolar laterals (whether clearer or darker) is to a large extent conditioned by laterality requirements” (p.33). He particularly criticizes the assignment of a tongue body gesture to clear /l/ with the argument that it shows less coarticulation than other consonants. We can adopt these ideas and suggest that, for Slovak, /l/ consists of one lateral gesture which requires a retracted tongue dorsum, but that the dorsal retraction is not separated from the apical gesture in that the strengthening of the dorsal gesture does not correlate with the weakening of the apical gesture, nor does the dorsal gesture function as the main carrier of prosodic information. Of course, this proposal has to be further tested for example for boundary effects and rime duration as done by Sproat and Fujimura (1993) and more carefully regarding the temporal coordination in different syllable contexts to examine whether the magnitude of the dorsal retraction and its temporal coordination with the anterior gesture or neighbouring consonant gestures changes or not.

To summarize, /l/ was largely unchanged in the different syllable positions, except for the unavoidable differences due to coarticulatory effects. Although it was not our primary focus, the main difference as an effect of syllable position seems to be the acoustic duration. /l/ in nucleus position was more than 20ms longer than /l/ in onset condition. Compared to /l/ in coda position, /l/ in nucleus was only 8ms longer, but the difference still proved to be significant. Upon visual inspection of Figures 4.10 and 4.18, the durational differences do not, however, seem to affect the articulatory velocity or magnitude of the gesture but rather the plateau duration, particularly when comparing the tongue movement captured at the posterior angle in onset and nucleus conditions. Accent had an effect on F1, indicating a lower jaw position. Contrary to our initial hypothesis that the dorsal gesture enables /l/ to function like a vowel, we saw that the dorsal gesture does not act separately from the anterior gesture and does not function as the carrier of accentual information. Instead, a difference in tongue movement captured at the anterior angle as an effect of accent was observed, which was strongest for the nucleus position, somewhat weaker for the onset condition and non-existent for the coda condition.

Chapter 5

Summary and conclusion

The main issue we tried to address is the difference between vowels and consonants and why, although not cross-linguistically rare, syllabic consonants are more restricted than vocalic nuclei. As outlined in the introduction, the difference between vowels and consonants cannot be explained based solely on their acoustic and articulatory properties, and is related to their functional differences. Based on the common claim that vowels carry the prosodic information and that syllabic consonants are cross-linguistically less common in stressed and accented syllables, we examined how phrasal accent is realized on syllabic consonants in Slovak, namely /l/ (and in Chapter 2 also /r/). Dark /l/ as it occurs in Slovak is described as consisting of two gestures, a consonantal apical gesture and a vocalic dorsal gesture. As such, it can be viewed as a segment which can vary easily between a more consonant-like and a more vowel-like form, as has been reported for American English (Sproat & Fujimura, 1993). Slovak /l/ can occupy the nucleus of a stressed syllable, and hence provides an ideal testing ground for the interaction of form (consonantal vs. vocalic) and function (nucleus and carrier of prosodic information vs. onset or coda). Pouplier and Beňuš (2011) found that the apical gesture of syllabic /l/ (and /r/) does not become more vowel-like kinematically. However, they found that the timing of syllabic consonants with adjacent consonants showed consistently greater lags than when both consonants were part of the onset or the coda cluster of a syllable with a vocalic nucleus. They suggested that the open transitions, resulting from the consonant gestures which are timed farther apart, provide the necessary modulation of the speech signal in vowelless sequences. The greater lag provides a sonorant portion in the acoustic signal, similar to a schwa, which is related to the retraction gesture. They further speculated that while both the apical and dorsal gesture form the nucleus, the 'vocalic' dorsal gesture makes liquids cross-linguistically more likely to occupy the nucleus than obstruents. Consequently, we hypothesized that the dorsal gesture, which is kinematically more similar to vowels, would also function like a vowel to signal prosodic emphasis.

This study can be divided in two parts: on the one hand, we compared the vowel /e/ and the consonant /l/ (and in Chapter 2 also /r/) in the nucleus position, where we examined whether and how they function as carriers of prosodic information (Chapters 2 and 3). On the other hand, we examined if and how /l/ differs depending on whether it occupies the syllable nucleus or the onset and coda positions (Chapter 4). In this chapter we will summarize the observations from previous chapters once again in order to address the overall question of whether the main difference between vowels and consonants lies in their ability to satisfy the function required by prosodic emphasis.

5.1 Vowels or consonants as syllable nuclei

First, in Chapter 2, we investigated the realization of F0, the main indicator for phrasal accent. The target nucleus was phonemically short or long and either vocalic (/e/), or consonantal (/l/ or /r/). While it is known that intonation contours can extend from vocalic nuclei to neighboring sonorant consonants (Fletcher et al., 2015; Jilka & Möbius, 2006), data from Tashlhiyt Berber suggested that consonantal nuclei are less preferred as carriers of phrasal accent (Grice et al., 2015). In Slovak, accented words are realized with an F0 peak on the first syllable of the word (Král', 2005), but to our knowledge this was not specifically tested for words with consonantal nuclei. Our data confirmed that the F0 peak was on the nucleus for both vocalic and consonantal nuclei and regardless of whether they were phonemically short or long. In the accented condition, the phonemic length contrast of the nucleus was also enhanced by lengthening the long nuclei (Chapter 2), a phenomenon comparable to what was observed in earlier studies as an effect of word stress (Beňuš & Pouplier, 2011; Beňuš & Mády, 2010). The findings in Chapter 2, which showed that consonantal nuclei pattern with vocalic nuclei regarding F0 and acoustic duration, served as a starting point to explore the requirements of prosodic emphasis for segments occupying the syllable nucleus.

In accented syllables, vowels occupying the syllable nucleus usually exhibit qualitative differences compared to in unaccented syllables. The explanation for the differences are provided by two main theories: the Sonority Enhancement Hypothesis (Beckman et al., 1992) and the Hyperarticulation Hypothesis (De Jong, 1995). Taking these as starting points, in Chapter 3 we examined the tongue movement at the anterior and posterior part of the oral cavity for /l/ and /e/. Both hypotheses predict a more retracted tongue back in accented conditions but make contradicting predictions for the apical gesture. The Hyperarticulation Hypothesis predicts a stronger apical constriction and thus a higher tongue position, while the Sonority Enhancement Hypothesis predicts a lower tongue position to allow for a wider oral cavity. We hypothesized that the contradicting predictions for the anterior part of the oral cavity might limit the effects of prosodic emphasis to the posterior part. Similarly to what has been proposed for high front vowels, this would explain why segments with a more narrow specification in the anterior part of the oral cavity are phonologically less preferred in stressed and accented syllables.

Contrary to expectation, the dorsal gesture of /l/ was affected by phonemic quantity but not by phrasal accent. However, an effect of accent was observed for the anterior region. The anterior part of the tongue was lowered more during the preceding consonant in the accented condition, but a stronger rising movement was observed during the acoustically determined /l/. We proposed that the more dynamic movement in accented conditions is needed for the reorganization of the tongue shape and allows a more curled tongue shape with a greater lateral opening. The following upwards movement suggests a stronger apical constriction. For the vocalic nucleus a more dynamic tongue movement in the form of the trough effect was observed. Acoustically, a higher F1 in the accented condition was observed for both /e/ and /l/ which we assigned to a lower jaw. For both the vocalic and consonantal nucleus, the sonority was enhanced in the accented condition as reflected in the acoustic data. The articulatory data revealed a stronger movement towards the target in accented condition, which can be interpreted as hyperarticulation. However, while accent affected the entire tongue movement of the vowel, it affected only the movement of the anterior part of the tongue for /l/. The dorsal gesture of /l/, which we predicted to function like a vowel, was not affected by phrasal accent. Possibly, the dorsal gesture was not further retracted due to a ceiling effect. Yet, altogether, syllabic /l/ functioned like a vowel in that it signalled prosodic emphasis.

5.2 /l/ as “vowel” or “consonant”

In Chapter 4, we compared /l/ in nucleus with /l/ in onset and coda to investigate whether differences can be observed when /l/ occupies different positions within a syllable. One of the questions we pursued was whether the dorsal gesture plays an important role in the different spacial and temporal coordination of gestures for different syllable positions, as in American English or Southern British English (Krakow, 1999; Scobbie & Pouplier, 2010). In general, dark /l/ in the onset position is associated with a stronger consonantal apical gesture and a weaker vocalic dorsal gesture, and the opposite in the coda condition. The temporal coordination of the two gestures in different varieties of English also differs as an effect of syllable position, corresponding to the general gestural coordination patterns (Nam et al., 2009). In onset position, the two gestures are coordinated in-phase, while in coda position the two gestures are coordinated sequentially, so that the dorsal gesture precedes the apical gesture. Furthermore, we investigated whether the gestures for /l/ were affected differently by phrasal accent depending on their syllable affiliation. Phrasal accent is expected to affect the vocalic nucleus, while the effect has been reported to not spread to nucleus adjacent consonants (Bombien, 2011; Cho & Keating, 2009). Consonant gestures occupying the onset position have been found to overlap less with each other when accented, but more with the vocalic nucleus, while for coda consonant gestures it has been found that they show less overlap with the vocalic nucleus (Bombien, 2011; Peters, 2015).

Indeed, the gestures of /l/ differed for the different syllable affiliations, and were affected to differing degrees by phrasal accent. The magnitude of the tongue movement in the anterior and posterior region differed as an effect of syllable position, and phrasal accent had a different effect on the gestures depending on their syllable affiliation. However, the main differences in syllable affiliation could be attributed to coarticulatory effects. /l/ in different syllable positions differed also in acoustic duration, which was much shorter in onset position and significantly but not as much shorter in coda position compared to the nucleus condition. The dorsal gesture was prominently visible in all three syllable positions, although its magnitude was smaller in onset and coda conditions. In onset and nucleus positions, phrasal accent did not affect the magnitude of the dorsal gesture, while in the coda position the tongue was less retracted when accented, most likely due to the stronger coarticulatory effect of the vowel /e/ occupying the nucleus, which requires a more fronted tongue. In the onset position, instead of the magnitude of the dorsal gesture being affected, it seems that its plateau duration was affected. At the anterior region the tongue was higher in the onset position and lower in the coda position compared to the nucleus position. Phrasal accent affected the movement of the anterior region of the tongue only for /l/ in onset and nucleus condition, in that the tongue reached a lower position, but after that also a greater upwards movement was observed. For the two syllable positions, a more dynamic tongue movement was observed in the anterior region when accented. The effect was greater for the nucleus condition. In coda position, the tongue movement of the anterior region was not affected by phrasal accent. Acoustically, the effect of phrasal accent was reflected in a higher F1, regardless of syllable position, which we attributed to a lower jaw. The temporal coordination of the two gestures of /l/ differed as an effect of syllable position, in contrast to preliminary results reported by Pouplier and Beňuš (2011) who did not find differences in the temporal coordination of the two gestures. In the onset position the two gestures were closer together than in the coda condition, while in the nucleus they patterned in between. However, in all cases the anterior gesture preceded the posterior gesture, contrary to what has been observed for different varieties of English. In part, this might have been caused by our method of determining the lag. As laid out in the discussion of Chapter 4 and visualized in Figure 4.31, the time point determined as the achievement of the target of the dorsal retraction gesture was towards the end of what looks like a plateau upon visual inspection, while for the apical gesture it was the beginning of the plateau. Consequently, durational differences of the plateau of the dorsal gestures might as well have played a role in the different coordination patterns obtained. The temporal coordination of the two gestures of /l/ with preceding or following consonant gestures confirmed the results provided for the apical gesture by Pouplier and Beňuš (2011): The lag was greater when /l/ occupied the nucleus. However, the difference was significant only for the dorsal gesture. Consequently, there is no support for the proposal that the greater lag between the apical gesture and the adjacent consonant gesture for /l/ occupying the nucleus provides a longer sonorous portion during which the dorsal gesture might provide for the vowel-like modulation of the speech signal. The proposal can further be refuted because the effect of phrasal accent was not reflected in the magnitude of the dorsal gesture.

Taking together all results, the two gestures of /l/ in Slovak do not seem to act autonomously, and the 'vocalic' dorsal gesture does not seem to function as the sole carrier of prosodic information. From what we could see in our data, the dorsal gesture is there, because /l/ in Slovak happens to be dark. While phrasal accent was signalled by the nucleus, it was coded in the jaw position as reflected acoustically in F1, and for syllabic /l/ in the 'consonantal' anterior part of the tongue.

5.3 Final remarks

The interpretation of the results did not provide clarity in establishing a better categorization of vowels and consonants. Contrary to expectations, /l/ did not weaken its apical gesture to become more vowel-like in the nucleus position, yet it functioned as a carrier of prosodic information. This could partly be attributed to the fact that most of our initial hypotheses were based mainly on observations made for American English. The allophonic variation of /l/ as observed for English, where differences between different syllable positions are great, is not universal. As already noted by (Sproat & Fujimura, 1993, p. 306), "It is generally assumed that the relative timing of multiple gestures (or targets) is specified by (possibly language specific) phonetic implementation rules, and is not specified at any level of phonological representation". As laid out in their paper, gestural coordination patterns for /l/ go hand in hand with general coordination patterns observed for American English. The comparisons of /l/s in different languages by Recasens (2012) showed that in some languages /l/ varies more depending on syllable affiliation, while in others it is affected more by vowel context, showing greater coarticulatory variation. Taking together the statement by Sproat and Fujimura and the observation by Recasens, the realization of /l/ is highly dependent on language specific spacial and temporal coordination of gestures and consequently differs also in the way it is affected by coarticulatory mechanisms. Pouplier and Beňuš (2011) found that Slovak syllables show coordination patterns differing from those according to the model by Nam et al. (2009) and observed among others for English or German. While in English adding consonants to the onset results in greater overlap of the preceding consonants and the vowel due to the in-phase coordination of the onset consonants with the nucleus, results by Pouplier and Beňuš (2011) suggest that this is not the case for Slovak. The different timing patterns between Slovak and English might be the reason why Slovak allows syllabic consonants like /l/ with both the 'consonantal' and 'vocalic' gestures, while in English syllabic /l/ is much more restricted and often vocalized. Our findings support the concluding remark by Pouplier and Beňuš (2011) that there is a relationship between phonotactics and consonant timing. It is possible that the important difference between vowels and consonants lies in their different possibilities of temporal coordination while remaining fully perceptible and at the same time also carrying information about the adjacent sounds and the prosodic context.

At this point we could just suggest that vowels are cross-linguistically most suited as syllable nuclei. However, languages seem to differ in their requirements for the nucleus. Not only might they differ in temporal coordination of gestures, but also in how stress or accent is coded. It might be interesting to carry out experiments mentioned in the introduction which provide evidence for functional differences between vowels and consonants (e.g. Bonatti et al. (2005); Bouchard et al. (2013); Carreiras and Price (2008); Nazzi (2005); Nespor et al. (2003)) using languages like Slovak or Tashlhiyt Berber. Future research on correlation between phonotactics and detailed spacio-temporal coordination of gestures might further allow us to understand the universal preference to treat vowels and consonants as categorically distinct in contrast to language specific exceptions.

References

Articulate Instruments Ltd. (2012a). Articulate assistant advanced user guide (2.14 ed.) [Computer software manual]. Edinburgh, UK.

Articulate Instruments Ltd. (2012b). Ultrasound stabilisation headset users manual: Revision 1.4 [Computer software manual]. Edinburgh, UK.

Barnes, J., Brugos, A., Veilleux, N., & Shattuck-Hufnagel, S. (2011). Voiceless intervals and perceptual completion in F0 contours: Evidence from scaling perception in American English. In *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 108–111).

Barnes, J., Brugos, A., Veilleux, N., & Shattuck-Hufnagel, S. (2014). Segmental influences on the perception of pitch accent scaling in English. In *Proceedings of 7th Speech Prosody Conference* (pp. 1125–1129).

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4 [Computer software manual]. Retrieved from `http://CRAN.R-project.org/package=lme4` (R package version 1.1-7)

Beckman, M. E., Edwards, J., & Fletcher, J. (1992). Prosodic structure and tempo in a sonority model of articulatory dynamics. *Papers in Laboratory phonology II: Segment, gesture, prosody*, 68–86.

Beckman, M. E., Swora, M. G., Rauschenberg, J., & Jong, K. d. (1990). Stress shift, stress clash, and polysyllabic shortening in a prosodically annotated discourse. In *First International Conference on Spoken Language Processing.*

Bell, A. (1978). Syllabic consonants. In J. H. Greenberg (Ed.), *Universals of human language* (Vol. 2: Phonology, pp. 153–201). Stanford University Press.

Beňuš, Š. (2014). Phonological structure and articulatory phonetic realization of syllabic liquids. In J. Emonds & M. Janebová (Eds.), *Language use and linguistic structure. Proceedings of the Olomouc linguistics colloquium 2013* (pp. 281–291).

Beňuš, Š., & Pouplier, M. (2011). Jaw movement in vowels and liquids forming the syllable nucleus. In *12th Annual Conference of the International Speech Communication Association.*

Beňuš, Š., & Mády, K. (2010). Effects of lexical stress and speech rate on the quantity and quality of Slovak vowels. In *Proceedings of the 5th International Conference on Speech Prosody.*

Beňuš, Š. (2011). Control of phonemic length contrast and speech rate in vocalic and consonantal syllable nuclei. *The Journal of the Acoustical Society of America*, *130*(4), 2116–2127.

Beňuš, Š., Reichel, U. D., & Mády, K. (2014). Modelling accentual phrase intonation in Slovak and Hungarian. In L. Veselovská & M. Janebová (Eds.), *Complex visibles out there. Olomouc modern language series, 4* (pp. 677–689). Palacký University, Olomouc, Czech Republic.

Boersma, P., & Weenink, D. (2014). *Praat: doing phonetics by computer.* Retrieved from http://www.praat.org/

Bombien, L. (2011). *Segmental and prosodic aspects in the production of consonant clusters* (Doctoral dissertation). Ludwig-Maximilians-Universität, Munich.

Bombien, L., Mooshammer, C., Hoole, P., Rathcke, T., & Kühnert, B. (2007). Articulatory strengthening in initial German /kl/ clusters under prosodic variation. In *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 457–460).

Bombien, L., Winkelmann, R., & Scheffers, M. (2016). wrassp: an R wrapper to the ASSP library [Computer software manual]. (R package version 0.1.4)

Bonatti, L. L., Peña, M., Nespor, M., & Mehler, J. (2005). Linguistic constraints on statistical computations the role of consonants and vowels in continuous speech processing. *Psychological Science*, *16*(6), 451–459.

Bouchard, K. E., Mesgarani, N., Johnson, K., & Chang, E. F. (2013). Functional organization of human sensorimotor cortex for speech articulation. *Nature*, *495*(7441), 327–332.

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

Browman, C. P., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology*, *6*(2), 201–251.

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

Browman, C. P., & Goldstein, L. (1995). Gestural syllable position effects in American English. *Producing speech: Contemporary issues*, 19–33.

Buschmeier, H., & Wlodarczak, M. (2013). TextGridTools: A TextGrid processing and analysis toolkit for Python. In *Tagungsband der 24. Konferenz zur elektronischen Sprachsignalverarbeitung* (pp. 152–157). Bielefeld, Germany.

Butcher, A., & Harrington, J. (2003). An instrumental analysis of focus and juncture in Warlpiri. In *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 321–324).

Butskhrikidze, M. (2002). *The consonant phonotactics of Georgian* (Doctoral dissertation).

Bučar Shigemori, L. S., Pouplier, M., & Beňuš, Š. (2015). Phonemic length and phrasal accent in Slovak consonantal and vocalic nuclei. In *Proceedings of the 18th International Congress of Phonetic Sciences.*

Byrd, D. (1995). C-centers revisited. *Phonetica*, *52*(4), 285–306.

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

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

Carreiras, M., & Price, C. J. (2008). Brain activation for consonants and vowels. *Cerebral Cortex*, *18*(7), 1727–1735.

Cederbaum, J. (2017). sparseFLMM: Functional linear mixed models for irregularly or sparsely sampled data [Computer software manual]. Retrieved from `https://CRAN.R-project.org/package=sparseFLMM` (R package version 0.1.0)

Cederbaum, J., Pouplier, M., Hoole, P., & Greven, S. (2016). Functional linear mixed models for irregularly or sparsely sampled data. *Statistical Modelling*, *16*(1), 67–88.

Chitoran, I. (1998). Georgian harmonic clusters: phonetic cues to phonological representation. *Phonology*, *15*(2), 121–141.

Cho, T. (2005). Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of/a, i/in English. *The Journal of the Acoustical Society of America*, *117*(6), 3867–3878.

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

Clements, G. N. (1990). The role of the sonority cycle in core syllabification. *Papers in Laboratory Phonology*, *1*, 283–333.

Coleman, J. (2001, 5). The phonetics and phonology of Tashlhiyt Berber syllabic consonants. *Transactions of the Philological Society*, *99*(1), 29–64. doi: 10.1111/1467-968X.00073

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. C., Liberman, A. M., & Cooper, F. S. (1955). Acoustic loci and transitional cues for consonants. *The Journal of the Acoustical Society of America*, *27*(4), 769–773.

Edwards, J., Beckman, M. E., & Fletcher, J. (1991). The articulatory kinematics of final lengthening. *The Journal of the Acoustical Society of America*, *89*(1), 369-382.

Erickson, D. (2002). Articulation of extreme formant patterns for emphasized vowels. *Phonetica*, *59*(2-3), 134–149.

Fletcher, J. (2010). The prosody of speech: Timing and rhythm. *The Handbook of Phonetic Sciences, Second Edition*, 521–602.

Fletcher, J., Stoakes, H., Loakes, D., & Singer, R. (2015). Accentual prominence and consonant lengthening and strengthening in Mawng. In *Proceedings of the 18th International Congress of Phonetic Sciences.*

Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *The Journal of the Acoustical Society of America*, *101*(6), 3728–3740.

Fougeron, C., & Ridouane, R. (2008). On the phonetic implementation of syllabic consonants and vowel-less syllables in Tashlhiyt. *Estudios de fonética experimental*, *17*, 140–175.

Fowler, C. A. (1995). Acoustic and kinematic correlates of contrastive stress accent in spoken English. *Producing speech: contemporary issues for Katherine Saffort Harris. American Institute of Physics, Woodbury/New York*, 355–373.

Fowler, C. A., & Brown, J. M. (1997). Intrinsic f0 differences in spoken and sung vowels and their perception by listeners. *Perception & Psychophysics*, *59*(5), 729–738.

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

Fox, J. (2003). Effect displays in R for generalised linear models. *Journal of Statistical Software*, *8*(15), 1–27.

Fuchs, S., Hoole, P., Brunner, J., & Inoue, M. (2004). The trough effect – an aerodynamic phenomenon. In *MIT-meeting From Sound to Sense.*

Gick, B., Campbell, F., Oh, S., & Tamburri-Watt, L. (2006). Toward universals in the gestural organization of syllables: A cross-linguistic study of liquids. *Journal of Phonetics*, *34*(1), 49–72.

Giles, S. B., & Moll, K. L. (1975). Cinefluorographic study of selected allophones of English /i/. *Phonetica*, *31*(3-4), 206–227.

Goldstein, L., Byrd, D., & Saltzman, E. (2006). The role of vocal tract gestural action units in understanding the evolution of phonology. In M. A. Arbib (Ed.), *Action to language via the mirror neuron system* (p. 215—249). Cambridge University Press.

Goldstein, L., Chitoran, I., & Selkirk, E. (2007). Syllable structure as coupled oscillator modes: evidence from Georgian vs. Tashlhiyt Berber. In *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 241–244).

Gordon, M., & Nafi, L. (2012). Acoustic correlates of stress and pitch accent in Tashlhiyt Berber. *Journal of Phonetics*, *40*(5), 706–724.

Gordon, M. K. (2016). *Phonological typology* (B. Vaux, Ed.). Oxford: Oxford University Press.

Greenberg, S., Carvey, H., Hitchcock, L., & Chang, S. (2003). Temporal properties of spontaneous speech – a syllable-centric perspective. *Journal of Phonetics*, *31*(3), 465–485. (Temporal Integration in the Perception of Speech)

Grice, M., Ridouane, R., & Roettger, T. B. (2015). Tonal association in Tashlhiyt Berber: Evidence from polar questions and contrastive statements. *Phonology*, *32*(02), 241–266.

Hanulíková, A., & Hamann, S. (2010). Slovak. *Journal of the International Phonetic Association*, *40*(3), 373–378.

Harrington, J., Fletcher, J., & Beckman, M. (2000). Manner and place conflicts in the articulation of accent in Australian English. *Papers in laboratory phonology*, *5*, 40–51.

Harrington, J., Fletcher, J., & Roberts, C. (1995). Coarticulation and the accented/unaccented distinction: Evidence from jaw movement data. *Journal of Phonetics*, *23*(3), 305–322.

Heldner, M., & Strangert, E. (2001). Temporal effects of focus in Swedish. *Journal of Phonetics*, *29*(3), 329–361.

Henke, E., Kaisse, E. M., & Wright, R. (2012). Is the sonority sequencing principle an epiphenomenon. In P. S. (Ed.), *The sonority controversy* (pp. 65–100). DeGruyter Mouton.

Hermes, A., Mücke, D., & Grice, M. (2013). Gestural coordination of Italian word-initial clusters: the case of 'impure s'. *Phonology*, *30*(1), 1–25.

Honda, K. (1983). Relationship between pitch control and vowel articulation. *Haskins Laboratories Status Report on Speech Research, SR*, *73*, 269–282.

Honorof, D. N., & Browman, C. P. (1995). The center or edge: How are consonant clusters organized with respect to the vowel. In *Proceedings of the 13th International Congress of Phonetic Sciences* (pp. 552–555).

Hoole, P., & Honda, K. (2011). Automaticity vs. feature-enhancement in the control of segmental f0. In G. N. Clements & R. Ridouane (Eds.), *Where do phonological features come from?: Cognitive, physical and developmental bases of distinctive speech categories* (Vol. 6, p. 131). John Benjamins Publishing.

Hyman, L. M. (2008). Universals in phonology. *The linguistic review*, *25*(1-2), 83–137.

Jilka, M., & Möbius, B. (2006). Towards a comprehensive investigation of factors relevant to peak alignment using a unit selection corpus. In *Annual Conference of the International Speech Communication Association.*

Kavitskaya, D., Iskarous, K., Noiray, A., & Proctor, M. (2009). Trills and palatalization: Consequences for sound change. In J. Reich, M. Babyonyshev, & D. Kavitskaya (Eds.), *Proceedings of the formal approaches to Slavic Linguistics* (pp. 97–110). Michigan Slavic Publications.

Kaye, J. (1995). Derivations and interfaces. In *Frontiers of phonology: Atoms, structures, derivations* (pp. 289–332). Addison-Wesley Longman Limited.

Keating, P. A. (1983). Comments on the jaw and syllable structure. *Journal of Phonetics*, *11*(4), 401–406.

Kenstowicz, M. (2004). Quality-sensitive stress. *Optimality Theory in Phonology: A Reader*, 191–201.

Kenstowicz, M., & Rubach, J. (1987). The phonology of syllabic nuclei in Slovak. *Language*, 463–497.

Kisler, T., Schiel, F., & Sloetjes, H. (2012). Signal processing via web services: the use case WebMAUS. In *Digital humanities conference* (pp. 30–34). Hamburg, Germany.

Kohler, K. J. (1990). Segmental reduction in connected speech in German: Phonological facts and phonetic explanations. *Speech production and speech modelling*, *55*, 69–92.

Krakow, R. A. (1999). Physiological organization of syllables: a review. *Journal of Phonetics*, *27*(1), 23–54.

Král', Á. (2005). *Pravidlá slovenskej výslovnosti: systematika a ortoepický slovník.* Matica slovenská.

Král', Á., & Sabol, J. (1989). *Fonetika a fonológia.* Slovenské pedagogické nakladatel'stvo.

Ladd, D. R., Mennen, I., & Schepman, A. (2000). Phonological conditioning of peak alignment in rising pitch accents in Dutch. *The Journal of the Acoustical Society of America*, *107*(5), 2685–2696.

Ladefoged, P. (2001). *Vowels and consonants: an introduction to the sounds of languages.* Malden, Mass: Blackwell.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages.* Blackwell Publishing.

Lehiste, I. (1964). *Acoustical characteristics of selected English consonants* (C. Voegelin, Ed.). Mouton & Co., The Hauge: Indiana University, Bloomington.

Lin, S. (2011). *Production and perception of prosodically varying inter-gestural timing in American English laterals.* (Doctoral dissertation). The University of Michigan.

Lin, S., Beddor, P. S., & Coetzee, A. W. (2014). Gestural reduction, lexical frequency, and sound change: A study of post-vocalic /l/. *Laboratory Phonology*, *5*(1), 9–36.

Lindblom, B. r., Sussman, H. M., Modarresi, G., & Burlingame, E. (2002). The trough effect: implications for speech motor programming. *Phonetica*, *59*(4), 245–262.

Lisker, L. (1995). English /w, j/: Frictionless approximants or vowels out of place. *Producing speech: Contemporary issues. For Katherine Safford Harris*, 129–142.

Lowenstamm, J. (1996). CV as the only syllable type. In J. Durand & B. Laks (Eds.), *Current trends in phonology: Models and methods* (pp. 419–441).

Marin, S. (2013). The temporal organization of complex onsets and codas in romanian: A gestural approach. *Journal of Phonetics*, *41*(3), 211–227.

Marin, S., & Pouplier, M. (2010). Temporal organization of complex onsets and codas in American English: Testing the predictions of a gestural coupling model. *Motor Control*, *14*(3), 380–407.

Moosmüller, S., Schmid, C., & Kasess, C. H. (2016). Alveolar and velarized laterals in Albanian and in the Viennese dialect. *Language and Speech*, *59*(4), 488–515.

Mücke, D., & Grice, M. (2014). The effect of focus marking on supralaryngeal articulation – is it mediated by accentuation? *Journal of Phonetics*, *44*(Supplement C), 47–61. (Dynamics of Articulation and Prosodic Structure)

Nam, H., Goldstein, L., & Saltzman, E. (2009). Self-organization of syllable structure: A coupled oscillator model. *Approaches to phonological complexity*, 299–328.

Narayanan, S. S., Alwan, A. A., & Haker, K. (1997). Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data. part I. the laterals. *The Journal of the Acoustical Society of America*, *101*(2), 1064-1077.

Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: Differences between consonants and vowels. *Cognition*, *98*(1), 13–30.

Nespor, M., Peña, M., & Mehler, J. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue e linguaggio*, *2*(2), 203–230.

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

Parker, S. G. (2002). *Quantifying the sonority hierarchy* (Doctoral dissertation). University of Massachusetts Amherst.

Peters, S. (2015). *The effects of syllable structure on consonantal timing and vowel compression in child and adult speakers of German* (Doctoral dissertation). Ludwig-Maximilians-Universität, Munich.

Pierrehumbert, J., & Talkin, D. (1992). Lenition of /h/ and glottal stop. *Papers in Laboratory phonology II: Gesture, segment, prosody*, 90–117.

Pouplier, M., & Beňuš, Š. (2011). On the phonetic status of syllabic consonants: Evidence from Slovak. *Laboratory Phonology*, *2*(2).

Pouplier, M., Cederbaum, J., Hoole, P., Marin, S., & Greven, S. (2017). Mixed model for irregularly sampled and correlated functional data: speech science applications. *The Journal of the Acoustical Society of America*, *142*(2), 935–946.

Prieto, P. (2005). Stability effects in tonal clash contexts in catalan. *Journal of Phonetics*, *33*(2), 215–242.

Proctor, M., & Walker, R. (2012). Articulatory bases of sonority in english liquids. *The sonority controversy*, 285–312.

R Core Team. (2015). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Recasens, D. (2004). Darkness in [l] as a scalar phonetic property: implications for phonology and articulatory control. *Clinical Linguistics & Phonetics*, *18*(6-8), 593–603.

Recasens, D. (2012). A cross-language acoustic study of initial and final allophones of /l/. *Speech Communication*, *54*(3), 368–383.

Recasens, D. (2013). Coarticulation in catalan dark [l] and the alveolar trill: General implications for sound change. *Language and Speech*, *56*(1), 45–68.

Recasens, D. (2016). What is and what is not an articulatory gesture in speech production: The case of lateral, rhotic and (alveolo)palatal consonants. *Gradus*, *1*(1), 23–42.

Ridouane, R. (2008). Syllables without vowels: phonetic and phonological evidence from Tashlhiyt Berber. *Phonology*, *25*(2), 321–359.

Ridouane, R., & Fougeron, C. (2011). Schwa elements in Tashlhiyt word-initial clusters. *Laboratory Phonology*, *2*(2), 275–300.

Roach, P., Sergeant, P., & Miller, D. (1992). Syllabic consonants at different speaking rates: A problem for automatic speech recognition. *Speech Communication*, *11*(4-5), 475–479.

Roettger, T. B. (2017). *Tonal placement in Tashlhiyt: How an intonation system accommodates to adverse phonological environments.* Berlin: Language Science Press.

Rose, S., & Walker, R. (2004). A typology of consonant agreement as correspondence. *Language*, *80*(3), 475–531.

Rubach, J. (1993). *The lexical phonology of Slovak.* Clarendon Press.

Rusko, M., Sabo, R., & Dzúr, M. (2007). Sk-ToBI scheme for phonological prosody annotation in Slovak. In V. Matoušek & P. Mautner (Eds.), *Text, speech and dialogue: 10th international conference* (pp. 334–341). Springer.

Sawicka, I. (2001). *An outline of the phonetic typology of the Slavic languages.* Torun: Wydawn. Uniwersytetu M. Kopernika.

Scheer, T. (2008). Syllabic and trapped consonants in (Western) Slavic: The same but yet different. In G. Zybatow, L. Szucsich, U. Junghanns, & R. Meyer (Eds.), *Formal description of Slavic languages: The fifth conference* (pp. 149–168). Frankfurt am Main.

Scobbie, J. M., & Pouplier, M. (2010). The role of syllable structure in external sandhi: An EPG study of vocalisation and retraction in word-final English /l/. *Journal of Phonetics*, *38*(2), 240–259.

Shaw, J., Gafos, A., Hoole, P., & Zeroual, C. (2009). Temporal evidence for syllabic structure in Moroccan Arabic: Data and model. *Phonology*, *26*(1), 187–215.

Siddins, J., Harrington, J., Reubold, U., & Kleber, F. (2014). Investigating the relationship between accentuation, vowel tensity and compensatory shortening. In *7th Speech Prosody Conference.* Dublin, Ireland.

Silverman, K. E. A. (1987). *The structure and processing of fundamental frequency contours* (Doctoral dissertation). Cambridge University, Cambridge, UK.

Šimko, J., Beňuš, Š., & Vainio, M. (2016). Hyperarticulation in Lombard speech: Global coordination of the jaw, lips and the tongue. *The Journal of the Acoustical Society of America*, *139*(1), 151–162.

Simons, G. F., & Fennig, C. D. (2017). *Ethnologue: Languages of the World, Twentieth edition.* `http://www.ethnologue.com`. Dallas, Texas: SIL International.

Smith, C., & Lammert, A. C. (2013). Identifying consonantal tasks via measures of tongue shaping: a real-time MRI investigation of the production of vocalized syllabic /l/ in American English. In *Annual Conference of the International Speech Communication Association* (pp. 3230–3233).

Sproat, R., & Fujimura, O. (1993). Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics*, *21*(3), 291–311.

Stevens, K. N. (1997). Articulatory-acoustic-auditory relationships. In W. Hardcastle & J. Laver (Eds.), *The handbook of phonetic sciences* (p. 462-506). Blackwell.

Stoll, G. (1984). Pitch of vowels: Experimental and theoretical investigation of its dependence on vowel quality. *Speech Communication*, *3*(2), 137–147.

Stone, M. (2005). A guide to analysing tongue motion from ultrasound images. *Clinical linguistics & phonetics*, *19*(6-7), 455–501.

Swadesh, M. (1947). On the analysis of English syllabics. *Language*, *23*(2), 137–150.

Toft, Z. (2002). The phonetics and phonology of some syllabic consonants in Southern British English. *ZAS Papers in Linguistics*, *28*, 111–144.

Toro, J. M., Nespor, M., Mehler, J., & Bonatti, L. L. (2008). Finding words and rules in a speech stream. *Psychological Science*, *19*(2), 137–144.

van der Hulst, H., & van de Weijer, J. (1995). Vowel harmony. In J. A. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 495–534). Cambridge: Blackwell.

Vazquez-Alvarez, Y., & Hewlett, N. (2007). The 'trough effect': an ultrasound study. *Phonetica*, *64*(2-3), 105–121.

Wells, J. C. (1965). The phonological status of syllabic consonants in English RP. *Phonetica*, *13*(1-2), 110–113.

Whalen, D., & Levitt, A. G. (1995). The universality of intrinsic F0 of vowels. *Journal of Phonetics*, *23*, 349-366.

Wiese, R. (2000). *The phonology of German.* Oxford University Press.

Winkelmann, R., Jaensch, K., Cassidy, S., & Harrington, J. (2016). emuR: Main package of the EMU speech database management system [Computer software manual]. (R package version 0.2.0)

Zhang, J. (2004). The role of contrast-specific and language-specific phonetics in contour tone distribution. In B. Hayes, R. Kirchner, & D. Steriade (Eds.), (pp. 157–190). Cambridge University Press.

Acknowledgements

I am extremely grateful to Marianne Pouplier for her extreme patience. Her guidance and support were essential for the accomplishment of this thesis. Thanks for allowing me the time to explore but always being there to help me understand the data and find the big picture.

I also thank Phil Hoole who always had an answer to my questions regarding the articulatory data. Štefan Beňuš helped setting up the experiment with his expertise on Slovak and assisted me greatly especially during the initial phase. To Anna Ratzinger and Dustin Young I am grateful for their help in acoustic and articulatory labelling and to Susanne Waltl for organizing the recording lab! I am grateful to the people from the IPS and guests for their valuable feedback during the PhD meetings. Great thanks especially to Jessica Siddins for proof reading, the cake eating, and everything else. I will never forget the big pot of chili by Raphael Winkelmann that nourished me during the last week. I felt lucky having been surrounded by many great people at the institute with whom I spent time also before or after conferences, who told me jokes at the office or were just great support. Thanks to Nikola Eger, Miquel Llompart, Manfred Pastätter, Thomas Kisler, Tom Lentz, Stefania Marin. Special thanks also to Uwe Reichel for giving me the first job at the IPS and motivating me throughout the years.

Without my friends who kept me up to date with the world outside phonetics, especially Nico and Danijel during the regular lunch meetings, I would not have enjoyed the past years half as much. Finally, thanks to my family, especially my grand mother who always let me know that she was proud of me studying.